Lecture Notes in Electrical Engineering 499

Jiadong Sun · Changfeng Yang
Shuren Guo
Editors

China Satellite Navigation Conference (CSNC) 2018 Proceedings

Volume III

Springer

Lecture Notes in Electrical Engineering

Volume 499

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

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
- 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.

For general information about this series, comments or suggestions, please use the contact address under "service for this series".

To submit a proposal or request further information, please contact the appropriate Springer Publishing Editors:

Asia:

China, *Jessie Guo, Assistant Editor* (jessie.guo@springer.com) (Engineering)

India, *Swati Meherishi, Senior Editor* (swati.meherishi@springer.com) (Engineering)

Japan, *Takeyuki Yonezawa, Editorial Director* (takeyuki.yonezawa@springer.com) (Physical Sciences & Engineering)

South Korea, *Smith (Ahram) Chae, Associate Editor* (smith.chae@springer.com) (Physical Sciences & Engineering)

Southeast Asia, *Ramesh Premnath, Editor* (ramesh.premnath@springer.com) (Electrical Engineering)

South Asia, *Aninda Bose, Editor* (aninda.bose@springer.com) (Electrical Engineering)

Europe:

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

(Applied Sciences and Engineering; Bio-Inspired Robotics, Medical Robotics, Bioengineering; Computational Methods & Models in Science, Medicine and Technology; Soft Computing; Philosophy of Modern Science and Technologies; Mechanical Engineering; Ocean and Naval Engineering; Water Management & Technology)

(christoph.baumann@springer.com)

(Heat and Mass Transfer, Signal Processing and Telecommunications, and Solid and Fluid Mechanics, and Engineering Materials)

North America:

Michael Luby, Editor (michael.luby@springer.com) (Mechanics; Materials)

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

Jiadong Sun · Changfeng Yang
Shuren Guo
Editors

China Satellite Navigation Conference (CSNC) 2018 Proceedings

Volume III

Springer

Editors
Jiadong Sun
China Aerospace Science
and Technology Corporation
Beijing
China

Shuren Guo
China Satellite Navigation
Engineering Center
Beijing
China

Changfeng Yang
China Satellite Navigation
Engineering Center
Beijing
China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-13-0028-8 ISBN 978-981-13-0029-5 (eBook)
https://doi.org/10.1007/978-981-13-0029-5

Library of Congress Control Number: 2018938375

Printed on acid-free paper

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

Editorial Board

Vice-Chairman

Yanhong Kou, Beihang University, Beijing, China
Zheng Yao, Tsinghua University, Beijing, China

Topic: S4: Satellite Orbit and Satellite Clock Error

Chairman

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

Vice-Chairman

Rongzhi Zhang, Xi'an Satellite Control Center, Shaanxi, China
Geshi Tang, Beijing Aerospace Control Center, Beijing, China
Maorong Ge, Geo Forschungs Zentrum (GFZ), Potsdam, Germany

Topic: S5: Precise Positioning Technology

Chairman

Qile Zhao, Wuhan University, Hubei, China

Vice-Chairman

Jianwen Li, The PLA Information Engineering University, Henan, China
Shuli Song, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China
Yanming Feng, Queensland University of Technology, Brisbane, Australia

Topic: S6: Time–Space Datum and Temporal Frequency Technologies

Chairman

Lianshan Gao, The 203th Research Institute of China Aerospace Science & Industry Corporation, Beijing, China

Vice-Chairman

Chunhao Han, Beijing Satellite Navigation Center, Beijing, China
Xiaohui Li, National Time Service Center, Chinese Academy of Sciences, Shaanxi, China
Nikolay Demidov, Vremya CH Corporation, Novgorod, Russia

Topic: S7: Satellite Navigation Augmentation Technology

Chairman

Junlin Zhang, OLinkStar Co., Ltd., Beijing, China

Vice-Chairman

Jinping Chen, Beijing Satellite Navigation Center, Beijing, China
Rui Li, Beihang University, Beijing, China
Shaojun Feng, Imperial College London Qianxun Positioning Network, Co., Ltd., Shanghai, China
Dun Wang, Space Star Aerospace Technology Applications Co., Ltd., Heilongjiang, China

Topic: S8: Test and Assessment Technology

Chairman

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

Vice-Chairman

Jun Yang, National University of Defense Technology, Hunan, China
Jianguang Qu, Heilongjiang Institute of Technology, Heilongjiang, China
Yang Gao, University of Calgary, Alberta, Canada

Topic: S9: User Terminal Technology

Chairman

Haibo He, Beijing Satellite Navigation Center, Beijing, China

Vice-Chairman

Baowang Lian, Northwestern Polytechnic University, Shaanxi, China
Hong Li, Tsinghua University, Beijing, China

Topic: S10: Multi-Source Fusion Navigation Technology

Chairman

Zhongliang Deng, Beijing University of Posts and Telecommunications, Beijing, China

Vice-Chairman

Hong Yuan, Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China
Yongbin Zhou, National University of Defense Technology
Jinling Wang, University of New South Wales, Australia

Topic: S11: PNT New Concept, New Methods and New Technology

Chairman

Mingquan Lu, Tsinghua University, Beijing, China

Vice-Chairman

Wei Wang, The 20th Research Institute of China Electronics Technology Group Corporation, Shaanxi, China
Ying Xu, Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China
Sang Jeong Lee, Chungnam National University, Daejeon, South Korea

Topic: S12: Policies and Regulations, Standards and Intellectual Properties

Chairman

Junlin Yang, Beihang University, Beijing, China

Vice-Chairman

Daiping Zhang, China Defense Science and Technology Information Center, Beijing, China
Yonggang Wei, China Academy of Aerospace Standardization and Product Assurance, Beijing, China
Huiying Li, Electronic Intellectual Property Center, Ministry of Industry and Information Technology, PRC Beijing, 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 global positioning system (GPS) and compatible with other global satellite navigation systems (GNSSs) worldwide. BDS will provide highly reliable and precise positioning, navigation and timing (PNT) services and 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 9th China Satellite Navigation Conference (CSNC 2018) is held during May 23–25, 2018, Harbin, China. The theme of CSNC2018 is Location, Time of Augmentation, including technical seminars, academic exchanges, forums, exhibitions, and lectures. The main topics are as followed:

Conference Topics

S1 Satellite Navigation Technology
S2 Navigation and Location Service
S3 Satellite Navigation Signal and Anti-Jamming Technologies
S4 Satellite Orbit and Satellite Clock Error
S5 Precise Positioning Technology
S6 Time–Space Datum and Temporal Frequency Technologies
S7 Satellite Navigation Augmentation Technology
S8 Test and Assessment Technology
S9 User Terminal Technology
S10 Multi-Source Fusion Navigation Technology
S11 PNT New Concept, New Methods and New Technology
S12 Policies and Regulations, Standards and Intellectual Properties

The proceedings have 208 papers in twelve topics of the conference, which were selected through a strict peer-review process from 588 papers presented at CSNC2018. In addition, another 274 papers were selected as the electronic proceedings of CSNC2018, 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 279 referees and 55 session chairmen who are listed as members of editorial board. The assistance of CNSC2018's organizing committees and the Springer editorial office is highly appreciated.

Beijing, China

Jiadong Sun
Changfeng Yang
Shuren Guo

Scientific Committee

Chairman

Jiadong Sun, China Aerospace Science and Technology Corporation, Beijing, China

Vice-Chairman

Rongjun Shen, China Satellite Navigation System Committee, Beijing, China
Jisheng Li, China Satellite Navigation System Committee, Beijing, China
Qisheng Sui, China Satellite Navigation System Committee, Beijing, China
Changfeng Yang, China Satellite Navigation System Committee, Beijing, China
Zuhong Li, China Academy of Space Technology, Beijing, China
Shusen Tan, Beijing Satellite Navigation Center, Beijing, China

Executive Chairman

Jingnan Liu, Wuhan University, Hubei, China
Yuanxi Yang, China National Administration of GNSS and Applications, Beijing, China
Shiwei Fan, China Satellite Navigation Engineering Center, Beijing, China
Jun Xie, China Academy of Space Technology, Beijing, China

Committee Members: (By Surnames Stroke Order)

Xiancheng Ding, China Electronics Technology Group Corporation, Beijing, China
Qingjun Bu, China National Administration of GNSS and Applications, Beijing, China
Weixing Wan, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China
Liheng Wang, China Aerospace Science and Technology Corporation, Beijing, China
Yuzhu Wang, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai, China

Guoxiang Ai, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China
Shuhua Ye, Shanghai Astronomical Observatories, Chinese Academy of Sciences, Shanghai, China
Daren Lv, The Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
Yongcai Liu, China Aerospace Science & Industry Corporation, Beijing, China
Zhaowen Zhuang, National University of Defense Technology, Hunan, China
Qifeng Xu, PLA Information Engineering University, Henan, China
Houze Xu, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Hubei, China
Tianchu Li, National Institute of Metrology, Beijing, China
Jiancheng Li, Wuhan University, Hubei, China
Guirong Min, China Academy of Space Technology, Beijing, China
Xixiang Zhang, The 29th Research Institute of China Electronics Technology Group Corporation, Sichuan, China
Lvqian Zhang, China Aerospace Science and Technology Corporation, Beijing, China
Junyong Chen, National Administration of Surveying, Mapping and Geo-information, Beijing, China
Benyao Fan, China Academy of Space Technology, Beijing, China
Dongjin Luo, China People's Liberation Army, Beijing, China
Guohong Xia, China Aerospace Science & Industry Corporation, Beijing, China
Shuren Guo, China Satellite Navigation Engineering Center, Beijing, China
Peikang Huang, China Aerospace Science & Industry Corporation, Beijing, China
Chong Cao, China Research Institute of Radio Wave Propagation (CETC 22), Beijing, China
Faren Qi, China Academy of Space Technology, Beijing, China
Rongsheng Su, China People's Liberation Army, Beijing, China
Ziqing Wei, Xi'an Institute of Surveying and Mapping, Shaanxi, China

Executive Members: (By Surnames Stroke Order)

Qin Zhang, Chang'an University, Shaanxi, China
Yamin Dang, Chinese Academy of Surveying & Mapping, Beijing, China
Xiaochun Lu, National Time Service Center, Chinese Academy of Sciences, Shaanxi, China
Xiaogong Hu, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China
Qile Zhao, Wuhan University, Hubei, China
Lianshan Gao, The 203th Research Institute of China Aerospace Science & Industry Corporation, Beijing, China
Junlin Zhang, OLinkStar Co., Ltd., Beijing, China

Xiaolin Jia, Xi'an Institute of Surveying and Mapping, Shaanxi, China
Haibo He, Beijing Satellite Navigation Center, Beijing, China
Zhongliang Deng, Beijing University of Posts and Telecommunications, Beijing, China
Mingquan Lu, Tsinghua University, Beijing, China
Junlin Yang, Beihang University, Beijing, China

Organizing Committee

Secretary-General

Haitao Wu, Satellite Navigation Headquarters, Chinese Academy of Sciences, Beijing, China

Executive Deputy Secretary-General

Weina Hao, Navigation Headquarters, Chinese Academy of Sciences, Beijing, China

Deputy Secretary-General

Wenhai Jiao, China Satellite Navigation Engineering Center, Beijing, China
Wenjun Zhao, Beijing Satellite Navigation Center, Beijing, China
Tonghuang Wang, Space Star Aerospace Technology Applications Co., Ltd., Heilongjiang, China
Bo Wang, Academic Exchange Center, China Satellite Navigation Office, Beijing, China

Committee Members: (By Surnames Stroke Order)

Qun Ding, The 20th Research Institute of China Electronics Technology Group Corporation, Shaanxi, China
Guang Yu, Harbin Institute of Technology, Heilongjiang, China
Li Wang, International Cooperation Research Center, China Satellite Navigation Office, Beijing, China
Chunsheng Wang, Haidian Investment Promotion Bureau, Beijing, China
Ying Liu, China Satellite Navigation Engineering Center, Beijing, China
Wanming Yang, Administrative Committee of Nanjing New & High Technology Industry Development Zone, Jiangsu, China
Jun Shen, Beijing UniStrong Science & Technology Co., Ltd., Beijing, China
Liquan Song, Heilongjiang Institute of Technology, Heilongjiang, China
Mingquan Lu, Tsinghua University, Beijing, China

Lu Chen, Beijing Institute of Space Science and Technology Information, Beijing, China
Xiuwan Chen, Peking University, Beijing, China
Gang Hu, Beijing Unicore Communications, Inc., Beijing, China
Qile Zhao, Wuhan University, Hubei, China
Min Shui, The National Remote Sensing Center of China, Beijing, China

Contents

Part I
Test and Assessment Technology

Accuracy Assessment of Broadcast Ephemerides for Quasi-Zenith Satellite System

Xuying Ma, Chengpan Tang, Xingyu Wang, Chuang Jin and Xiaping Ma

Abstract In this paper, we presented a method for assessing the broadcast ephemerides accuracy of QZSS by comparing the measured broadcast ephemerides with precise orbit and clock products provided by MGEX. By using this method, we carried out assessment of broadcast ephemerides for QZS-1 satellite and also compared the results with that of GPS system during the same data collection duration. The results show that: the broadcast ephemerides of QZSS can achieve an orbit accuracy of 0.42, 1.33 and 0.89 m in the direction of radial, along-track and cross-track and a clock accuracy of 1.87 ns respectively; the SISRE-orb and SISRE-orb&clk can achieve accuracies of 0.48 and 0.55 m respectively; currently the accuracy of broadcast ephemerides of QZSS still lags behind that of GPS system. The results presented in this paper may provide some reference information for multi-GNSS application in the near future.

Keywords QZSS · Broadcast ephemerides · Precise products · Accuracy analysis

X. Ma (✉) · C. Jin
University of Auckland, 23 Symonds Street, Auckland 1010, New Zealand
e-mail: maxuying10086@163.com

C. Jin
e-mail: cjin987@aucklanduni.ac.nz

C. Tang
Shanghai Astronomy Observatory, Chinese Academy of Science,
No. 8 NanDan Road, 200030 Shanghai, China
e-mail: cptang@shao.ac.cn

X. Wang
Xi'an SiYuan University, No. 28 ShuiAn Road, 710038 Xi'an, China
e-mail: 809914450@qq.com

X. Ma
Xi'an University of Science and Technology, No. 58 YanTa Road,
710054 Xi'an, China
e-mail: celiang0321@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_1

1 Introduction

GNSS (Global Navigation Satellite System) has been widely used in many fields of research study and industry [1]. More and more countries begin to launch their own global or regional navigation systems with realizing its great potential in the future. Quasi-Zenith Satellite System (QZSS) is the latest regional satellite navigation and augmentation system launched by Japan Aerospace Exploration Agency (JAXA) [2]. The original intention of launching QZSS is to build a regional system covering Japan and its surrounding area, which it could provide users with not only better usability but also augmentation service of multi-GNSS [3]. Therefore, QZSS has a high level of compatibility and interoperability with GPS and other GNSS systems [4]. Currently, there are four Quasi-Zenith satellites (QZS) in orbit. However, there was only one satellite providing formal service at the time of writing down this paper. This first satellite QZS-1 'Michibiki' was launched on 11th Sep, 2010; its PRN is 193 and shown as J01 in RINEX observation file [5]. The accuracy of broadcast ephemeris is a key factor influencing the performance of a navigation system since currently most real-time navigation and positioning services are based on information transmitted by broadcast ephemeris. Although on current stage QZSS still cannot provide stand-alone navigation and positioning service, its good compatibility and interoperability has created a golden opportunity for multi-GNSS applications. Assessing and knowing the accuracy of QZSS's broadcast ephemeris would lay a solid foundation for further multi-GNSS combination in the near future. There are just handful documents about the topic of QZSS's broadcast ephemeris accuracy as this system is still under construction. Kishimoto et al. assessed the accuracy of broadcast ephemeris of QZS-1 by using data of seven months from Jun, 2011 to Jan, 2012 and got a conclusion that the SISRE (Signal-In-Space Range Error) of QZSS was about 0.4 m. However, the method of assessment and procedure details were not given in this paper [6]. Wang and Kuang assessed the accuracy of LEX augmentation products of QZS-1 based on the precise ephemeris of JAXA and made a conclusion that the accuracy of LEX real-time orbit and clock products are better than 11 cm and 0.56 ns respectively. However, the accuracy of broadcast ephemeris of QZS-1 is not analysed in this paper [7]. In this contribution, first, the orbital period of satellite QZS-1 is analysed using measured QZSS data of broadcast ephemeris; and then, assessment of the accuracy of QZS-1's broadcast ephemeris is carried out using the precise orbit and clock products from MGEX as a reference. The results of QZSS are compared with that of GPS system during the same data collection duration. The result of this study could give some useful information for future multi-GNSS applications.

2 Methodology

2.1 Satellite orbital period

According to Kepler's law of motion of satellite orbit, the orbital period of one satellite can be calculated if satellite orbit parameters of major semi-axis and mean angular velocity are already known. These parameters are exactly contained in broadcast ephemeris message, therefore, the orbital period of one satellite can be calculated as the following formulas with broadcast ephemeris information [8]:

$$\begin{cases} n = \sqrt{G \cdot M}/a^{\frac{3}{2}} + \Delta n \\ T_0 = 2\pi/n \end{cases} \tag{1}$$

In formula (1), Δn represents the correction of mean angular velocity; G represents the gravitational constant; M represents the mass of the earth; and $G \cdot M = 398,600.4418 \times 109\ \mathrm{m}^3/\mathrm{s}^2$; a represents the square root of major semi-axis; T_0 represents the time of orbiting around the earth of a satellite; T_R represents the orbital period of the satellite.

2.2 The Calculation Method of SISRE

The SISRE (Signal-In-Space Range Error) is a value that it reflects the influence of the accuracy of broadcast ephemeris orbit and clock products on the precision of pseudo range measurements [9]. In general, the accuracy of SISRE is affected by factors from both space and control segments such as orbit motion, clock stability of satellite, precision of orbit determination, the quality of uploading information, and the distribution of monitoring stations. Therefore, SISRE is a vital index for estimating the performance of one GNSS system. Currently, the constellation of QZSS consists of three IGSO satellites and one GEO satellite. According to formula derivation processes presented in Refs. [9, 10], the SISRE of QZS-1 can be expressed as [10]:

$$\begin{cases} SISRE(orb) = \sqrt{[0.99 \cdot RMS(dR)]^2 + \dfrac{1}{126} \cdot [RMS^2(dA) + RMS^2(dC)]} \\ SISRE = \sqrt{RMS^2(0.99 \cdot dR - \Delta c \cdot dT) + \dfrac{1}{126} \cdot [RMS^2(dA) + RMS^2(dC)]} \end{cases} \tag{2}$$

In formula (2), $SISRE(orb)$ represents the SISRE just considering the influence of broadcast ephemeris orbit error; $SISRE$ represents the situation considering the influence of both broadcast ephemeris orbit and clock errors. dR, dA and dC represent orbit errors in the directions of radial, along-track and cross-track respectively. $\Delta c \cdot dT$ denotes the clock error correction of broadcast ephemeris.

2.3 Important Notes in Data Processing

The compatibility of orbit and clock products between broadcast ephemeris and precise ephemeris should be considered while doing a comparison. This section will discuss this issue from aspects of coordinate reference systems, time systems, and relative error corrections.

The coordinate reference system of QZSS is called JGS (Japan satellite navigation Geodetic System) [2]. The difference within coordinate reference system between JGS and WGS-84 of GPS is less than 2 cm. The precise ephemeris of QZSS provided by JAXA takes the ITRF08 as its coordinate reference frame which its difference between JGS or WGS-84 is at the level of several centimetre. On current stage, the accuracy of broadcast ephemeris for each GNSS system is about meter level. We can just ignore the difference of coordinate reference while doing comparisons between the broadcast ephemeris and the precise ephemeris since their difference is about two orders of magnitude [11].

The time system of QZSS is called QZSST, which its one second span is the same as the International Atomic Time TAI. QZSST lags behind TAI by 19 s, which is the same with that of GPST. In the aspect of interface with GPST, the satellite borne atomic clock of QZSS is controlled by GPST, which is the same with that of GPS [2].

The difference of satellite reference centre for different orbit products should be considered while doing a comparison between broadcast ephemeris and precise ephemeris. The precise ephemeris of QZSS calculated by JAXA from the MGEX is based on a satellite reference centre of CoM (Center-of-Mass); while satellite positions in the broadcast ephemeris of QZSS is calculated based on a satellite reference centre of APC (Antenna Phase Center). Therefore, this fact also raises an issue of antenna phase center correction. A set of antenna phase center correction parameters are needed to transfer the satellite positions from precise ephemeris into a reference frame of APC to match the standard of broadcast ephemeris. According to parameters provided on the website of MGEX, the antenna phase center offsets of QZS-1 in the direction of x, y and z are 0.00, 0.00 and +3.50 m respectively [9].

Similar to satellite orbit, the time reference difference of two clock products should also be considered while doing a comparison of clock parameters between broadcast ephemeris and precise ephemeris. The time references used in broadcast ephemeris and precise ephemeris are different; moreover, the bias arisen from the difference of two time references is much larger than the precision of clock itself. Therefore, this bias must be dealt with before further comparison and analysis of

different clock products. For general navigation constellation, this bias affects each satellite in one constellation in the same way and can be treated as a systematic bias. Usually, this bias can be mitigated by applying the following method: first, an ensemble clock bias is calculated at each epoch by averaging results of 'broadcast ephemeris clock—precise ephemeris clock' from all satellites in the constellation; and then, each satellite clock is corrected by this ensemble clock bias [9]. However, this common method cannot be applied to QZSS as there is only one single satellite providing formal service in its constellation. In this study, a method of averaging the long-term time series of 'broadcast ephemeris clock—precise ephemeris clock' is used to calculate the ensemble clock bias. For instance, one month data is used to calculate a monthly ensemble clock bias; and then, corrected clock at each epoch can be calculated by subtracting the ensemble clock bias.

3 Data Collection

The website of MGEX can provide both broadcast and precise ephemeris of QZSS. The interval of each broadcast message is 15 min; that means a total of 96 epochs each day. The precise ephemeris provided by MGEX is generated by JAXA. In order to assess the broadcast ephemeris accuracy of QZSS, 8 weeks data of broadcast ephemeris from 21st Oct, 2016 to 15th Dec, 2016 (*doy*295–*doy*350) was downloaded; at the mean while, precise ephemeris data of the same duration was also downloaded.

4 Assessment and Analysis Results

First, the orbital period of satellite orbit of QZS-1 is calculated based on the data of broadcast ephemeris from *doy*29 to *doy*350 according to formula (1). The interval of each broadcast message is 15 min, so there are 96 epochs of orbital period results per 24 h. Figure 1 shows the time series of 5472 epochs orbital period results of QZS-1 from *doy*295 to *doy*350. According to statistics, the average orbital period of satellite QZS-1 during this period is about 86152 s, which it is quite close to the average repeat time of satellite orbit of GPS system (85154 s) released in Ref. [13] by Agnew and Larson. The consistency of these two results also proves the high level of compatibility and interoperability between GPS and QZSS from the aspect of the orbit period. For instance, there is no need to consider the orbit period difference between GPS and QZSS while applying combined sidereal filtering (SF) algorithm because of the excellent compatibility. Figure 1 also indicates that the time series of orbital period results of QZS-1 have an obvious periodic characteristic. Within one day, there is a short periodic trend of first increased and then decreased; within a long term of two weeks, there is a long periodic trend of first increased and then decreased. In addition, the value of orbital period of satellite QZS-1 increases gradually as time goes on; this characteristic is slightly different with that of GPS system.

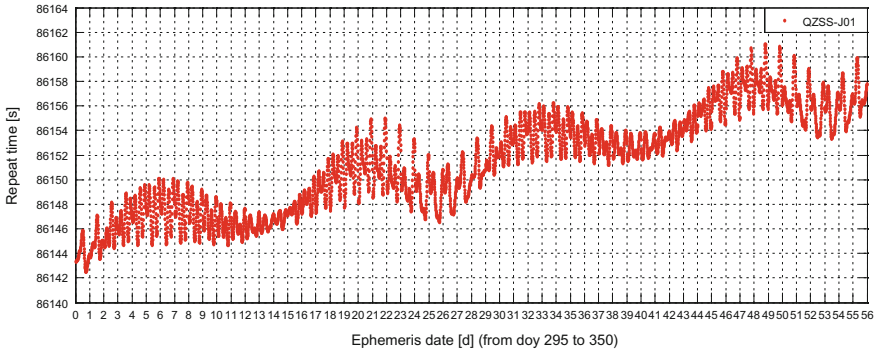


Fig. 1 Time Series of orbital period results for satellite QZS-1

Comparisons of orbit and clock products of QZSS between broadcast and precise ephemeris from *doy*295 to *doy*350 are carried out at an interval of 15 min. In addition, the initial orbit difference results in the direction of x, y and z in the geocentric coordinate system are transformed correspondingly into results in the direction of radial (dR), along-track (dA) and cross-track (dC) in the satellite coordinate system. The corresponding results of GPS system during the same period are also calculated to use as a reference. In Table 1 are RMS statistics of differences of orbit in each direction and clock between broadcast and precise ephemeris from *doy*295 to *doy*350 for different kinds of satellites. The broadcast ephemerides of QZSS can achieve an orbit accuracy of 0.42, 1.33 and 0.89 m in the direction of radial, along-track and cross-track and a clock accuracy of 1.87 ns respectively. At the meanwhile, according to the comparison of broadcast orbit and clock products with GPS-IIA, GPS-IIR and GPS-IIF these three kinds of satellites from GPS system, QZSS still has an obvious gap with GPS in the aspect of broadcast ephemeris accuracy.

The SISRE performance of QZSS broadcast ephemeris is analysed according to formula (2) at an interval of 15 min for data from *doy*295 to *doy*350. In addition, the corresponding SISRE results of GPS system during the same period are also calculated to use as a reference. In Table 2 are SISRE results of different kind of satellites. The SISRE-orb and SISRE-orb&clk of QZS-1 broadcast ephemeris can achieve accuracies of 0.48 and 0.55 m respectively. At the meanwhile, according to the comparison of broadcast ephemeris SISRE accuracy with GPS-IIA, GPS-IIR

Table 1 Comparison results between broadcast and precise products

Satellite type	RMS_dR (m)	RMS_dA (m)	RMS_dC (m)	RMS_dT (ns)
QZS-1	0.42	1.33	0.89	1.87
GPS-IIA	0.27	1.22	0.44	3.55
GPS-IIR	0.16	1.09	0.39	1.74
GPS-IIF	0.15	0.80	0.34	0.96

Table 2 SISRE results of different kind of satellites

Satellite type	SISRE-orb (m)	SISRE-orb&clk (m)
QZS-1	0.48	0.55
GPS-IIA	0.33	1.06
GPS-IIR	0.24	0.54
GPS-IIF	0.19	0.37

and GPS-IIF these three kinds of satellites from GPS system, QZSS still has an obvious gap with GPS in the aspect of broadcast ephemeris SISRE accuracy, especially in the situation of just considering the influence of broadcast ephemeris orbit error (SISRE-orb).

5 Conclusion

QZSS is the latest regional satellite navigation and augmentation system launched by JAXA. The accuracy of the broadcast ephemeris of QZSS system at current stage has been a highlighted issue concerned by GNSS users from all over the world. According to this issue, in this paper we assessed the accuracy of the broadcast ephemeris of QZSS system by comparing with precise orbit and clock products downloaded from the MGEX. The corresponding results of GPS system during the same duration are also calculated as a reference. Some initial conclusions are made as follows:

(1) The average orbital period of satellite QZS-1 is about 86152 s, which is quite close to the repeat time of GPS system (85154 s); the time series of orbital period results of QZS-1 have an obvious periodic characteristic.
(2) The broadcast ephemerides of QZS-1 can achieve an orbit accuracy of 0.42, 1.33 and 0.89 m in the direction of radial, along-track and cross-track and a clock accuracy of 1.87 ns respectively.
(3) The SISRE-orb and SISRE-orb&clk of QZS-1 broadcast ephemeris can achieve accuracies of 0.48 and 0.55 m respectively.
(4) QZSS still has an obvious gap with GPS in the aspect of broadcast ephemeris accuracy.

References

1. Gong Z, Jin C, Zhang C, Zhou Z, Zhang Q, Li H (2017) Temporal and spatial morphological variations along a cross-shore intertidal profile, Jiangsu, China. Cont Shelf Res
2. Ma X, Ma X, Tang C, Wang X (2017) The evaluation of improvement of gps performance combined with quasi-zenith satellite system in Japan Area[C]//China satellite navigation conference. Springer, Singapore, pp 525–532

3. Jaxa (2016) Interface specification for Quasi-Zenith satellite system. In: IS-QZSS_18D (ed), Tokyo, Japan
4. Kishimoto M, Hase H, Matsumoto A et al (2007) QZSS System design and its performance. Proc Nat Tech Meet Inst Navig Ntm 6682(1):405–410
5. Hauschild A, Steigenberger P, Rodriguez-Solano C (2012) Signal, orbit and attitude analysis of Japan's first QZSS satellite Michibiki. GPS Solutions 2012(16):127–133
6. Kishimoto M, Myojin E, Kawate K, Miyoshi M, Kogure S, Noda H (2012) Technical verification status of Quasi-Zenith satellite system. Proc Int Tech Meet Inst Navig
7. Wang G, Kuang C, Wang G (2016) QZSS Michibiki Real-Time satellite orbit and clock data accuracy assessment. J. Geomatics Sci Technol 4:379–382
8. Ma X, Shen Y (2013) BeiDou positioning and multipath analysis for short baselines. In: Sun J, Jiao W, Wu H, Shi C (eds) China satellite navigation conference (CSNC) proceedings. Lecture Notes in Electrical Engineering, vol 243. Springer, Berlin, Heidelberg
9. Montenbruck O, Steigenberger P, Hauschild A (2015) Broadcast versus precise ephemerides: a multi-GNSS perspective. GPS Solutions 19(2):321–333
10. Chen L, Jiao W, Huang X, Geng C, Ai L, Lu L, Hu Z (2013) Study on Signal-In-Space errors calculation method and statistical characterization of Beidou navigation satellite system. In: Sun J, Jiao W, Wu H, Shi C (eds) China satellite navigation conference (CSNC) proceedings. Lecture notes in electrical engineering, vol 243. Springer, Berlin, Heidelberg
11. Guo F, Li X, Zhang X, Wang J (2016) Assessment of precise orbit and clock products for Galileo, BeiDou, and QZSS from IGS Multi-GNSS experiment (MGEX). GPS Solutions 1–12
12. Prange L, Orliac E, Dach R, Arnold D, Beutler G, Schaer S, Jäggi A (2017) CODE's five-system orbit and clock solution—the challenges of multi-GNSS data analysis. J Geodesy 1–16
13. Agnew DC, Larson KM (2007) Finding the repeat times of the GPS constellation. GPS Solutions 11(1):71–76

Research on Prediction and Characteristics of Periodic in BDS/GPS Satellite Clocks

Jie Zheng, Xiangyu Zhu, Meijun Guo, Guang Sun, Xiaolin Jia and Shuangqin Li

Abstract The prediction accuracy of satellite clock bias (SCB) directly affects the performance of navigation system. According to the spectrum analysis method, the periodic characteristics of satellite clocks are analyzed, and using precise SCB data of BDS/GPS satellite the prediction tests are conducted, and then we build the forecast model considering periodic items. Based on this, the influence of the number of periodic items on the result of SCB prediction is analyzed. The effect of SCB prediction between periodic term model and quadratic polynomial model is compared, and the SCB prediction performance of different types of satellites is summarized. Experimental results show that the periodic variation of satellites are significant. The periodic items of BDS are 24, 12 and 8 h, and the periodic items of GPS are 24, 12 and 6 h. The number of periodic items affects the prediction results. The prediction accuracy of BDS using 2–3 main periodic items is the best, and the prediction accuracy of GPS using 3–4 main periodic items is the best. Compared with the quadratic polynomial prediction model, the prediction accuracy of the proposed model with additional periodic items can be improved by 1–6 ns.

Keywords BDS · GPS · Satellite atomic clock · Clock offset prediction

1 Introduction

As an important part of time synchronization, the SCB prediction of satellite-borne atomic clocks is particularly important in satellite navigation systems. The accuracy of SCB prediction directly affects the positioning and timing of satellite navigation system. The study of SCB prediction is beneficial to improve the reliability and accuracy of parameter prediction. With the increasing requirements of satellite

J. Zheng (✉) · X. Zhu · M. Guo · G. Sun · S. Li
Xi'an Aerospace Remote Sensing Data Technology Co., Ltd, Xi'an, Shaanxi, China
e-mail: 40735701@qq.com

X. Jia
Xian Research Institute of Surveying and Mapping, Xi'an, Shaanxi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_2

navigation positioning accuracy, the SCB prediction has become an extremely important work. The domestic and foreign scholars have done a lot of research on clock error forecasting [1–8]. Various forecast models have been proposed including polynomial model [1], gray model [2], time series model [3], wavelet neural network model [4], etc. However, these studies are mainly focused on the satellite-borne clock of GPS system and GLONASS system.

In recent years, China is vigorously developing the BeiDou-2 satellite navigation system. Currently, the BeiDou satellite navigation system in the Asia Pacific region has been completed, and 14 satellites have been networked and can provide all-weather service. With the application of the BeiDou satellite navigation system, Chinese scholars have studied the performance evaluation of the BDS satellite-borne atomic clock and made some achievements [9−11]. However, the research on BDS SCB prediction is relatively rare. Therefore, in this paper, the spectrum analysis method is used to analyze the periodic terms variation of BDS/GPS satellite-borne atomic clock for the problem of different types of SCB prediction, and the forecast model of periodic term is constructed as well as the effect of the number of periodic items on the results of SCB prediction is analysed. Lastly, the prediction results of the periodic term model and the quadratic polynomial model are compared as well as the difference of prediction accuracy between different types of satellites.

2 Algorithm Principle

2.1 Polynomial Model

For the GNSS SCB prediction, the first order model and the second order model are respectively applied to the real-time SCB prediction of broadcast ephemeris in GPS system and GLONASS system, and the order selection of polynomial model mainly depends on the frequency stability and frequency drift characteristics of the atomic clock. In general, cesium atomic clocks are predicted by first order polynomial model, and rubidium atomic clocks are predicted by quadratic polynomial model [8]. The expression of the quadratic polynomial model is:

$$x = a + b\Delta t + c\Delta t^2 + \int_{t_0}^{t_s} f(t)dt \tag{1}$$

where, $\Delta t = t_s - t_0$, t_0 is the reference moment of the satellite clock, x is the observed value of SCB at time t_s. a, b and c are the parameters to be estimated which representing phase, frequency and frequency drift of the reference time t_0, respectively. $\int_{t_0}^{t_s} f(t)dt$ is a random clock error caused by random error of frequency, which can be used to describe its statistical properties [5, 12].

2.2 Spectrum Analysis

The method of spectrum analysis is used to extract the periodic term of SCB. First, the quadratic polynomial is used to fit the SCB of each day to eliminate the trend item of clock error, and then the periodic item of the fitting residual is extracted. In this paper, the Fourier series expansion is used to calculate the spectrum values of each numerical value in the fitted residuals sequence. Its expression is as follows:

$$X(k) = \sum_{n=0}^{N-1} x(n) \mathrm{e}^{-j\frac{2\pi}{N}kn} \tag{2}$$

where, X(k) is the spectrum value of time k, x(n) is the fitting residual sequence of SCB, n is the sequence number of the element in the residual sequence, j is the imaginary unit, e is the constant, N is the number of the residuals. $N = 2^L$, L = 0, 1, …, n is usually required when the solution is solved. The 0 element sometimes need to be increased when the number does not meet the requirements. Through this formula, we can find the spectrum value corresponding to each point in the residual sequence, and finally through the sequence of spectrograms to determine the significant periodic items.

2.3 Spectrum Analysis Model (Periodic Term Model)

The spectrum analysis model is based on the polynomial model, by increasing the periodic term correction on the basis of Eq. (1) the model can be obtained as follows:

$$x = a + b\Delta t + c\Delta t^2 + \sum_{m=0}^{n} \{A_m \cos(2\pi f_m \Delta t) + B_m \sin(2\pi f_m \Delta t)\} + \int_{t_0}^{t_s} f(t)dt \tag{3}$$

where, a, b and c are the model parameters of the clock error, n is the number of the main periodic function, f_m is the frequency of the corresponding periodic term, A_m and B_m are the amplitude of the corresponding periodic term. The model parameters are solved by the least square method, and the clock error is predicted by the calculated parameters.

Figure 1 shows the whole process to analyze and evaluate the periodic characteristics and prediction of the BDS/GPS satellite-borne atomic clock.

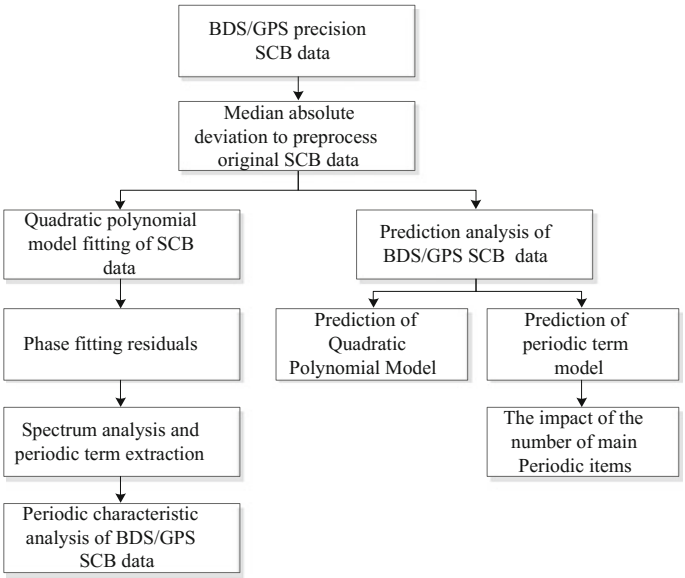


Fig. 1 Evaluation analysis flow chart

3 Experimental and Analytical Results

3.1 Periodic Characteristic Analysis

The periodic characteristics of BDS/GPS satellite clocks is verified and analyzed by the SCB data, which is provided by the Wuhan University with a sampling rate of 300 s. First, the gross error is detected by using the median absolute method (MAD) to preprocess original SCB data. Then, the clock sequence of a single satellite is extracted, and we performed spectrum analysis of the fitted residuals after deduction of quadratic terms and extracted periodic terms. In this paper, the 2016 annual of BDS/GPS SCB data is taken as the research object, and the period of fitting residual is analyzed day by day. Take the C03 (GEO Rb), C07 (IGSO Rb), G11 (IIR Rb), G15 (IIR-M Rb) and G27 (IIF Rb) satellites as example, we get the results of the satellite spectrum analysis of Figs. 2, 3, 4, 5 and 6. As can be seen from these figures, there are significant periodic items in the SCB data, and the satellites of different orbit types show similar periodic characteristics. With the increase of frequency, the amplitude of the corresponding periodic items gradually decreased.

In this paper, the first six items are extracted when extracting periodic items. Tables 1 and 2 give the results of periodic items extraction for all satellites of BDS and GPS, and the corresponding reference amplitude of periodic items.

As can be seen from Tables 1 and 2, the periodic items of different orbital types of BDS/GPS satellites are different, while the periodic items of satellites of the same

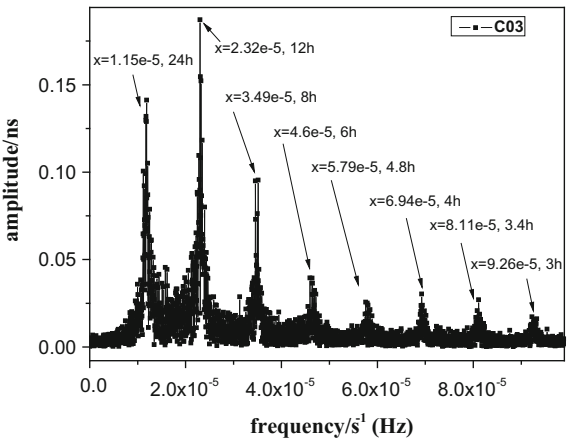


Fig. 2 The spectral analysis results based on SCB data of satellite C03

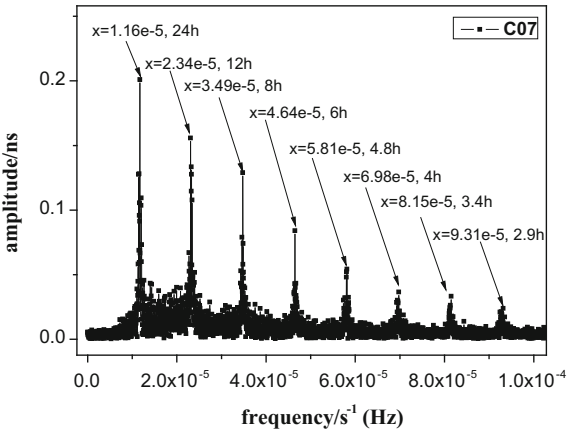


Fig. 3 The spectral analysis results based on SCB data of satellite C07

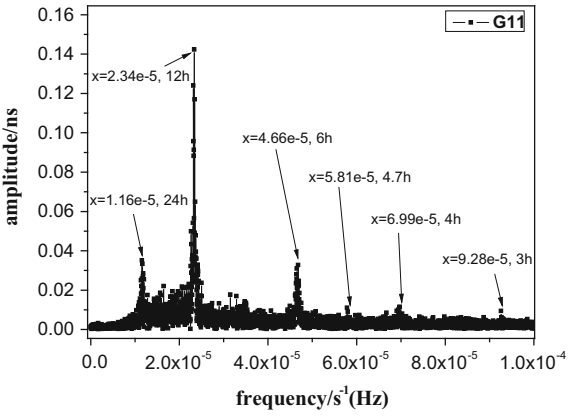


Fig. 4 The spectral analysis results based on SCB data of satellite G11

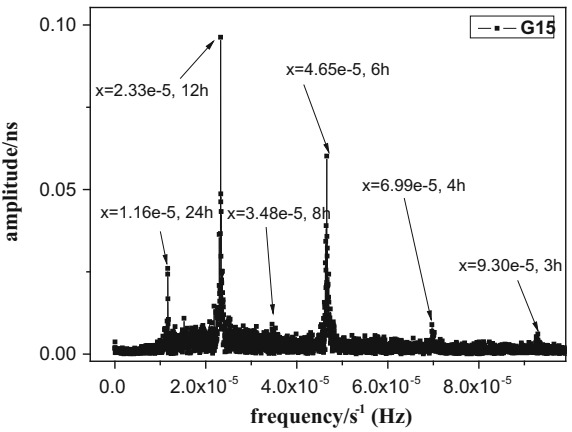


Fig. 5 The spectral analysis results based on SCB data of satellite G15

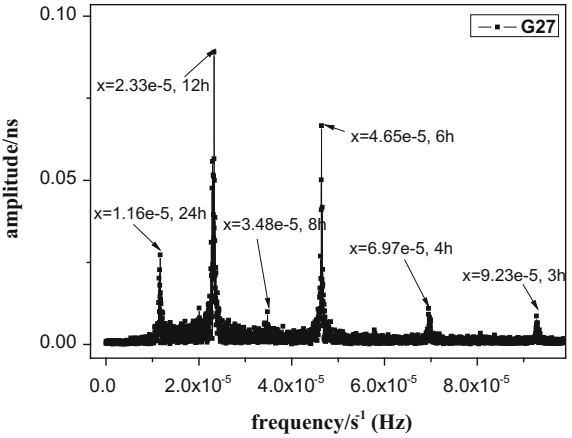


Fig. 6 The spectral analysis results based on SCB data of satellite G27

orbital type are also different. The concrete manifestation is that the two main periodic terms of all BDS IGSO satellites and GEO C02 satellite are 24 and 12 h, and the two main periodic terms of the remaining GEO satellites are 12 and 24 h. The two main period items of the four satellites of MEO are different. The C11 satellite period item is 12.8 and 14.6 h. The C12 satellite period term is 12.8 and 16.4 h. The C13 satellite period term is 21.2 and 16.8 h. The C14 satellite period term is 12.8 and 15.8 h. The two main periodic terms of G17, G02, G14, G20, G25 and G26 satellite are 6 and 12 h. The G24 satellite period item is 12 and 15.3 h. The G11 satellite period term is 12 and 15.3 h. The G24 satellite period term is 12 and 15.3 h. The G24 satellite period term is 12 and 15.8 h. The two main periodic terms of the remaining GPS satellites are 12 and 6 h.

Experimental results show that different types of BDS satellites show more significant 12 and 24 h period terms, and different types of GPS satellites show more significant 12 and 6 h period terms. Obviously, the periodic items of BDS

Table 1 Periodic terms extracted from BDS SCB data (periodic terms listed according to their amplitudes from large to small)

Type	PRN	T1	A (ns)	T2	A (ns)	T3	A (ns)	T4	A (ns)	T5	A (ns)	T6	A (ns)
GEO	C01	12	0.0924	24	0.0712	8	0.0421	6	0.0292	4.8	0.0278	4	0.0260
	C02	24	0.1256	12	0.1191	8	0.0837	6	0.0340	4.8	0.0249	4	0.0198
	C03	12	0.1872	24	0.1412	8	0.0955	6	0.0394	4	0.0303	3.4	0.0270
	C04	12	0.1550	24	0.0849	4.8	0.0502	8	0.0481	6	0.0328	3.4	0.0308
	C05	12	0.2163	24	0.1299	8	0.0534	6	0.0376	4.8	0.0297	3.4	0.0245
IGSO	C06	24	0.5282	12	0.4026	8	0.1962	6	0.1359	4	0.0967	4.8	0.0901
	C07	24	0.2009	12	0.1555	8	0.1286	6	0.0837	4.8	0.0540	4	0.0036
	C08	24	0.1006	12	0.0830	8	0.0490	6.4	0.0383	4.8	0.0354	3	0.0293
	C09	24	0.0979	12	0.0623	8	0.0339	6	0.0272	3.4	0.0255	4.8	0.0211
	C10	24	0.1798	12	0.1178	8	0.0633	6	0.0578	4.8	0.0326	4	0.0333
MEO	C11	12.8	0.0806	14.6	0.0266	6.4	0.0250	20.7	0.0195	8.8	0.0195	28.1	0.0183
	C12	12.8	0.0453	16.4	0.0239	11.1	0.0221	20.2	0.0202	7.6	0.0196	9	0.0193
	C13	21.2	0.5035	16.8	0.4870	15.4	0.4727	12.1	0.4500	13.1	0.4357	11.3	0.3164
	C14	12.8	0.0231	15.8	0.0228	17.1	0.0208	6.4	0.0171	17.9	0.0165	22.9	0.0125

Table 2 Periodic terms extracted from GPS SCB data (periodic terms listed according to their amplitudes from large to small)

Type	PRN	T1	A (ns)	T2	A (ns)	T3	A (ns)	T4	A (ns)	T5	A (ns)	T6	A (ns)
IIR	G02	6	0.0335	12	0.0262	24	0.0132	4	0.0128	3.3	0.0059	3	0.0046
	G11	12	0.1415	24	0.0342	6	0.0307	4	0.0108	4.7	0.0102	3	0.0084
	G13	12	0.0759	6	0.0462	24	0.0225	8	0.0122	3	0.0083	4	0.0070
	G14	6	0.0282	12	0.0214	15.6	0.0186	24	0.0090	4	0.0079	3	0.0055
	G16	12	0.1988	6	0.1468	24	0.0590	4	0.0290	3	0.0144	8	0.1191
	G18	12	0.0868	6	0.0569	24	0.0285	8	0.0111	4	0.0087	3	0.0052
	G19	12	0.0363	6	0.0356	24	0.0191	3	0.0116	8	0.0101	4	0.0097
	G20	6	0.0489	12	0.0283	14.9	0.0200	24	0.0137	8	0.0118	3	0.0063
	G21	12	0.1108	6	0.0592	24	0.0328	8	0.0156	4	0.0132	3	0.0122
	G22	12	0.0633	6	0.0589	24	0.0234	8	0.0111	4	0.0067	3	0.0064
	G23	6	0.0262	13.4	0.0164	8.6	0.0114	24	0.0091	4	0.0090	3	0.0046
	G28	12	0.0480	12.8	0.0391	16.7	0.0360	20.1	0.0304	9.3	0.0236	6	0.0228
IIR-M	G05	12	0.0880	6	0.0722	24	0.0308	4	0.0145	3	0.0084	2.7	0.0055
	G07	12	0.0606	6	0.0452	15.7	0.0209	24	0.0161	8	0.0131	4	0.0096
	G12	12	0.1910	6	0.0894	24	0.0401	8	0.0189	4	0.0158	3	0.0090
	G15	12	0.0961	6	0.0592	24	0.0255	4	0.0090	8	0.0082	3	0.0058
	G17	6	0.0741	12	0.0703	24	0.0356	4	0.0169	8	0.0163	3	0.0105
	G29	12	0.0683	6	0.0470	24	0.0322	7.8	0.0124	4	0.0083	3	0.0077
	G31	12	0.0719	6	0.0716	15.2	0.0261	24	0.0249	3	0.0196	4	0.0074
IIF	G01	12	0.0835	6	0.0390	24	0.0191	4	0.0645	3	0.0035	2.8	0.0028
	G03	12	0.0523	6	0.0229	24	0.0150	4	0.0057	3	0.0032	2.8	0.0031
	G06	12	0.0845	6	0.0361	24	0.0183	4	0.0106	3	0.0043	3.4	0.0036
	G08	12	0.0950	6	0.0910	24	0.0747	4	0.0653	3	0.0638	3.4	0.0428
	G09	12	0.0672	6	0.0642	24	0.0231	4	0.0081	3	0.0054	3.4	0.0033

(continued)

Table 2 (continued)

Type	PRN	T1	A (ns)	T2	A (ns)	T3	A (ns)	T4	A (ns)	T5	A (ns)	T6	A (ns)
	G10	12	0.0432	6	0.0235	24	0.0012	4	0.0074	3	0.0040	3.3	0.0035
	G24	12	0.1338	15.3	0.1310	18.8	0.1036	24	0.0913	8.3	0.0846	6	0.0824
	G25	6	0.0984	12	0.0837	24	0.0346	4	0.0112	8	0.0108	3	0.0084
	G26	6	0.0930	12	0.0767	24	0.0275	4	0.0144	8	0.0105	3	0.0069
	G27	12	0.0882	6	0.0666	24	0.0260	4	0.0103	8	0.0094	3	0.0078
	G30	12	0.0858	6	0.0707	24	0.0285	4	0.0115	8	0.0107	3	0.0055
	G32	12	0.0899	6	0.0614	24	0.0215	5.5	0.0176	4	0.0094	3	0.0063

satellite clock and GPS satellite clock are different. It can be seen from the amplitude values that the amplitudes of MEO satellite in different periods are significantly lower than that of GEO and IGSO satellites. This may be due to the fact that MEO satellite orbit determination accuracy is higher than that of GEO and IGSO satellites, which leads to the small error in the unmodeled orbit.

3.2 Analysis of SCB Prediction

3.2.1 Analysis of the Influence of the Number of Main Periodic Item

The spectrum corresponding to the peak frequency is the main periodic item, the number of periodic items depends on the power spectrum. However, does the selection of the number of periodic terms have an effect on the accuracy of the prediction? In this paper, 1–4 periods are selected to model the experiment. The SCB data of one week is used as an example. After fitting the SCB of 24 h, the following 24 h prediction is made. The average RMS values of the satellite prediction results under each prediction model are obtained in this period. Tables 3 and 4 show the prediction accuracy of periodic models for different periods of all satellites of BDS and GPS, respectively, and calculate the result of the quadratic polynomial prediction accuracy (unit: ns). The average value represents the average value of the forecast results of all the satellite prediction models under this type.

Table 3 Results of BDS prediction accuracy by using different models

Type	PRN	Polynomial	Number of main periodic items			
			1	2	3	4
GEO	C01	6.2125	5.6900	3.8303	4.8180	5.5731
	C02	6.0926	3.2826	3.9839	2.4505	2.0631
	C03	6.7567	6.6230	6.2205	6.3594	6.7111
	C04	3.7360	3.6385	3.0762	1.7558	3.7279
	C05	12.2691	12.1834	7.0882	5.8538	11.3007
Average		7.0134	6.2835	4.8398	4.2475	5.8752
IGSO	C06	29.7755	24.4090	16.2522	22.5168	9.9052
	C07	11.2862	6.0344	3.7891	2.7201	5.8324
	C08	8.6752	6.8563	8.0595	7.5346	8.7501
	C09	8.2622	5.4283	3.2039	3.3127	3.3434
	C10	17.8365	10.1782	8.8204	7.0312	9.4055
Average		15.1671	10.5812	8.0251	8.6231	7.4473
MEO	C11	6.0124	5.6118	2.5284	2.7689	6.0060
	C12	3.8669	3.7803	2.8556	2.1411	3.7865
	C14	6.0857	5.9071	5.1349	4.1923	3.8779
Average		5.3217	5.0997	3.5063	3.0341	4.5568

Because the C13 satellite has not been served in the experimental data, it is not involved in the subsequent analysis.

In order to analyze the prediction results of different periodic term models, we use 24 h data as the historical data to predict. Taking C07 (IGSO Rb) and G09

Table 4 Results of GPS prediction accuracy by using different models

Type	PRN	Polynomial	Number of main periodic items			
			1	2	3	4
IIR	G02	3.8279	3.7153	3.5891	3.5764	3.6091
	G11	5.8978	5.1405	3.2537	2.8847	2.8106
	G13	4.6613	4.3507	4.3102	2.3844	2.2142
	G14	4.9897	4.9863	4.9568	4.8875	4.8301
	G16	4.8246	4.7260	4.5591	3.7616	4.1319
	G18	4.0442	4.0203	3.9644	3.6900	3.9223
	G19	3.9137	3.4890	3.4799	2.0302	2.3343
	G20	4.6502	4.6357	4.5816	4.0905	3.8268
	G21	4.4107	3.5051	3.4223	2.9276	3.1631
	G22	4.6206	4.3193	4.3674	4.5666	4.3426
	G23	4.3201	4.3022	4.0785	4.0969	2.7942
	G28	4.2385	4.2889	3.8548	3.8959	4.2362
Average		4.5333	4.2899	4.0348	3.5660	3.5180
IIR-M	G05	4.6428	4.6321	4.4375	2.5536	2.3594
	G07	3.9879	3.9004	3.9350	3.9793	3.5377
	G12	3.7217	3.3207	3.0028	1.3935	3.1534
	G15	3.1014	2.9336	2.9579	1.8288	1.8139
	G17	5.7863	5.7452	5.2850	4.3891	4.4357
	G29	5.4262	4.8997	4.9568	2.0822	4.3308
	G31	4.6456	4.5925	4.6178	4.9754	4.1740
Average		4.4731	4.2892	4.1704	3.0288	3.4007
IIF	G01	5.3544	5.1239	4.8071	4.6025	4.6074
	G03	4.3707	4.3275	4.4432	2.0503	1.9172
	G06	4.9033	4.6737	4.0213	2.2772	2.3595
	G08	4.2835	4.1876	4.1127	3.2927	4.1724
	G09	3.0836	3.0365	3.100	1.0488	1.335
	G10	4.4314	4.4057	4.3911	3.9393	4.0926
	G24	4.8017	4.4092	4.8004	3.9526	4.1641
	G25	4.5278	4.0002	3.900	3.3989	3.7653
	G26	4.2647	4.2274	3.5124	1.686	2.6913
	G27	4.075	4.0672	3.6352	3.0014	3.9482
	G30	4.7457	4.5593	4.5445	3.6101	4.1944
	G32	4.7392	4.4125	4.5673	3.4635	3.3741
Average		4.4650	4.2858	4.1529	3.0269	3.3851

(IIF Rb) satellites as an example, Figs. 7 and 8 show the prediction residuals compared with different forecast models.

It can be seen from Table 3 that the prediction accuracy of the periodic term model is better than that of the quadratic polynomial model. However, the prediction accuracy of the periodic correction model with different number of main periodic terms is different. Based on the prediction accuracy of different types of BDS satellites, it is found that the accuracy of introducing 2–3 main period items is well. Comparing the prediction characteristics of satellite clocks with different orbital types, the prediction accuracy of the two forecast models of MEO satellite is the best.

As can be seen from Table 4, comparing the results of the prediction accuracy of each forecast model of GPS satellite, it is found that the prediction accuracy of adding the periodic items is more accurate than that of the quadratic polynomial model. The GPS Block IIF satellite have the best prediction accuracy of the forecast models, followed by Block IIR-M satellite, and the worst Block IIR satellite. The accuracy of introducing 3–4 main period terms in different types of GPS satellites is improved better.

From Figs. 7 and 8, it is obviously seen that the variation trend of prediction residual accuracy after the addition of different periodic terms. The prediction results of the same satellite using different period term forecast models are quite different. The same forecast model varies with the prediction time, and the prediction result also has obvious change. In general, the different number of periodic items has an effect on the prediction accuracy of the periodic term model. But it does not decrease or increase gradually as the number of period terms increasing. Therefore, the number of periodic items should be selected according to the detailed analysis of different satellite and different clock error data.

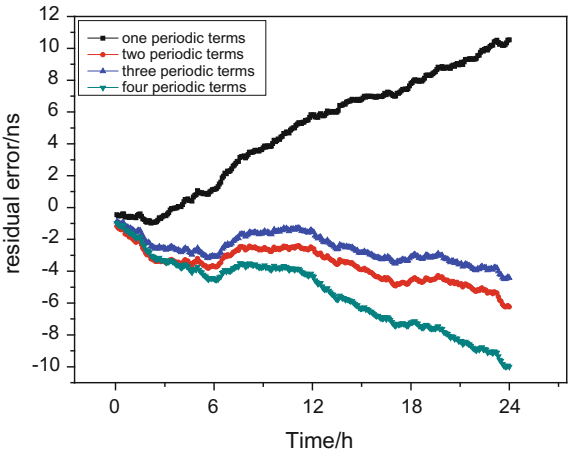


Fig. 7 Residual sequence prediction of C07 satellites with different periods

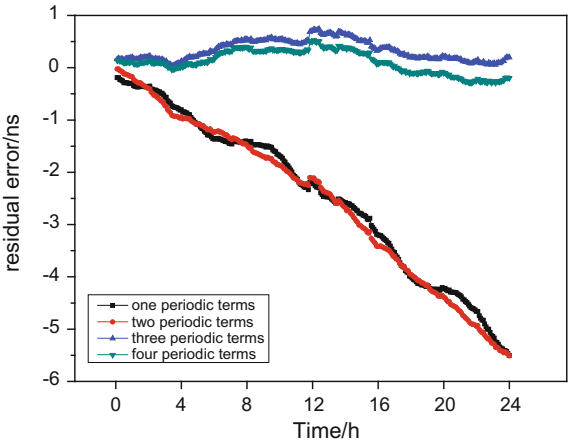


Fig. 8 Residual sequence prediction of G09 satellites with different periods

3.2.2 Analysis of the Improvement of Periodic Items

In order to analyze the improvement effect of the model after adding periodic items, we choose the first three power spectrum frequency to construct the periodic item model. The SCB data of 24 h were used to fit by quadratic polynomial model and periodic item model respectively. Figures 9, 10, 11 and 12 show the comparison result of quadratic polynomial model fitting residual and periodic term model fitting residual of BDS/GPS satellite. In order to clearly see the model features, taking GPS Block IIR-M as an example, the fitting results of two models are compared. Figure 13a is the quadratic polynomial model fitting, and Fig. 13b is the periodic term model fitting.

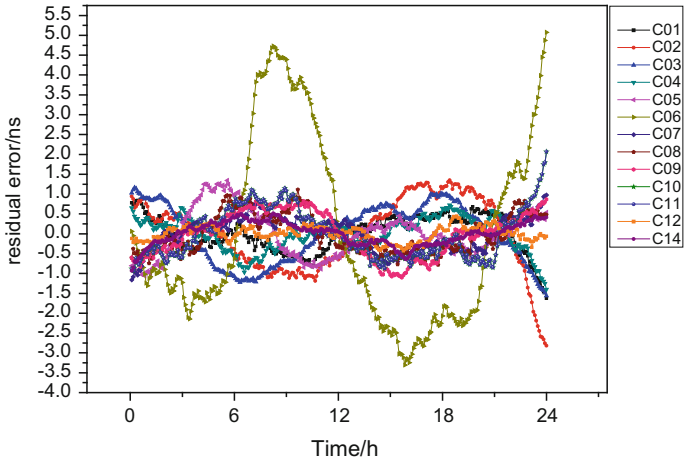


Fig. 9 Quadratic polynomial fitting residuals for BDS satellite

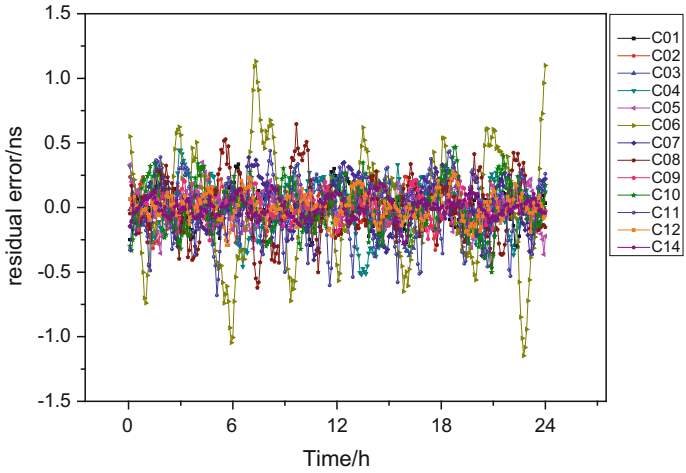


Fig. 10 Periodic item fitting residuals for BDS satellite

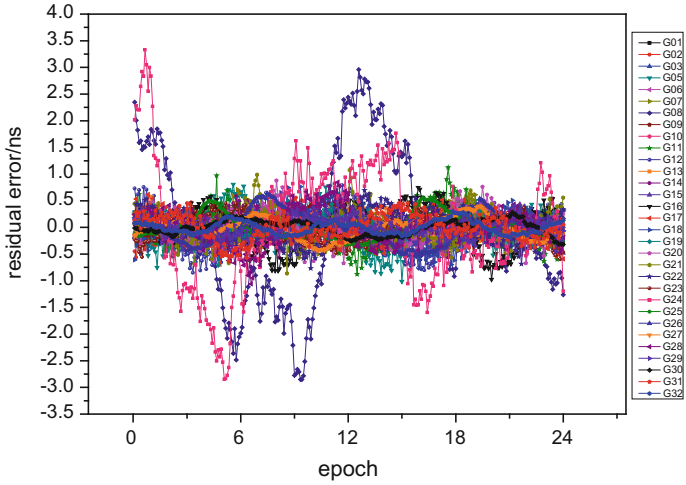


Fig. 11 Quadratic polynomial fitting residuals for GPS satellite

Figure 9 shows the comparison results of BDS satellite fitting residuals. It can be seen that the quadratic residuals of almost all BDS satellites show some periodic characteristics, and their periodicity changes basically accord with 12 and 24 h. In addition, there are great differences in the residual characteristics of different atomic clocks, in which the fitting accuracy of C06 is relatively poor. Compared with Fig. 10, it can be found that the residual accuracy of BDS satellite after eliminating the influence of period term is obviously improved, and the fitting residuals are all

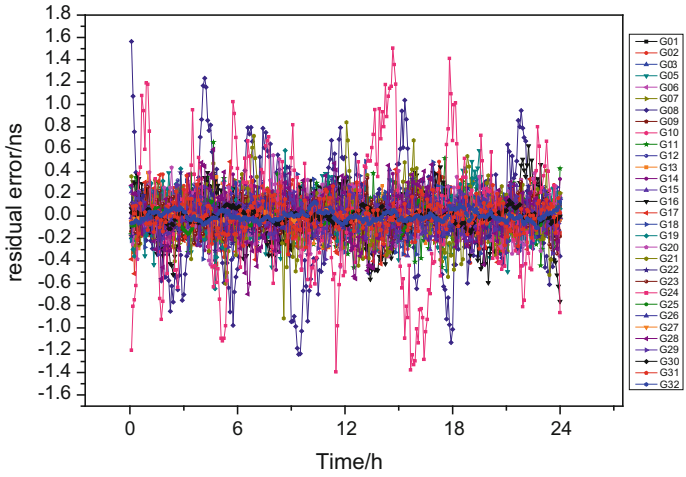


Fig. 12 Periodic item fitting residuals for GPS satellite

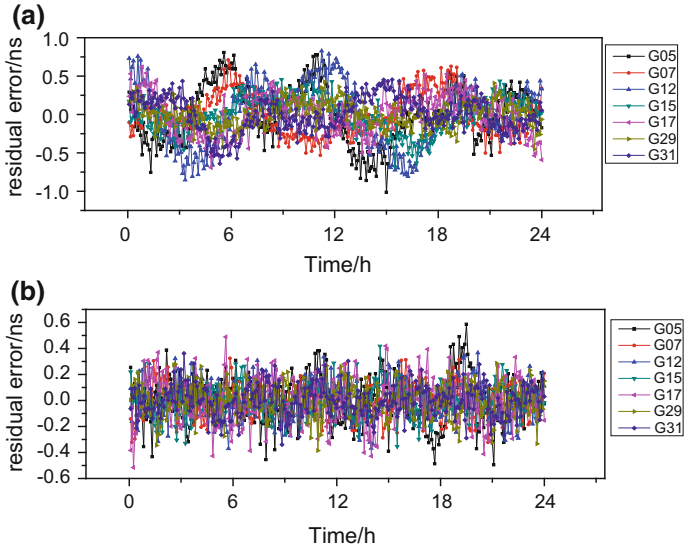


Fig. 13 Comparison of GPS Block IIR-M satellite fitting residuals

within 1 ns, except for C06. While the corrected residuals are still poor for satellites that are not stable by themselves.

Figure 11 shows the comparison results of GPS satellite fitting residuals. It can be seen that almost all GPS satellites quadratic residuals have periodic characteristics with the period value of 12 and 6 h. Compared with Fig. 12, it can be found that when the periodic residuals are eliminated, the precision of other satellites is

obviously improved with the exception of G08 and G24, and the residual values are within 1 ns.

In general, the periodic characteristics of BDS/GPS satellites are significant. The fitting accuracy of the model after periodic term correction is better than that of quadratic polynomial model, but the correction of periodic term is not obvious for satellite with poor stability.

3.2.3 Prediction Accuracy Analysis

As can be seen from Figs. 7 and 8, with the increase of forecasting time, the prediction results of the periodic term model of the same period term vary significantly with time. It can be seen that the prediction effect of the BDS/GPS SCB are related to the prediction period. Therefore, the fitting residuals of the model are used to prediction, and the measured clock error is taken as the benchmark, and the root-mean-square error between the prediction clock error and the reference is taken as the basis for evaluating the prediction accuracy. Taking the 3 weeks SCB data of 20,160,901–20,160,922 as an example, the prediction results of 3, 6, 12 and 24 h are calculated respectively, as shown in Tables 5 and 6.

Figures 14 and 15 show the comparison of the prediction results of the two forecast models of 3, 6, 12, and 24 h for different types of BDS/GPS satellites respectively. The figure shows that the forecast results of same satellite clock error vary with the length of time, and the forecast results of different satellites under the same forecasting time are also different. Compared with different forecasting time, the prediction accuracy of 3 h is the best, and the prediction accuracy is relatively worse with the increase of forecasting time. Compared with the average RMS of the same forecasting time, it is found that the prediction accuracy of the BDS MEO is the best, followed by the IGSO and the worst is the GEO, and the prediction accuracy of the GPS IIF is the best, followed by the IIR-M and the worst is the IIR.

Compared with the statistical results of each model, it is found that the prediction accuracy of periodic term model is better than that of quadratic polynomial model. Specifically, periodic model analysis results of 3 h prediction accuracy is about 2 ns for BDS satellite, and the prediction accuracy of 24 h is about 10 ns. The periodic model analysis results of 3 h prediction accuracy is about 3 ns for GPS satellite, and the prediction accuracy of 24 h is about 4 ns.

Table 5 RMS values of SCB prediction results for the satellite of BDS

Type	PRN	3 h prediction accuracy (ns)		6 h prediction accuracy (ns)		12 h prediction accuracy (ns)		24 h prediction accuracy (ns)	
		Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model
GEO	C01	1.2328	0.7399	2.2274	1.6609	5.3068	3.5194	6.2125	4.8181
	C02	2.3774	1.2500	3.3100	1.6249	5.3795	3.0993	8.2533	5.5076
	C03	4.3956	2.1052	5.2813	1.8911	5.9864	1.5321	6.0607	4.2272
	C04	5.3374	1.2524	6.4037	2.4408	9.2989	4.3753	18.6025	10.4158
	C05	3.8682	2.7278	6.4651	3.6936	12.4391	6.3961	23.4061	12.4050
Average		3.4423	1.6151	4.7375	2.2623	7.6821	3.7844	12.5070	7.4747
IGSO	C06	6.1466	3.0597	12.7802	3.3401	18.5766	8.1319	26.4505	15.2704
	C07	2.0936	1.3414	3.2961	1.6716	6.3699	2.2439	11.4376	3.7936
	C08	2.2086	1.3842	3.6624	2.7376	6.5712	6.0102	10.5900	8.2841
	C09	1.6374	0.5088	2.7123	0.6251	4.8055	1.1604	5.6647	2.7849
	C10	0.9941	0.5747	2.1633	1.6215	4.8613	2.0046	8.2791	3.5788
Average		2.6161	1.3738	4.9229	1.9992	8.2369	3.9102	12.4844	6.7424
MEO	C11	1.8216	0.9217	1.9968	1.2104	2.1513	1.2761	13.0314	11.1076
	C12	3.0172	1.8411	2.9018	1.7423	2.6354	1.8331	3.5379	2.3531
	C14	2.5041	1.9071	2.6943	2.0545	4.8211	3.1186	6.0857	4.2461
Average		2.4476	1.5566	2.5310	1.6691	3.2026	2.0759	7.5517	5.9023

Table 6 RMS values of SCB prediction results for the satellite of GPS

Type	PRN	3 h prediction accuracy (ns)		6 h prediction accuracy (ns)		12 h prediction accuracy (ns)		24 h prediction accuracy (ns)	
		Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model
IIR	G02	2.6606	2.0897	2.8429	2.1290	3.2229	2.4443	4.2469	3.3160
	G11	3.8278	1.4906	4.1657	1.9519	4.4854	2.7460	5.2215	3.4233
	G13	3.5671	3.2557	3.6493	3.2706	3.9038	3.4266	4.8600	4.0565
	G14	4.0327	3.8835	4.0584	3.8640	4.1703	3.9472	4.6617	4.2264
	G16	4.3954	2.6320	4.4696	2.7401	4.5513	3.0702	4.8681	4.1880
	G18	3.0025	2.1266	3.1357	2.5779	3.4865	3.0846	4.2057	4.1624
	G19	3.7292	2.8869	3.7860	3.2671	3.8969	3.4714	4.4188	4.2213
	G20	4.2379	3.7944	4.3168	4.0217	4.5817	4.3520	5.3648	5.2433
	G21	4.0594	3.3596	4.2273	3.4150	4.4722	3.6790	5.1331	4.6915
	G22	3.2718	3.1965	3.9950	3.3571	4.0754	3.9917	4.5320	4.3580
	G23	3.8916	2.5407	4.0335	2.7109	4.2704	2.9838	5.0295	3.7389
	G28	4.6638	3.5712	4.6937	4.4382	5.6536	5.0447	6.6081	5.1478
Average		3.7783	2.9023	3.9478	3.1453	4.2309	3.5201	4.9292	4.2311
IIR-M	G05	3.7887	1.3818	3.9272	1.3685	4.0422	1.6766	4.2423	3.2337
	G07	3.9856	3.8627	4.0634	3.9848	4.1670	4.0723	4.7478	4.3846
	G12	3.4300	2.0285	3.5234	2.2326	3.5701	2.7083	4.0440	3.8731
	G15	2.7369	2.1711	2.8952	2.3640	3.1633	2.4642	3.9060	2.5123
	G17	4.3479	3.5864	4.4507	3.7793	4.7669	4.6031	5.8695	4.6994
	G29	3.3703	2.1685	3.5230	2.5959	3.7692	3.5556	4.6981	4.0523
	G31	3.3704	3.1775	3.5143	3.3938	3.9304	3.7779	4.9371	4.4902
Average		3.5757	2.6252	3.6996	2.8170	3.9156	3.2654	4.6350	3.8922

(continued)

Table 6 (continued)

Type	PRN	3 h prediction accuracy (ns)		6 h prediction accuracy (ns)		12 h prediction accuracy (ns)		24 h prediction accuracy (ns)	
		Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model	Quadratic polynomial model	Periodic term model
IIF	G01	4.6840	3.7817	4.7728	3.8178	4.8540	3.8660	4.9690	4.0059
	G03	2.5200	0.7459	2.6433	0.7397	2.7944	0.8767	3.1113	1.2231
	G06	3.6783	2.2756	3.7664	2.5074	3.9468	2.9774	4.3852	3.8541
	G08	2.8466	1.5251	2.2911	1.7027	2.5250	2.3206	4.2835	3.9392
	G09	2.5131	1.2520	2.5943	1.3879	2.8090	1.7754	3.4397	2.6132
	G10	3.6144	2.5669	3.6903	2.7208	3.8461	3.0193	4.2258	3.7654
	G24	3.1896	3.1145	3.0036	3.1065	4.4813	4.3577	8.9682	7.8817
	G25	4.2917	3.1915	4.4304	3.3045	4.4743	3.6425	4.7528	4.3592
	G26	2.5333	2.5200	2.6508	2.5627	2.9085	2.6893	3.6688	3.1045
	G27	3.7963	3.0166	3.8880	3.1062	3.9219	3.3175	4.1471	3.7920
	G30	4.1003	3.2499	4.1687	3.3407	4.2846	3.5306	4.5103	3.8700
	G32	4.5672	3.6300	4.6482	3.8106	4.6908	4.1616	4.8417	4.1688
Average		3.5279	2.5725	3.5457	2.6756	3.7947	3.0446	4.6086	3.8814

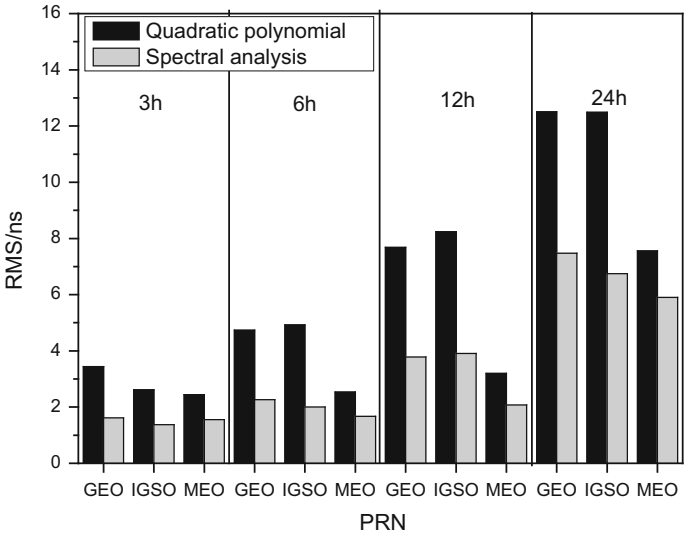


Fig. 14 Prediction errors of different types of BDS satellites

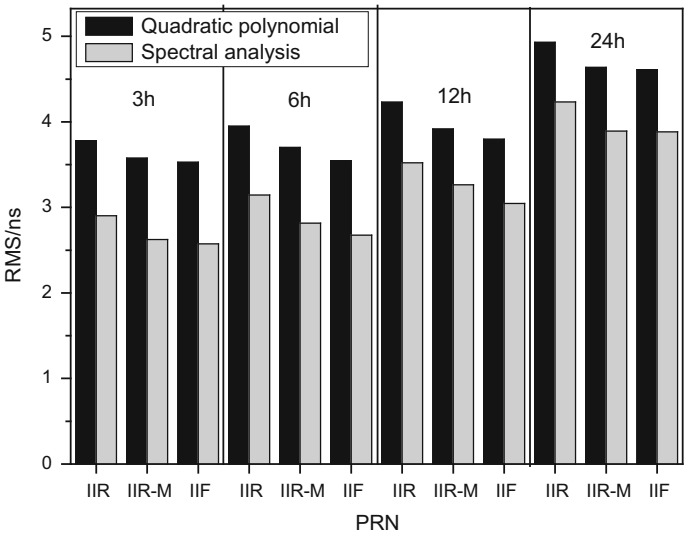


Fig. 15 Prediction errors of different types of GPS satellites

4 Conclusion

In this paper, we analyzed the periodic characteristics of BDS/GPS SCB, established the prediction model of periodic items, and verified that the model in short-term prediction is superior to quadratic polynomials. The conclusions are as follows:

1. There are significant periodic variations in the clock sequence of BDS satellite and GPS satellite, but the significant periodic items of different orbital types are different, and the periodic items of the same orbital type also have differences.
2. The number of periodic items has some effect on the forecast results of periodic items. With the increase of the number of periodic, the improvement of the accuracy may be improved, and may also be reduced. The accuracy improvement of BDS satellite is better when using 2–3 main period items. The accuracy improvement of GPS satellite is better when using 3–4 main period items.
3. Prediction precision of periodic term model is significantly improved compared with quadratic polynomial model. However, for the satellite with poor stability, the corrected prediction residual is still poor.
4. The prediction results of the same satellite using different forecast models are quite different. For the same satellite of BDS, the accuracy of 24 h prediction can be better than 7 ns when using periodic item model, while the accuracy of 24 h prediction can be better than 13 ns when using quadratic polynomial model. For the same satellite of GPS, the accuracy of 24 h prediction can be better than 4 ns when using periodic model, while the accuracy of 24 h prediction can be better than 5 ns when using quadratic polynomial model.
5. The prediction accuracy of the same satellite at different forecast times is quite different. The accuracy of 3 h prediction of BDS is 8 ns higher than that of 24 h. The accuracy of 3 h prediction of GPS is 1.5 ns higher than that of 24 h.

Acknowledgements Project supported by the National Natural Science Foundation of China (Grant No. 41774018).

References

1. Fu W (2014) Research on prediction and characteristics of satellite clock on orbit. Chang'an University
2. Wang YP, Lv ZP et al (2014) A method of satellite clock bias prediction based on grey model and wavelet neural network. J Geodesy Geodyn 34(3):155–159
3. Xu JY, Zeng AM (2009) Application of ARIMA(0,2, q) model to prediction of satellite clock error. J Geodesy Geodyn 29(5):P116–120
4. Wang YP, Lv ZP, Chen ZS et al (2013) Research the algorithm of wavelet network to predict satellite clock bias. Acta Geodaetica Cartogr Sin 42(3):323–330
5. Tian J (2015) Study on the GPS/BDS a tomic clock performance and clock offset prediction model. Chang'an University

6. Davis J, Bhattarai S, Ziebart M (2012) Development of a Kalman filter based GPS satellite clock time-offset prediction algorithm. In: European frequency and time forum
7. Huang GW et al (2014) Real-Time clock offset prediction with an improved gps satellite clock model. GPS Solutions 18(1):95–104
8. Huang GW (2012) Research on algorithms of precise clock offset and quality evaluation of GNSS satellite clock. Chang'an University
9. Wang YP, Lv ZP, Wang N (2017) The long-term performance analysis for on-board atomic clocks of BDS. Acta Geodaetica Cartogr Sin. https://doi.org/10.11947/j.AGCS.2017.20160369
10. Chen ML, Zhan XQ, Du G et al (2015) Compass/BeiDou-2 Spaceborne clock performance assessment and its error detection, Mitigation. Trans Electr Electronic Eng 10(4):438–446
11. Zhou PY, Du L, Lu Y et al (2015) Periodic variations of BeiDou satellite clock offsets derived from multi-satellite orbit determination. Acta Geodaetica Cartogr Sin. https://doi.org/10.11947/j.AGCS.2015.20150183
12. Zheng ZY, Lu XS (2008) Comparison and precision analysis of several GPS satellite clock bias prediction methods. J Shandong Univ Sci Technol 27(4):6–12

Analysis of the Impact of Small Magnetic Storms on the Evaluation of Broadcast Ionospheric Model

Teng Peng, Xiao Lin Jia and Yue Mao

Abstract Using the Space Environment Prediction Center's related products, performance and analysis on broadcast ionosphere model during the small magnetic storm on April 4, 2017. The result shows that storm have had some impact on the three models, but in the degree of influence, The correction rate of GPSK8 model is around 15%, while BDSSH and NeQuick2 are within 10%. In comparison of the RMS of the mean square root error, the mean square root of the GPSK8 model is the largest, and the BDSSH model is slightly less than the NeQuick2 model. It is indicated that the ionosphere model is the best for the adaptability of magnetic storm, and the best of BDSSH is NeQuick2, and the GPSK8 model is the worst.

Keywords NeQuick2 · BDSSH · Ionosphere magnetic storm · Datum VTEC

1 Introduction

The geomagnetic storm is a global phenomenon of intense disturbances of the Earth's magnetic field, which is caused by high-speed plasma clouds are ejected from the solar corona, carrying a coronal magnetic field that impacts the Earth's magnetosphere, compressing the magnetosphere so that the magnetic layer is compressed and deformed. The causes ionospheric anomaly disturbances.

T. Peng (✉)
College of Geology Engineering and Geomantic, Chang'an University, Xi'an 710054, Shanxi, China
e-mail: 2298452014@qq.com

X. L. Jia (✉) · Y. Mao
State Key Laboratory of Geo-Information Engineering, Xi'an 710054, Shanxi, China
e-mail: 13891907401@139.com

X. L. Jia · Y. Mao
Xi'an Research Institute of Surveying and Mapping, Xi'an 710054, Shanxi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_3

The geomagnetic storm has the characteristics of large variation, complex manifestation and long duration. The impact of a magnetic storm on the accuracy of communication systems, space vehicles, and GNSS navigation positioning and timing is severely affected [1, 2]. After years of development, people have a certain understanding of the geomagnetic, at the same time, domestic and foreign scholars on this also has in-depth research [3–10], But it mainly focuses on two aspects: on the one hand, it analyzes the influence of the quality anomaly of observational data caused by geomagnetic storm on communication and navigation positioning accuracy, on the other hand, the influence of geomagnetic storms on the perturbation characteristics of the global ionosphere TEC is analyzed. The effect of the correction effect of the ionospheric model caused by the geomagnetic storm and the adaptability analysis of the model to geomagnetic storms are relatively few.

It is necessary to analyze the correction effect of ionospheric model to consider selection of baseline information, using the ionospheric grid products provided by the CODE analysis center as the reference information, the distribution of global stations are uneven and the results are bias [11]. Therefore, in the selection of datum, this paper adopt the unified High-precision reference value as the baseline, it is the VTEC information obtained from the measured data of the station as the benchmark. To utilize IONAPS software extract the ionospheric pierce point information of the station requires to the satellite and receiver DCB correction information, The weakness of the DCB product information provided by the CODE analysis center has a certain lag, It can be caused not good to reflect the daily variation rules of receiver and satellite DCB information, such as use the monthly value information is difficult to effectively detect the station receiver terminal replacement. For this reason, this paper extracts the VTEC information by using the day value information instead of the average monthly value.

Currently International GNSS Monitoring Assessment work is under implementation, analysis the factors influencing of broadcast ionospheric model correction is also one of the tasks, ionospheric storm is an external factor to be considered. Therefore, the research work in this paper can provide some reference for the analysis of related work in International GNSS Monitoring Assessment.

2 Observation Stations Select and Strategy Analysis

2.1 Selection the Observation Stations

In the AR2033 (S12° W65°) area UTC 20:38 on 2 April. take place Flux M 5.7 event, this event last 44 min, The affected area is shown in Fig. 1.

According to the Fig. 1, the Fountainhead flux located in the equatorial area and affected the local area of the Americas American. So this paper select five IGS stations in the American area (ALBH,AMC2,GOLD,DRAO,QUIN).

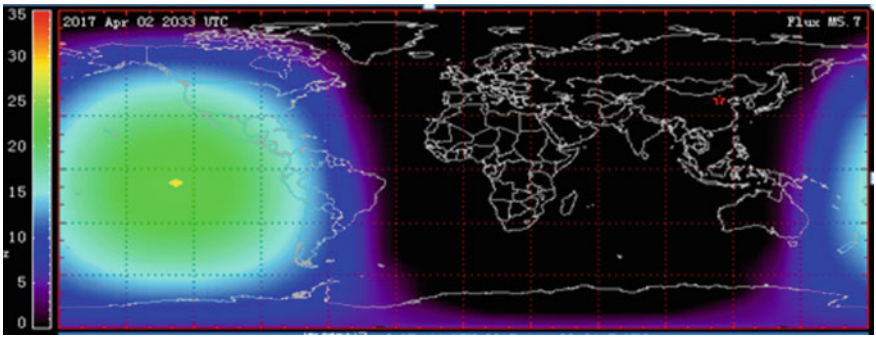


Fig. 1 Flux M 5.7

2.2 Processing Strategy

2.2.1 Benchmark VTEC

Using the dual-frequency pseudo-distance observation value to obtain the propagation path. The formula is:

$$d_{ion} = \frac{f_2^2}{f_1^2 - f_2^2}(p_1 - p_2) \tag{1}$$

Formula 1, f_i is frequency, p_i is the code observation value, considering the hardware delay of satellite and receiver, the DCB information must be corrected in order to obtain more accurate ionospheric information. The formula is as follows:

$$p_4 = p_1 - p_2 = 40.28(\frac{1}{f_1^2} - \frac{1}{f_2^2}) \cdot vtec + c \cdot DCB_r + c \cdot DCB_s \tag{2}$$

Formula 2, the DCB_r is hardware latency on the receiver side, DCB_s is the hardware delay on the satellite side. Because p_i is the absence of geometrical combined observations. p_4 contains a lot of noise, so need to use the carrier phase smoothing pseudo-distance technology for pseudo-range smoothing, the formula is:

$$\begin{aligned} \widetilde{p_1}(t) &= \varphi_1(t) + \overline{p_1} - \overline{\varphi_1} + 2 \cdot \frac{f_2^2}{f_1^2 - f_2^2} \cdot [(\varphi_1(t) - \overline{\varphi_1}) - (\varphi_2(t) - \overline{\varphi_2})] \\ \widetilde{p_2}(t) &= \varphi_2(t) + \overline{p_2} - \overline{\varphi_2} + 2 \cdot \frac{f_1^2}{f_1^2 - f_2^2} \cdot [(\varphi_1(t) - \overline{\varphi_1}) - (\varphi_2(t) - \overline{\varphi_2})] \end{aligned} \tag{3}$$

Formula 3, $\widetilde{p_i}$ is the epoch time frequency after point smoothing code observations, $\varphi_i(t)$ is the epoch time frequency for carrier phase observations. $\overline{p_i}$ is the Observation segmental arc for frequency Average value of observation. At last,

using the $\widetilde{p_i}$ replace the p_i, this can be used to accurately extract the VTEC values of the observation epoch, as a baseline information.

2.2.2 Comparison Between the Day DCB and the Monthly Mean

Using the Fortran program, the DCB information of the satellite and receiver of the ionospheric grid file is extracted. The DCB information of the stations selected from April 1, 2017 to April 4, 2017 (the day of year 91–94) are compared with the monthly mean value, and the results are shown in Tables 1 and 2.

According to Tables 1 and 2, in addition to GLOD station day value information and the average monthly information in a few days outside the slightly different, the other stations different are within 0.2 ns. It shows that the information of the days value is more stable in this time period. At the same time, in the comparison of the same station, there is a 0.3–0.6 TECU improvement in the extraction of baseline information using the day value information than the monthly average information. The baseline information obtained by using the day value information has 3–5% improvement in the model correction rate, correction accuracy of 0.4–0.6 TECU improvements. Therefore, use day value information can be more accurate VTEC station ionospheric pierce point, and improve the result of the model evaluation.

Table 1 Daily DCB and monthly average DCB (ns)

Station	ALBH	QUIN	AMC2	GOLD	DRAO
91	14.154	−7.69	4.948	−8.31	7.755
92	14.082	−7.79	4.801	−8.38	7.667
93	14.173	−7.62	4.843	−8.12	7.781
94	14.282	−7.61	4.959	−8.28	7.872
In the mean	14.241	−7.59	4.797	−8.22	7.853

Table 2 Difference value and RMS (ns)

Station	ALBH	QUIN	AMC2	GOLD	DRAO
91	−0.087	−0.1	0.151	−0.09	−0.098
92	−0.159	−0.2	0.004	−0.16	−0.187
93	−0.068	−0.03	0.046	0.991	−0.072
94	0.041	−0.02	0.162	−0.06	0.019
RMS	0.099	0.113	0.114	0.109	0.112

2.2.3 Evaluation Indicators

Using the stations observed data and broadcast ionospheric parameters to evaluate the accuracy of the different broadcast ionospheric models, use the correction rate P and Root Mean Square (RMS) as a statistic. The formula is

$$P = \sum_{i=1}^{n} (1 - \frac{|VTEC_{评估} - VTEC_{基准}|}{VTEC_{基准}}) / n$$
$$RMS = \sqrt{\sum_{i=1}^{n} (VTEC_{评估} - VTEC_{基准})^2 / n} \quad (4)$$

Formula 4, $VTEC_{评估}$ is ionospheric delay correction models VTEC, $VTEC_{基准}$ is the baseline information. For statistics of RMS and P, use day and night as a Statistical strategy, the day is (07:00–19:00, local time), Night is (19:0—07:00, local time).

3 Experimental Results and Analysis

On April 4, 2017, the KP index and DST index of the magnetic storm day are shown in Figs. 2 and 3.

It can be seen from Figs. 2 and 3, the magnetic field at 6:00–11:00 on April 4 reaches the small magnetic storm level (KP = 5, DST is greater than −50 nt less than −30 nt), and at the same time, KP = 4 at 3:00–5:00, and the geomagnetic field is in active state.

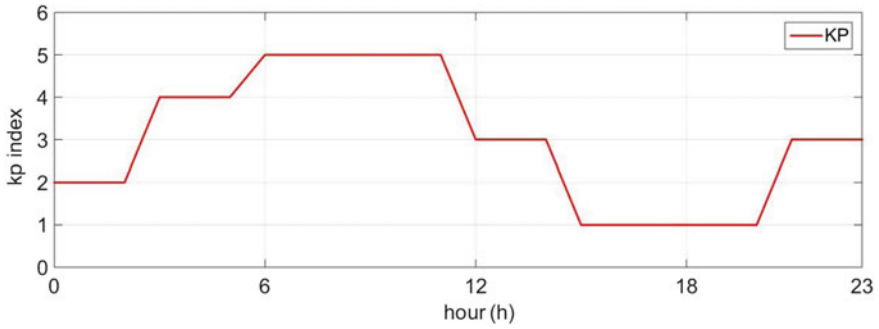


Fig. 2 KP index on April 4

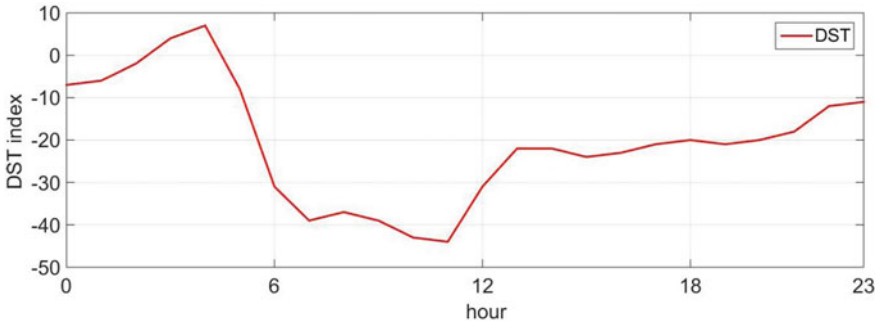


Fig. 3 DST index on April 4

3.1 Impact on the Baseline VTEC

This article on the extraction of the ionosphere VTEC when analysis information, select the most intense magnetic disturbances 09:00–11:00 time period analysis, Due to there are several satellites at a certain time, it's not very good to reflect the change of the baseline information. Therefore, this paper only analyzes the changes of single satellite in multiple station. Figure 4 show the change of the G21 satellite during the most active period of the magnetic storm.

As shown in Fig. 4, due to the stations in different latitude, The variation trend of VTEC is slightly different. During the maximum period of time of geomagnetic disturbance at 09:00–11:00, there is no significant fluctuation of datum VTEC information in the station single satellite. To comparing the changes of VTEC during periods of calm and magnetic storms. Chosen April 7 as a comparison day, also use the G21 information from 9:00 to 11:00 time period, the information comparison results of VTEC in the period of April 4 and April 6 from 09:00–11:00 are shown in Table 3.

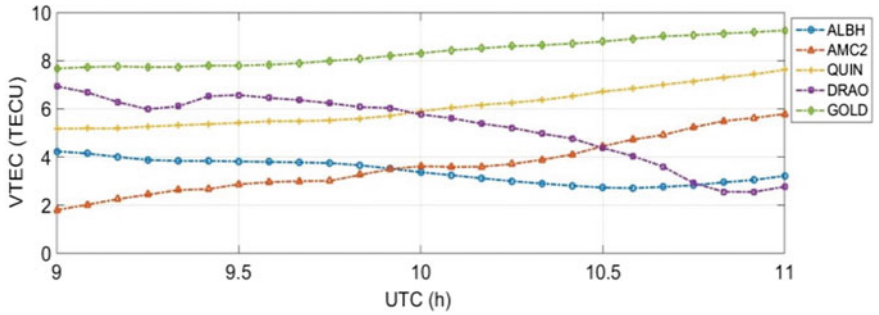


Fig. 4 G21 analysis results from VTEC at 5 stations

Table 3 Comparison of the magnetic storm day and the calm day VTEC (TECU)

	April 4		April 6	
Station	Max	Min	Max	Min
ALBH	4.24	2.91	3.88	2.78
AMC2	5.79	2.19	5.63	2.14
QUIN	7.63	5.18	6.75	4.29
GOLD	9.27	7.68	7.42	5.54
DRAO	6.94	2.77	6.51	2.92

Table 3 shows, during the period of small magnetic storms, except for the GOLD station's baseline information difference more than one TECU, (The maximum and maximum ratio, the minimum and minimum ratio), the VTEC information of other stations are changed within 1 TECU. Zhang et al. [12] believe that the influence of the M5.1 and M4.6 flares on the station's VTEC are within 1.5 TECU, the results are consistent with this article. It is shown that the actual VTEC information of the small magnetic storm is not significant.

3.2 *The Effect of the Modification Broadcast Ionosphere Model*

In order to analyze the effect of geomagnetic storm on model performance of broadcast ionosphere, the performance of GPSK8, BDSSH and NeQuick2 broadcast ionospheric model were analyzed during the geomagnetic storm on April 4, and the results are shown in Figs. 5 and 6.

The correction rate and root-mean-square error of the three models in Figs. 5 and 6 can be seen:

(1) In the correction rate, the magnetic storm has the greatest influence on the correction rate of the GPSK8 model, and the overall correction rate is less than 60%, while the BDSSH model and NeQuick2 model have relatively small influence.

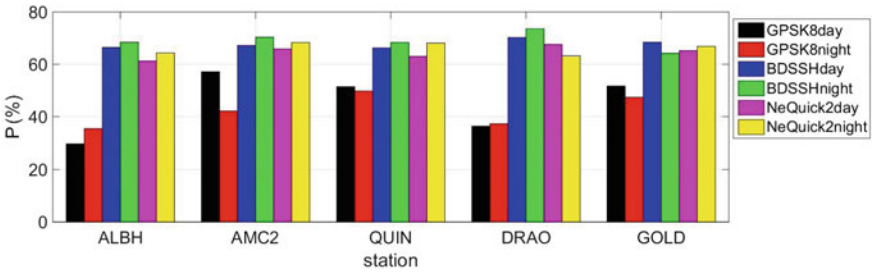


Fig. 5 Statistics of correction rate of station

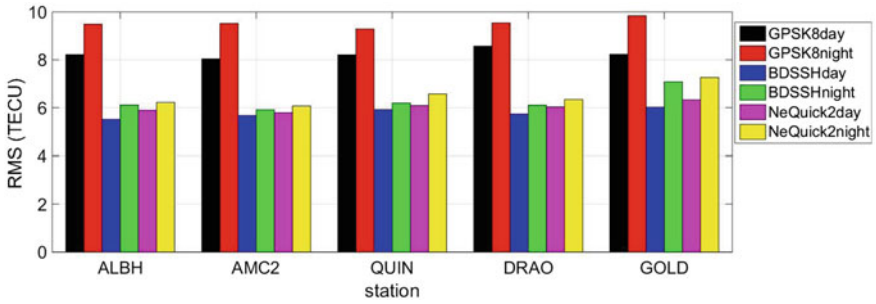


Fig. 6 RMS statistics of station

(2) In RMS comparison, the correction error of GPSK8 model is greater than that of BDSSH model and NeQuick2 model.
(3) In the same model, the correction effect of the model during the day is slightly better than that in the night.

The above analysis is just comparison of the three models information on the day of the magnetic storm., In order to better reflect the change of model correction effect in the period after the geomagnetic storm, analysis of the correction effect of the three models from April 1 to April 5 for analysis. Due to space limitations, this paper only shows the modification effect of the three models of the QUIN station, the results are shown in Tables 4, 5 and 6.

According to the information of the time period after the magnetic storm, the magnetic storm has a certain influence on the performance of three models. This is due to the emergence of M5.7 flares on April 2, leading to geomagnetic activity become more active and perturbation during the most of period.

Table 4 GPSK8 model performance analysis

Date	Day p	Night p	Day RMS	Night RMS
April 1	55.43	41.26	7.59	9.72
April 2	48.47	37.74	8.17	10.12
April 3	54.98	49.47	8.04	9.86
April 4	48.2	46.58	8.32	9.57
April 5	51.07	45.66	7.82	9.34

Table 5 BDSSH model performance analysis

Date	Day p	Night p	Day RMS	Night RMS
April 1	79.58	78.54	4.29	4.98
April 2	77.62	82.85	4.43	4.61
April 3	72.56	75.63	4.97	5.42
April 4	69.32	71.15	5.93	6.22
April 5	79.57	83.56	4.86	5.23

Table 6 NeQuick2 model performance analysis

Date	Day p	Night p	Day RMS	Night RMS
April 1	62.83	74.75	6.03	5.74
April 2	66.53	69.69	5.74	5.57
April 3	68.47	78.56	5.67	5.28
April 4	63.68	70.29	5.87	5.49
April 5	70.55	79.67	5.35	4.96

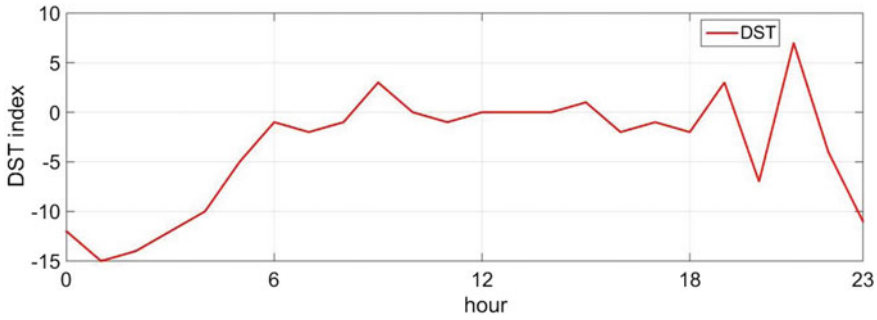


Fig. 7 DST index on April 7

In order to further analyze the effect of magnetic storm on the model, The comparative analysis was conducted on April 7, when the geomagnetic activity was in quiet period (the day of year is 97). The DST index on April 7th is shown in Fig. 7.

Also select the QUIN station for analysis, the comparison results of three models between the quiet period and the magnetic storm period are shown in Figs. 8 and 9.

It can be seen from Figs. 8 and 9 that:

(1) In QUIN station, the BDSSH model was better than NeQuick2 and GPSK8 models in the comparison of the three models.

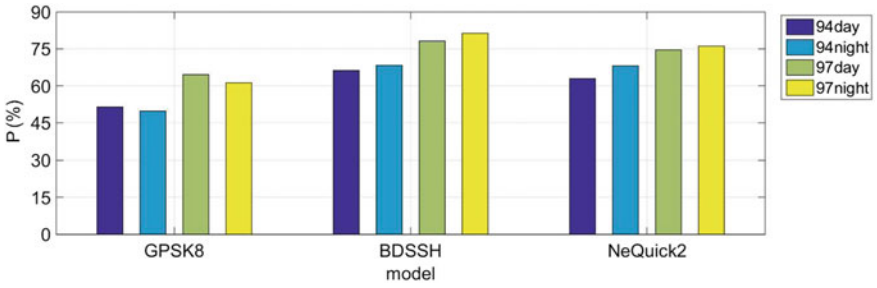


Fig. 8 Correction rate comparison

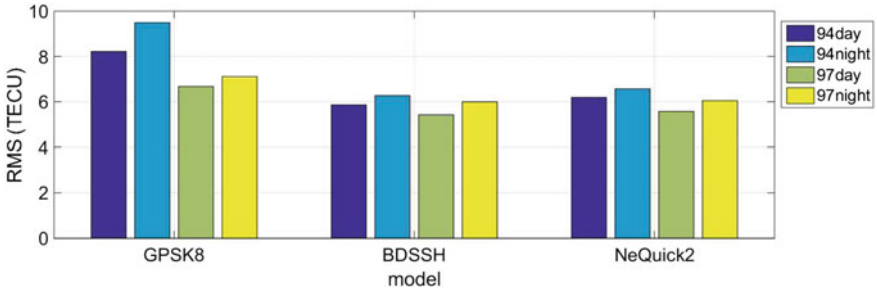


Fig. 9 RMS comparison

(2) there is a certain difference in the influence degree of magnetic storm on the three models. GPSK8 the daytime difference between calm and geomagnetic storms is reached 15.8% The difference at night is in 17.11%, the NeQuick2 model and the BDSSH model are within 10%. The other stations during the period of magnetic storm were similar to QUIN station. In the comparison between the BDSSH model and the NeQuick2 model, the correction rates are similar, and the BDSSH model is superior to the NeQuick2 model in the root mean square. The BDSSH model is better than the NeQuick2 model and the GPSK8 model for adaptability of small magnetic storm.

4 Conclusion

This paper take example of the small magnetic storm on April 4, 2017, analysis the influence of geomagnetic storm on the modified effect of broadcast ionospheric model, The following conclusions can be drawn:

1. Analysis of the baseline VTEC information, the small magnetic storm did not have a significant impact on the baseline VTEC information of the station.
2. Through the comparison analysis of several days before and after the magnetic storm, it can be seen that, When the local magnetic activity is in a more active and perturbation state, the effect of the modification on the broadcast ionospheric model is not obvious.
3. It can be seen from the comparison between the magnetic storm day and the calm day, Magnetic storms have varying degrees of influence on the three models. In GPSK8 model, the period of calm and geomagnetic storm, the change in correction rate has reached 10–17%, the change in the night has reached 12–19%. The changes in the NeQuick2 model and the BDSSH model were between 5 and 10%. In the comparison of root-mean-square errors, the change of BDSSH model was 0.3–0.8 TECU, while the NeQuick2 model was slightly larger, 0.6–1.1 TECU. It is indicated that BDSSH is better than NeQuick2 model and GPSK8 model for small magnetic storm.

Acknowledgements Thanks for the product and data provided by the CODE analysis center and the space environment prediction center of the Chinese academy of sciences, xi'an Aerospace Data Technology Co. Ltd provided ionospheric parameters.

National Natural Science Foundation: 41604024

References

1. Xie YQ (2007) Compressive studies on solar storm. Center of space science and applied Chinese Academy of science
2. Li CB (2012) Statistical study of ionosphere Mid/low-latitude anomalous disturbances. Submitted to Xi Dian University. XI'an
3. Li Q, Zhang D (2006) Analysis of the Global Ionospheric TEC disturbance during the magnetic storm to NOV 24. CHIN J Space Sci 26(6):440–445
4. Yang Y (2014) Characteristics of the ionosphere affected by small and medium geomagnetic storms GNSS Data Analysis. Northeastern University
5. Lu F, Xu JS, Zou YH (2004) Analysis of observations of ionospheric storms TID Using a GPS–array. J Wuhan University (Science EDITION) 03:365–369
6. Xu-chun L, Zheng Z, Jun Y (2007) Monitoring results and analysis of nonblank different solar flares using GPS. Sci Surv Mapp 32(5):220–226
7. Ding Y, Yang SG, Wang H (2017) Statistical study on the response of TEC to geomagnetic storms. Chinese J Space Sci 37(5):524–530
8. Wang XY, Liu L, Guo CJ (2015) In procures 17th during a magnetic storm of temporal and spatial variation of European regional ionospheric electron density Analysis of Satellite navigation System management office of academic exchange center. The eighth Chinese satellite navigation symposium. Chinese satellite navigation system management office of academic exchange center: China Satellite Navigation Academic Annual meeting of the organizing committee, 2017:5
9. Peng D, Zhang YB, Hu YL, Sheng ZH, Hu YL (2016) Xu Tong. Influence of solar and geomagnetic activities on SPORADIC-E Layer over different latitudinal stations. Chinese J Space Sci 36(06):881–887
10. Cheng N (2015) Study on methods of precision monitoring and assessment of broadcast ionospheric model of GNSS. Xi'an, Chang'an university
11. Liu S, Jia XL (2016) GNSS ionospheric model correction accuracy assessment and analysis. J Space Sci 297–306
12. Zhang K, Lgarashi K, Zuo S, Dong He (2002) The observation of SCM scale travelling ionospheric disturbances based on GPS receptacle. Chinese J Geophys 45(4):0001–5733

Design and Analysis of Chain Construction of Navigation Constellation

Canyou Liu, Wei Zhang and Baojun Lan

Abstract The research goal of this paper is the design analysis of the construction of navigation constellation stars in time division system to meet the requirements of navigation constellation in terms of measurement and data transmission. By uniformly integrating the continuously visible links between the stars into the slot cycle, the non-continuously visible links of the remaining slots are prioritized with high probability of not being in the same territory at the same time, and increasing the use of high-orbit satellites to establish domestic and foreign links, Inter-satellite link and route planning principles Constellation in each star's time slot table and routing table to complete the pre-distribution of each node connection. The statistical PDOP values of all the satellites and any two-way time slots show that the range of the PDOP is 0.47–1.42, meeting the measurement requirement of PDOP < 1.5. The average transmission delay between stars is 3.5 s, meeting the requirements of data transmission delay.

Keywords Interstellar links · Position dilution of precision · Continuously visible link · Time-varying link · Time slot table · Routing table

C. Liu (✉) · B. Lan
State Key Laboratory of Astronautic Dynamics, Xi'an Satellite Control Center, Xi'an, China
e-mail: liu_yusi1@163.com

B. Lan
e-mail: lanbaojun@163.com

W. Zhang
Troops61363, Xi'an, China
e-mail: 945052495@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_4

1 Introduction

The use of inter-satellite links is an inevitable trend in the development of satellite systems. Without the support of inter-satellite links, many satellite stations must be established in the world due to the large number of satellites in some satellite constellation systems. The use of roads, you can better solve this problem. The research goal of this paper is to design and analyze the chain building between navigation constellations (24MEO + 3GEO + 3IGSO) to meet the requirements of navigation constellation in terms of measurement and data transmission, such as the lack of satellite observation under autonomous navigation, Orbit and autonomous navigation on the observation geometry requirements, the satellite link to the ground calibration, the anchor mode of operation requirements, as well as the requirements of the inter-satellite time-lapse, etc., for the navigation star constellation of global interstellar construction chain to provide reference.

2 Requirements of Chain Building Between Stars

2.1 Measurement Requirements of Chain Building

Considering the lack of satellite observations in the case of autonomous navigation, high and low building chains are required [1]; in view of joint orbit determination, autonomous navigation requires low and low building chains [2]; in consideration of the calibration of the inter-satellite link to the ground, Anchoring work mode, requiring the construction of the chain to the ground [3]; taking into account the inter-satellite autonomous punctuality requirements, requiring no more than one transfer node between any satellite [4].

2.2 Transmission Requirements of Chain Building

Interstellar link data transmission and control of the main betting, remote telemetry and other information between the distribution and return of the interstellar, focusing on the ground visible satellites and invisible satellite transmission between the issue of information, taking into account the real-time remote control information The requirements are high, requiring that the number of hops between the visible satellite and the invisible satellite should be as small as possible, and the one-hop connection should be established as much as possible.

3 Rules of Chain Building Between Stars

Time division system, with flexible time slot allocation [5], each node access network pre-allocated time slots, synchronous access, access slot allocation determines the entire constellation at any time the connection status, a direct impact on application services and network transmission performance [6], access slot allocation and finalization should be achieved through cross-layer optimization design. The relationship between the satellites is calculated by the terrestrial inter-satellite link control center and injected into the satellites in the form of a slot table. The satellites perform the satellites-building operations of different satellites according to the received time slot table.

3.1 Principles of Chain Building

Priority to establish a continuous link visible, co-ordinate the establishment of homochronous and off-track links to ensure that the observation geometry to meet the measurement needs.

(1) Under the circumstance of ensuring the measurement requirement, the link planning should follow the principle of smaller link priority of PDOP;
(2) As far as possible in the territory of the stars—the establishment of links between foreign stars;
(3) The principle of one-hop data transmission using domestic and overseas satellites as far as possible;
(4) Try to ensure that each link-building chain nodes and satellite outside the satellite evenly distributed time slots;
(5) Minimize the inter-satellite link between domestic satellites and domestic satellites.

3.2 Inter-Satellite Link Time Slot Planning and Design

As can be seen from the inter-satellite visibility analysis, each MEO satellite removes currently invisible MEO satellites, plus the currently visible GEO/IGSO high-orbit satellites, with an average of about 20 particles. Among all the visible satellites of each MEO satellite, the satellites can be divided into two types depending on the visibility: continuously visible satellites and non-continuously visible satellites. Among them, 8 satellites are continuously observable with the source satellites during the period and 12 satellites are intermittent with the source satellites visible. Table 1 shows the continuous visible and intermittent satellite link comparison.

Table 1 Continuously visible and non-continuously visible links comparison

	Continuously visible links	Non-continuously visible links
Visibility	Continuously visible	Non-continuously visible
Topology	Basic fixed	In accordance with certain laws at any time change
Number of links	Less	Many
Performance comparison	DOP value changes small, the link selectivity is weak	DOP value changes, the link selectivity
Conclusion	A certain measurement and communication foundation for the link plan is possible	As a complement to time-invariant links, optimize link planning performance

In accordance with the requirements of the construction chain, the need to comprehensively use the continuous visible link and non-continuous visible link. Based on the characteristics of the two links, the architecture design will continue to see the link as the basis for the periodic link planning, and improve the performance of the optimized link planning by using the non-persistent visible link.

(1) Determination of continuously visible links

Visibility analysis shows that in the navigation satellite constellation, there are a total of eight continuously visible satellites for each MEO satellite. Due to the parity of Walker constellations, the spatial characteristics of each MEO are consistent with those of other satellites. Taking the first MEO satellite MEO11 on the first orbital plane as an example, the eight satellites continuously observable with the MEO11 satellite are MEO13, MEO14, MEO16, MEO17, MEO21, MEO24, MEO35 and MEO38, of which four are satellites in orbit, The other four are evenly distributed on the other two orbital surfaces of the MEO satellite.

The following separately from the two perspectives of positioning accuracy and data transmission performance, the performance of different numbers of continuously visible links are analyzed, the positioning accuracy is taken as the basis of the satellite's geometric distribution factor, and the data transmission performance is obtained from the number of forwarding hops of the overseas satellites.

- Geometry factors of links

The geometric factor PDOP (Position Dilution of Precision) is one of the important criteria to measure the positioning accuracy of a positioning system [7]. When the PDOP is too large, there will be a large probability of a large positioning error when positioning the satellite. Due to the close relationship between the PDOP and satellite's spatial distribution, the PDOP value is usually used to measure whether the current satellite has a good geometric distribution [8].

The PDOP values of 8 continuous visible links (4 co-link links + 4 heterogeneous links) and 7 non-continuously visible links (4 co-link links + 3 heterogeneous links) Calculated, the location to take MEO11 satellite. It is calculated that

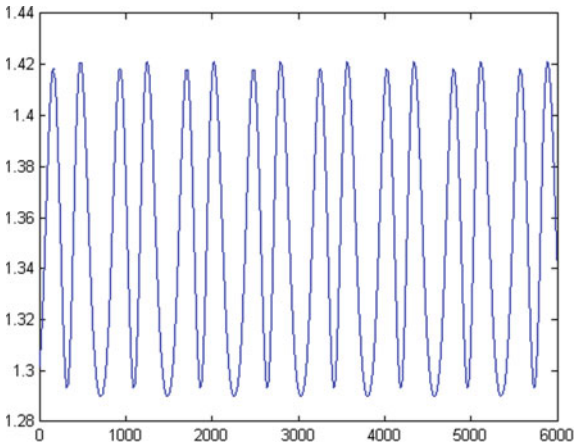


Fig. 1 The PDOP value of the same track 4 and the different track 4

when the number of continuous links is eight, the PDOP value can be maintained at 1.29–1.42, which meets the requirement of PDOP < 1.5. When the time-varying link takes seven, the PDOP value ranges from 1.36 to 2.31 fulfil requirements. Figure 1 shows the calculation results of PDOP values of eight time-varying links in one cycle (6000 ms, the same below), and Fig. 2 shows the calculation results of PDOP values of seven non-time-varying links in one cycle.

If the number of continuous links is too small, the positioning accuracy requirement of PDOP < 1.5 will not be met. Therefore, the number of links that can be continuously seen should be at least 8 from the viewpoint of satisfying the geometric factors.

- Foreign stars for the number of hops

Since the main application data transmission mode ISLs measured shipped overseas distribution and control business to return. From the perspective of digital

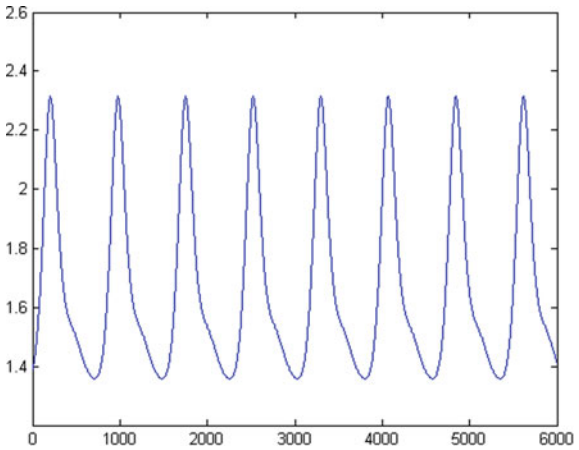


Fig. 2 The PDOP value of the same track 4 and the different track 3

transmission, the fewer the number of times of forwarding, the less delay of data transmission between the stars and the less complicated the route calculation, and the less pressure on the data buffer on the satellite. Therefore, the fewer the number of interstellar forwarding, the better the link design.

On the number of different links on the number of forwarding on the line analysis. The simulation conditions are 24 MEO satellites of Walker constellation, Beijing and Sanya ground stations, 5°–10° elevation angles of the earth station, and a simulation period of one week. Table 2 shows the simulation results of 3/4/8 of the percentage of hops transmitted by both domestic and overseas satellites under continuously visible links.

It can be seen from the table that considering only 3 and 4 continuous visible links, the average number of hops is too high, and the probability of two hops is relatively large; while considering only eight continuously visible links, a hop probability of 99.7%. The probability of two jumps is only 0.3%.

To sum up, the 8 continuously visible links obtain better performance in terms of positioning accuracy and data transmission, and provide a good foundation for the time-slot planning of the inter-satellite links. Therefore, the number of continuously-visible links is taken as 8 Article.

(2) Planning of inter-satellite continuously visible links

Because of the symmetric nature of the Walker constellation, the persistent visible satellites for each MEO satellite are symmetrically distributed, with identical orbital MEO 13 exhibiting exactly the same link attributes as 17, MEO 14, and 16, and the orbital MEO 21 and 38, MEO 24 and 35. The law of change is basically the same. Table 3 shows the attribute values of the continuously visible links of the MEO11 satellite.

The last column of Table 3 shows the probability of non-simultaneous presence of MEO 11 and target satellites in the territory, that is, one of the two satellites is located within the antenna beam range of the earth station and the other is located offshore. According to the above design principle of establishing a domestic-overseas link as

Table 2 Percentage of minimum hops from ground stations to off-shore satellites for 3/4/8 time-varying links

	Beijing + Sanya (elevation $\leq 5°$)	
3 links (2 on the same track + 1 offtrack)	One-hop probability	Two-hop probability
	83.06%	16.94%
4 chains (2 on the same track + 2 offtrack)	One-hop probability	Two-hop probability
	92.94%	7.06%
8 chains (4 on the same track + 4 different track)	One-hop probability	Two-hop probability
	99.7%	0.3%

Table 3 MEO11 satellites continuously visible link attributes at a glance

Satellite No.	Distance (km)	Pitch angle (deg)	Probability of not exist at home at the same time (%)
M13,17	37,304	45.0	62
M14,16	48,740	67.5	90
M21,38	29,466–47,562	34.1–64.4	66
M24,35	32,386–49,225	38.0–69.4	75

much as possible, the link with the greater probability P is used with a higher priority than the link with a smaller probability.

These links are classified. The link between MEO11 and MEOs 13 and 17 and other MEOs is called link A. The link corresponding to MEOs 14 and 16 and other MEOs is called link B. Corresponding to MEOs 21 and 38 and other MEOs Is called link C, and links corresponding to MEOs 24 and 35 and other MEOs are referred to as links D. FIG. Figure 3 reflects the topology of the same-link links A and B. It is not difficult to see from Fig. 4 that the on-link links A (blue) and B (red) in the continuously visible links are all typical star topology. According to the principle of priority-building of high-quality links, the eight links should be evenly discharged into the interstellar-linked timeslots as much as possible.

(3) Planning of inter-satellite time-varying links

Since there are eight consecutively visible links in a slot cycle, the remaining is a non-continuously visible link. The main difference between a continuously visible link and a non-continuously visible link of a satellite lies in that the persistent link can be permanently seen, but the non-continuously visible link exists only in part of the period. Therefore, for non-continuously visible links, Visibility is one of the key evaluation factors of its advantages and disadvantages.

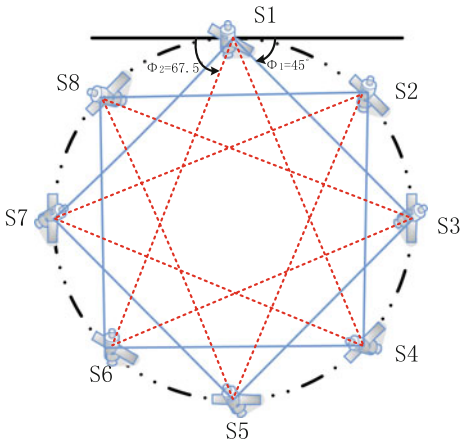


Fig. 3 Topology of the same track in a time-varying link

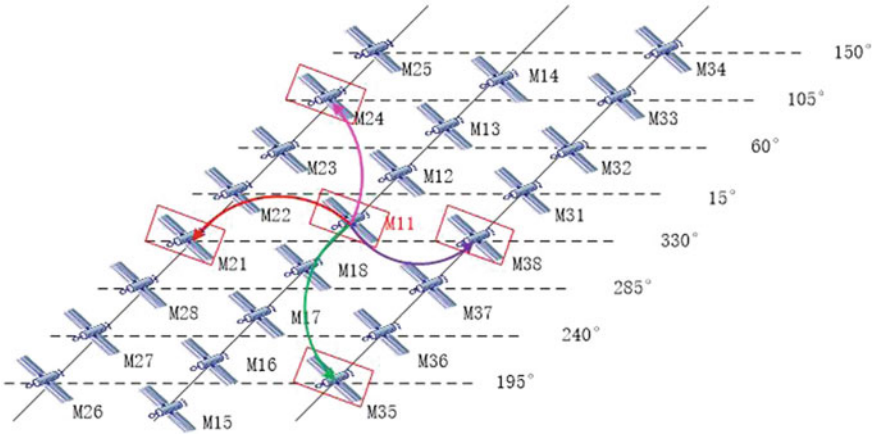


Fig. 4 Topology of a non-transitory link in a non-time-varying link

The same MEO11 example MEO satellite non-continuously visible link, as shown in Table 4. It can be seen that both the probability that a satellite is not in the same territory within a time period is basically the same as the change rule of its probable length of time, and the link with greater probability in different occasions can be seen for a longer time, i.e., the link quality is higher, low.

From the above analysis, it can be concluded that the method for planning a non-continuously visible link can be similar to the method for planning a continuously visible link, and follows the principle of prioritizing a high-quality link. However, since the non-continuously visible link is not always visible and the rate of change Fast, unstable, should try to meet the domestic star - outside the principle of building the chain star. In order to increase the distribution of satellites in China, more domestic and foreign links will be established and based on the non-continuously visible MEO domestic and foreign links, the use of high-orbit satellites will be increased to establish domestic-overseas links.

According to the principle of inter-satellite links, we can make a comprehensive planning of continuously visible and non-continuously visible links to get a better inter-satellite chain planning.

Table 4 MEO11 satellite non-continuously visible link attributes list

Satellite no	Distance (km)	Pitch angle (deg)	Probability of not exist at home at the same time (%)	Visible time (%)
M22,37	35,974–51,195	43.0–76.0	77	75
M23,36	36,923–51,195	44.5–76.0	80	69
M25,34	26,378–43,701	30.0–55.9	66	67
M26,33	26,378–38,667	30.0–47.1	55	47
M27,32	26,378–37,670	30.0–45.6	42	44
M28,31	26,378–41,676	30.0–52.2	52	58

4 Routing Planning of Inter-Satellite

The routing table specifies the data transmission path from the source node to the destination node. The design is based on the planning of the link establishment. Taking into account the working characteristics of the time division system, the routing plan follows the following principles:

Static routing principles. The relations between the stars and the dynamic changes of the connections make the static route be used to reduce the difficulty in the inter-satellite route design. The route does not change during the effective time of the time slot table, that is, a time slot table corresponds to a routing table.

The principle of minimum delay. Choose the most massive route from outside to the inside of China and ensure the minimum delay between domestic and overseas destinations so as to achieve one-hop data transmission between domestic and overseas satellites.

On the basis of time slot planning, according to the principles of routing planning, better routes can be obtained. Slot table, routing table design is as follows.

4.1 Time Slot Table Design

The slot table is used to indicate the order in which each satellite is linked to other satellites. Taking 24 MEO satellites as an example, each time slot table row represents satellites, columns represent time slots, and each row represents the order in which each satellite is linked to other satellites in a time slot cycle. The data in the cells in the table Significance: 1–30 represent the satellite number, of which 1–24 for the medium-range satellites, 25–30 for the high-orbit satellites, −1 for the satellite refers to, 99 for standby. In order to highlight the table can be yellow on behalf of domestic stars, white represents the outer stars, blue represents two-hop transmission.

4.2 Routing Table Design

The routing table is generated from the slot table, indicating whether each satellite can transmit data with other satellites. Each routing table also uses rows to represent satellites, and columns represent time slots. When the content in a cell is not 0, it indicates that data transmission is possible between two satellites. When the content in the cell is 0, it indicates two Data can not be transmitted between satellites. It is highlighted that the blue shades in the table represent transit stars with two-hop routes when overseas satellite data are transmitted back.

5 Network Performance of Inter-Satellite

5.1 Inter-Satellite PDOP Value

In the time division system, MEO, GEO and IGSO are considered as peer nodes. The PDOP values of all satellites are calculated based on a chain-building schedule (one table per minute) of the regression period. The results show that the PDOP ranges from 0.47 to 1.42, PDOP < 1.5 measurement requirements. MEO11 which PDOP value of 0.78–0.98, as shown in Fig. 5.

5.2 Transmission Delay Between Stars

Based on the full constellation orbit data of a regression period and the location of the earth stations, a slot table and a routing table (one table per cycle) are obtained according to design principles and ideas. Regardless of on-board data processing time, statistics of any one two-way time slot, any one of the outer satellite to the node star's end-to-end transmission delay, the results shown in Table 5.

It can be seen that the probability of the data transmission delay of the inter-satellite link network is less than or equal to 3 s is 84.91%, the probability of less than or equal to 6 s is 98.47% and the probability of less than 9 s is 100%, and the average transmission delay of the entire network is 3.5 s, satisfying the data Transmission delay needs.

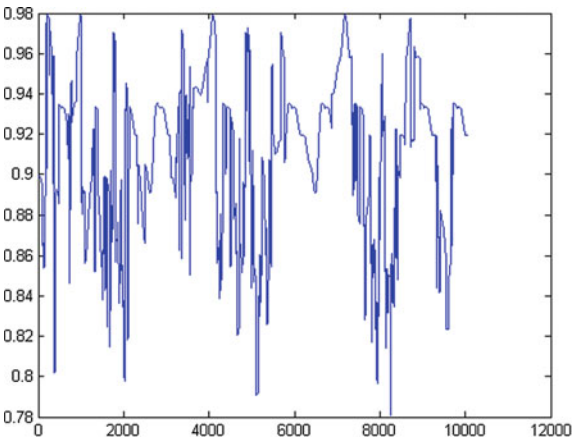


Fig. 5 MEO11 PDOP changes during a return period

Table 5 Analysis of network delay performance

Delay (s)	Probability (%)
≤ 3	84.91
≤ 6	98.47
≤ 9	100

6 Conclusions

In this paper, under the system of concurrent interspaces and inter-duplex links, this paper proposes the optimal link-building scheme based on the requirements of ISM for inter-satellite link: Based on the time division system, 8 pieces per MEO are continuously visible The links are uniformly programmed into their time slot periods. The remaining non-continuously visible links of the time slots are prioritized with high probability of not being in the same territory at the same time, and an increase of the use of high-orbit satellites to establish an inbound-outbound link is established. Then, And routing principles to establish the constellation of each star in the time slot table and routing table to complete the pre-assignment of each node in the inter-satellite chain. Based on the data of the entire constellation orbit of a regression period and the location of the earth stations, the slot tables, routing tables and statistics of the PDOP values of all the satellites are obtained according to the above scheme. The results show that the PDOP ranges from 0.47 to 1.42 and meets PDOP < 1.5. Measurement requirements. Regardless of the processing time of the data on the satellite, an arbitrary two-way time slot is counted, and the average transmission delay of the entire network is 3.5 s, meeting the requirement of data transmission delay.

References

1. Dang Y, Min J, Cheng Y (2007) Global navigation satellite system principles and applications. Beijing: Survey and Mapping Press
2. Guo F, Zhang X (2009) GALILEO system simulation and analysis based on STK software. Geomatics Inf Sci Eng 34(1):15–19
3. Eissfeller B, Zink T, Wolf R (2000) Autonomous satellite state determination by use of two-directional links. Int J Satell Commun 18(1):325–34
4. Menn M (1986) Autonomous navigation for GPS via crosslink ranging. In: Proceedings of the IEEE position, location, and navigation symposium, Las Vegas, NV, 4–7 Nov 1986, p 143–146
5. L Fan, Y Zhang (2007) Regional coverage hybrid constellation design. Aerosp Control 25 (6):52–55
6. He JF, Jiang Y (2009) A walker constellation network topology and route generation scheme with off-track Inter–Satellite link. PLA University of Science and Technology, Nanjing
7. Xiang J (2007) Satellite constellation configuration control and design. National University of Defense Technology, Changsha
8. Zhang T, Chongkan L (2006) Research on shortest path algorithm in satellite time-varying network topology. J Comput China 29(3):71–377

High-Quality BDS Navigation Signal Simulator Based on GPU Optimized Design

Lei Wang, Xiaomei Tang, Baiyu Li, Yangbo Huang and Feixue Wang

Abstract High-quality navigation signal simulator plays an important role as the verifying method for the function and capacity of navigation receivers. With the development of GPU, software navigation signal simulator based on CPU + GPU is developing rapidly because it's easy to transplant, update and maintain. However, current signal generation algorithms are not fully optimized. The generalized signal generation architecture based on GPU is built in this paper, and optimized GPU thread design is obtained by theoretical analysis and testing based on the architecture. The test results prove that the optimized algorithm is able to generate multiple channels of high-dynamic and high-quality GNSS signals. The time consumption is separately 93.6% and 1.73–7.73% less than the two methods before and the optimized capacity does not change while number of signal channels change. The optimized algorithm is already applied in BDS satellite simulator system.

Keywords GPU · GNSS signal simulator · CUDA · High-dynamic

1 Introduction

Software Defined Radio (SDR) has been widely used in simulation and testing for navigation system, because of its convenience in transplanting, upgrading and maintaining. Before GPU is widely used, navigation signal simulator and receiver based on CPU have been researched deeply. In paper [1] the algorithm of signal simulator is described. The software signal simulator and software receiver verified the function of each other. With the development of general purpose GPU,

L. Wang · X. Tang · B. Li · Y. Huang · F. Wang (✉)
School of Electronics Science, National University of Defense Technology, Changsha 410073, China
e-mail: wangfeixue365@sina.com

L. Wang
e-mail: w_lei_81@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_5

calculating signal samples by GPU is of great advantage because GPU is composed of multiple threads working parallel. Software navigation signal simulator based on CPU + GPU is developing rapidly.

Paper [2] described the direct digital synthesizer (DDS) algorithm of calculating the code phase and carrier phase in software navigation signal simulator. Signal samples are allocated to different threads of the GPU to calculate, and multiple samples can be calculated through looping. However, in this paper, the simulating of one signal is launched to one SM (Stream Multi-processor), with other SMs waiting. In the following part, this will be referred to as method one.

Paper [3] proposed an optimised algorithm for signal simulator based on GPU. In this algorithm, multiple signal channels are allocated to one BLOCK of GPU to calculate, so signal collaboration can be done through shared memory. This algorithm provides better performance when there are many signal channels. However the block design depends on the channel number, and when the channel number is between 8 and 16, the number of active warps is not optimal. In the following part, this will be referred to as method two.

Based on the research before, this paper analysed the optimising method of CUDA based GPU algorithm. The usage of varieties of memory is optimised. A generalized signal generation architecture based on GPU is established and optimal parameters are chosen aiming to least time consume calculating one channel of signal. The high-dynamic and high-quality BDS navigation signal simulator is build based on research result and it has been used in BDS satellite simulator system.

2 Structure of Software Simulator

The flow chart of the software simulator is showed in Fig. 1. It generates high-dynamic and high-quality digital signal. Simulator starts at 1PPS and calculates 1 s of signal samples. So the end time t should be less than 1 s. Since the software may work on different devices, the optimization target is to generate one channel of signal in least time and be able to generate multiple channels of signals on high-performance devices.

3 Signal Model

Signal simulator is the inverse process of receiver. Firstly, generate signal dynamic model. Secondly, generate the signal according to the dynamic model.

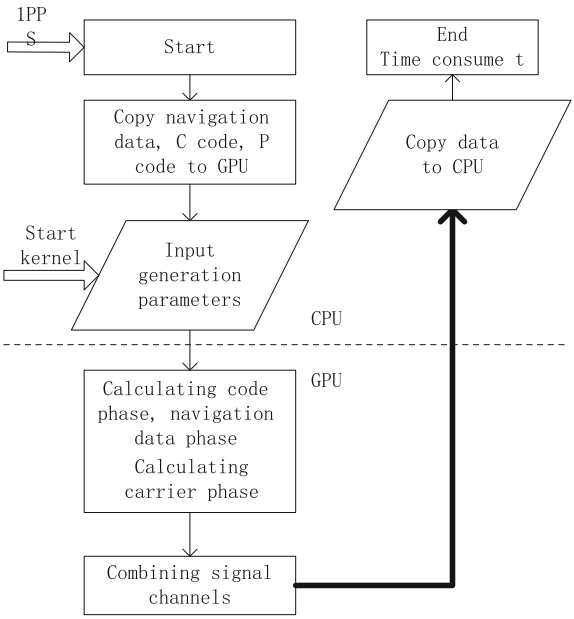


Fig. 1 Flow chart of software simulator

3.1 Mathmatical Model

BDS navigation system is composed of B1, B2, B3 signals. The algorithm is introduced considering B1. B1 consists of B1A, B1C and B1I. The signal expression is showed below:

$$s(t) = A(t)\left(\begin{array}{c}\left(S_{B1A\_data}(t-\rho_p)SC_{\cos}(t-\rho_p) + S_{B1A\_pilot}(t-\rho_p)SC_{\sin}(t-\rho_p)\right)e^{j\pi/4} \\ + \left(S_{B1C\_data}(t-\rho_p) + jS_{B1C\_pilot}(t-\rho_p)\right) + S_{B1I}(t-\rho_p)e^{j\omega_0(t-\rho_c)}\end{array}\right)e^{j(\omega(t-\rho_c)+\varphi)} \quad (1)$$

where s(t) is the signal, t is simulation time, A is amplitude, $S_{B1A\_data}(t)$, $S_{B1A\_pilot}(t)$, $S_{B1C\_data}(t)$, $S_{B1C\_pilot}(t)$, $S_{B1I}(t)$ are separately the base band expression of B1A data, B1A pilot, B1C data, B1C pilot and B1I and they are expressed in Eq. (2) below, $SC_{\cos}(t)$, $SC_{\sin}(t)$ are separately the cos subcarrier and sin subcarrier and they are expressed in Eq. (3) below, $e^{j(\omega t+\varphi)}$ is the carrier in which ω is the carrier frequency and φ is the carrier phase, ρ_p, ρ_c are separately the code propagation delay and carrier propagation delay, $e^{j\pi/4}$ indicates the phase difference between B1A and B1C, $e^{j\omega_0 t}$ indicates the frequency difference between B1I and others.

$$\begin{aligned} S_{B1A\_data}(t) &= D_{B1A\_data}(t)C_{B1A\_data}(t) \\ S_{B1A\_pilot}(t) &= C_{B1A\_pilot}(t) \\ S_{B1C\_data}(t) &= D_{B1C\_data}(t)C_{B1C\_data}(t) \\ S_{B1C\_pilot}(t) &= C_{B1C\_pilot}(t) \\ S_{B1I}(t) &= D_{B1I}(t)C_{B1I}(t) \end{aligned} \tag{2}$$

$$\begin{aligned} SC_{\cos}(t) &= sign(\cos(2\pi f_{SC}t)) \\ SC_{\sin}(t) &= sign(\sin(2\pi f_{SC}t)) \end{aligned} \tag{3}$$

In Eq. (2), D is the navigation data bit, and C is the pseudo code.
In Eq. (3), f_{SC} is the frequency of the subcarrier.
Based on the expressions above, the discrete signal model is expressed below:

$$S = \sum_{n=0}^{num} \sum_{i=1}^{satnum} s_i(t)\delta(t - n\Delta t) \tag{4}$$

where S is the discrete signal, *num* is number of samples, *satnum* is the channel number of signal, $\delta(t)$ is the Dirac delta function, Δt is sampling interval, $s_i(t)$ is the ith channel of signal.

3.2 Signal Dynamic Model

A 3-order interpolation polynomial is adopted to express signal dynamic, and the expression is as below:

$$\begin{aligned} \rho_p &= \rho_{p0} + \rho_{pv}(\Delta t) + \rho_{pa}(\Delta t)^2/2 + \rho_{pj}(\Delta t)^3/6 \\ \rho_c &= \rho_{c0} + \rho_{cv}(\Delta t) + \rho_{ca}(\Delta t)^2/2 + \rho_{cj}(\Delta t)^3/6 \end{aligned} \tag{5}$$

where Δt is the sampling interval, ρ_{p0}, ρ_{pv}, ρ_{pa}, ρ_{pj} are separately the start value, change rate, 2-order change rate, 3-order change rate for code propagation delay, ρ_{c0}、ρ_{cv}、ρ_{ca}、ρ_{cj} are the corresponding values for carrier propagation delay.

4 Optimization of GPU Algorithm

The hardware platform adopted in this paper is shown in Table 1.1. For specification, Tesla K80 is composed of two GPUs.

The generalized signal generation architecture based on GPU is established in Fig. 2.

Table 1 Hardware platform

CPU:Xeon E52640		GPU:K80(one GPU)			
Number of cores	16	Compute capability	3.7	Number of register	65,536
Clock rate	2.60 GHz	Number of SMXs	13	Shared memory per block	16–48 KB
Memory	32 GB	Number of SPs	2496	Clock rate	0.82 GHz

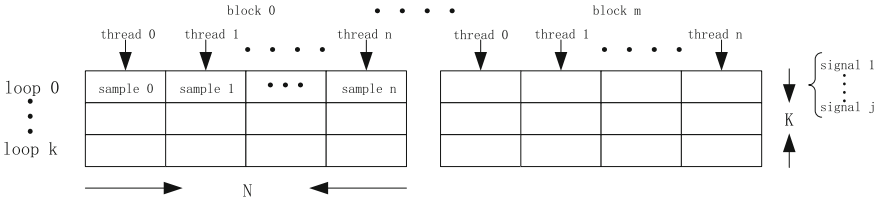


Fig. 2 Generalized signal generation architecture

M, N and K are used to separately express block numbers, thread numbers per block, loop times for every thread and they are bounded by Eq. (6).

$$\begin{cases} temp = \lceil num/K \rceil \\ M = \lceil temp/N \rceil \\ mKN + kN + n < num \\ 0 \le m \le M, 0 \le k < K, 0 \le n < N \end{cases} \tag{6}$$

where $\lceil\ \rceil$ means the ceil integer. From Eq. (6), it is evident that only K and N are free variables.

Based on the architecture, the discrete signal is expressed as below in Eq. (7).

$$S = \sum_{m=0}^{M-1} \sum_{k=mK}^{(m+1)K-1} \sum_{n=kN}^{(k+1)N-1} \sum_{i=1}^{satnum} s_i(t)\delta(t - n\Delta t) \tag{7}$$

Compared to signal combination through shared memory, it is proposed to do it through registers in this paper. So samples of different signals are calculated in the same thread and added together through registers.

The optimization of GPU algorithm includes memory optimization and parallel optimization. The content is showed below in two sections.

4.1 Memory Optimization

Memory of GPU includes registers, shared memory, texture memory, constant memory, global memory and L1, L2 Cache. The access delay varies greatly. Optimized utilization of GPU memory will improve the calculating efficiency [5].

Considering the data access frequency and data quantity, the memory usage is showed below:

a. Digital signal and P code are stored on global memory because the data size is big;
b. Registers are used to combine signals and to store intermediate variables and global memory pointers;
c. Rational utilization of shared memory is able to improve the processing efficiency. But the following two aspects should be considered, firstly, data is copied from global memory before the kernel start calculating which means that if the access time is small compared to the data size there won't be any advantage; secondly, the size of shared memory is limited, the usage size will affect the number of active warps which is described in detail in Sect. 4.2;

C code and navigation bit is the main concern considering the usage of shared memory. So in the following analysis, there will be three states: *state 1* means that C code and navigation bit is carried to shared memory, *state 2* means that navigation bit is carried to shared memory while C code is not, *state 3* means that both are stored in global memory.

4.2 Parallel Optimization

4.2.1 Thread Parallel

The following aspects are the main consideration in thread parallel optimization:

(a) Make sure there is enough calculation in one block to hide memory access delay. This corresponds to increasing K;
(b) Make sure there is enough active warps to keep SMs busy. Warp is the unit of GPU parallel processing. The only way to hide global memory access delay is launching multiple active warps to SMs [4]. So the block design should correspond to more active warps. This corresponds to the value of N;
(c) The calculation size of different SMs should be nearly averaged in lest to some SMs waiting. This corresponds to M, that is to say K and N;

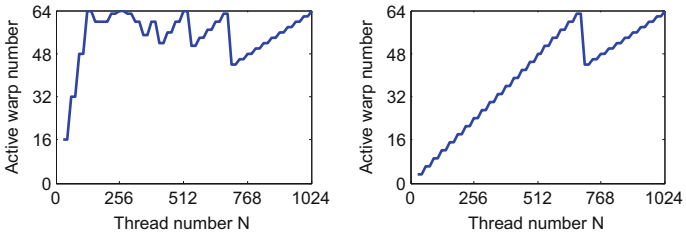


Fig. 3 Number of active warps versus thread number per block N and usage of shared memory

Figure 3 is the number of active warps versus thread number per block N and shared memory usage. State 1 corresponds to the right figure because of large shared memory usage, while state 2 and state 3 corresponds to the left one.

To compare the affection of K, N and shared memory usage, different K and N are chosen to test kernel processing time under three states. The values of N are [128, 256, 512, 1024] to get the most active warps and the values of K are [1, 5, 10, 15, 20, 25, 30, 40, 50, 5860/2930/1465/733]. The last number of K means that the calculation is launched to one block.

When calculating one channel of signal, method one corresponds to K as the last number, while method two corresponds to K as the first number.

The processing time of kernel is tested with the parameters (Fig. 4).

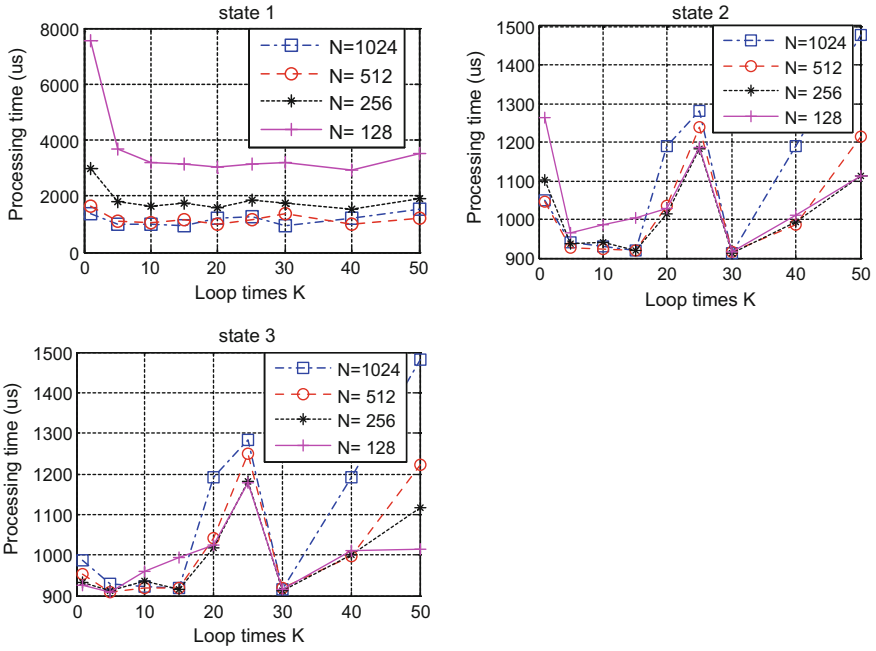


Fig. 4 Kernel processing time under different K, N and states

Table 2 Kernel processing time with one block

State (N = 1024)	State 1	State 2	State3
Processing time (μ)	14245.4	14561.5	14738.8

Table 2 is the kernel processing time under different states using method one. From Fig. 4, some conclusions can be got:

(a) Processing time is not proportional to K or N because they affect the calculation balance among SMs and number of active warps. So test is needed;
(b) Storing navigation bits in shared memory does little affect to processing time, but when N is small, such as 128 or 256, processing time may go up, because copy process consumes time;
(c) Storing C code in shared memory leads to decrease of active warp number and large time consume to copy data. But when calculation is launched to one block, it will lead to less time consumption which is showed in Table 2.

In summary, N = 512 or 128, K = 5 and storing C code and navigation bit in global memory is the best choice for this paper and the processing time is about 910 μ. The processing time of method one is 14245.4 μ for the best design and for method two it is separately 926.1, 931.7, 958.5, 986.1 μ with different N.

So compared to method one, the processing time is 93.6%. Compared to method two, the processing time is 73–7.73% less and number of active warps does not change with channel number of signals so keeping optimized capacity for multiple channels of signal.

4.2.2 Task Parallel

For GPU of Kepler architecture, stream can be scheduled to perform multiple CUDA operations simultaneously. Through streams, kernels are able to overlap data copy delay and kernel processing, saving total time consumption.

The procedure to process streams is showed in Table 3.

Table 3 Procedure to process streams

CudaStreamCreate (&stream[i]);//Create streams
cudaMemcpyAsync (d_ptr, h_ptr, size, cudaMemcpyHostToDevice, stream[i]);//Asynchronous data copy
compute_kernel < ≪grid_size, block_size, 0, stream [i] ≫ > (d_ptr);//Signal computing kernel
cudaStreamDestroy (stream);//Destroy streams

Before using stream, the total time of generating one second length of one channel of B1 signal is 94.1 ms in average, while using stream, the time goes down to 87.5 ms, saving 7.0% of time.

4.2.3 Multi-GPU Parallel

It is known that K80 is composed of two GPUs, and they can work at the same time. So two GPUs double the number of channels simulated real-time. And this principle applies to multiple GPUs.

5 Analysis of Signal Quality

To analysis the quality of the signal generated, the power spectrum is compared to the theoretical data. The tracking result of software receiver is also showed in Fig. 5.

Figure 5 shows that the power spectral density is just the same with theoretical data and the correlation domain is also of perfect property.

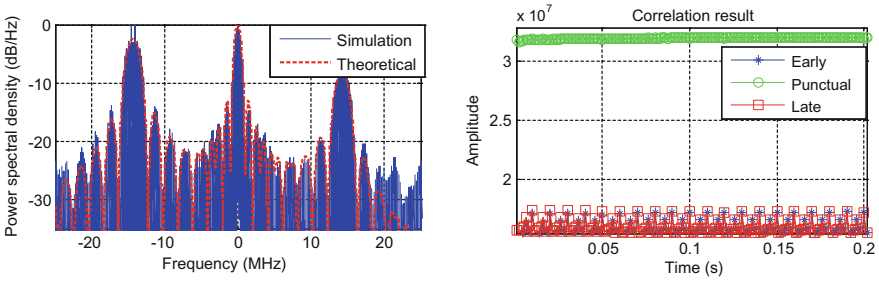


Fig. 5 The power spectrum and tracking result of the simulated signal

6 Conclusion

A high-dynamic and high-quality BDS navigation signal simulator based on GPU is introduced in this paper. Compared to the former algorithms, combining multiple signals through registers is proposed. A generalized signal generation architecture based on GPU is established and optimized parameters can be got through test. Compared to the two methods before, there will be separately 93.6% and 1.73–7.73% capacity improvement, and a better performance for multiple channels of signal. 3-level parallel optimization is introduced. After optimization, 22 channels of B1 signal can be generated real-time on the hardware platform used. The analysis result proves the correctness and quality of the signal.

Acknoeledgements Project 61601458 supported by NSFC.

References

1. Dong L (2003) IF GPS signal simulator development and verification. University of Calgary, Calgary
2. Zhang B., Guang-bin, L., Dong, L. Zhi-liang, (2010). Real-time software GNSS signal simulator accelerated by CUDA. Future computer and communication (ICFCC), 2nd international conference, vol 1, pp 100–104
3. Li Q, Yao Z, Li H, Lu M (2012) A CUDA-based real-time software GNSS IF. China Satell Navig Conf (CSNC) 359–369
4. Rob Farber. CUDA application design and development
5. 高性能CUDA应用设计与开发.于玉龙,唐堃.机械工业出版社, 2013
6. Shane Cook. CUDA programming: a developer's guide to parallel computing with GPUs
7. CUDA并行程序设计GPU编程指南.苏统华,李东,李松泽,译.机械工业出版社, 2016

Verification of Receiver Dynamic Measurement Accuracy Based on Precision Trigger Time Record

Haisong Jiao, He Huang, Chong Zhang, Liangjian Jiang, Qingliang Chen, Feng Li, Dexin Zhang and Li Shui

Abstract In this paper, GNSS receiver dynamic accuracy verification environment is constructed. High precision trigger time recording circuit is designed with the precise time synchronization function of the timing receiver. By measurement and recording of the precise time of the receiver antenna's touch position under the dynamic condition, the "true value" of the location of the receiver measuring time is calculated, the verification of the dynamic measurement accuracy of the receiver is finally realized.

Keywords Trigger time record · Receiver · Dynamic · Accuracy verification

1 Introduction

In recent years, with the continuous improvement and construction of Beidou global navigation satellite system, positioning and velocity measurement accuracy of measurement receiver under dynamic condition is an important assessment index [1]. At present, the test method of outdoor environment is usually adopted for "car comparison method", which is using a high accurate receiver as the "truth value" of dynamic sports car comparison test. In view of the two receivers being measured using satellite navigation system, with a strong correlation error, and comparison of the results can only be the relative accuracy between devices. It is difficult to accurately give absolute accuracy of the receiver in the condition of moving [2]. Therefore, for the verification of the accuracy of the high-precision measurement receiver, if it can be separated from the satellite location and by comparing with the method of providing "truth value" of the space location, it will greatly improve the reliability and credibility of the GNSS receiver dynamic accuracy detection.

H. Jiao (✉) · H. Huang · C. Zhang · L. Jiang · Q. Chen · F. Li · D. Zhang · L. Shui
Luoyang Electronic Equipment Test Center of China, Luoyang 471003, China
e-mail: jiaohs888@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_6

2 Design of Calibration System for Receive Dynamic Measurement

2.1 Design Idea

To give the dynamic measuring accuracy of GNSS receiver, the receiver need real-time measurement of value and dynamic conditions through the true position of a spatial location were compared, and the spatial location of the receiver antenna through a time pulse along the time to accurately determine on the sensor light through the antenna touch photoelectric [3], by the precise time the antenna can be calculated through a measurement position of the "true value". A GNSS receiver through space point measurements and the point of the "true value" comparison, can obtain the dynamic measurement accuracy of the receiver, and Fig. 1 shows the compositions of receiver dynamic precision calibration system and working principle.

2.2 System Composition

The structure of the receiver's dynamic accuracy verification system is shown in Fig. 1. The system consists of a dynamic precision calibration field, a precise time recording system, a differential information transmission link and data processing and display system.

The dynamic accuracy verification field is mainly composed of horizontal straight line, two-dimensional and three-dimensional dynamic accuracy verification field. It mainly completes all kinds of dynamic environment of verification receiver, including different speed and different acceleration state.

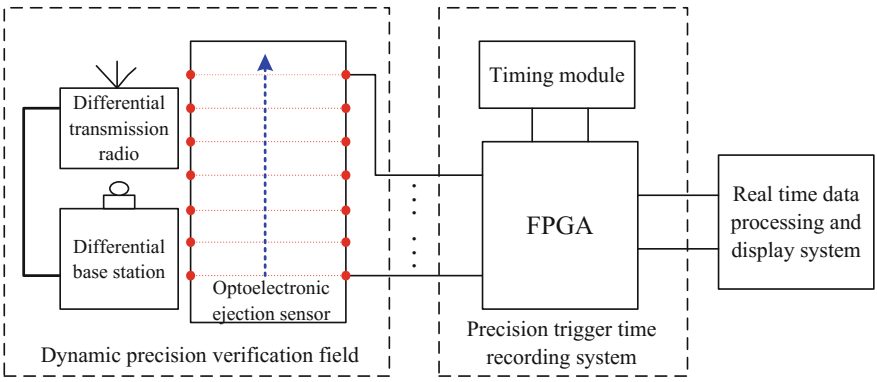


Fig. 1 Compositions and working principle of dynamic precision verification system

The precise trigger time recording system is mainly composed of time service module, FPGA and auxiliary circuit. It mainly completes the measurement of the precise time when the antenna pole of the receiver touches the light when the light is transmitted to the sensor, and it is used to calculate the coordinate "true value" of the real-time data of the receiver. The main timing module for FPGA high precision time synchronization and timing signal, the rich pin resources of FPGA can be pulse output signals of the sensor. when FPGA detects a pin is triggered, It records the number of sensor and trigger time to provide reference for the calculation time to measure the position of "true value".

The differential information transmission link is mainly composed of data transmission radio, antenna and power supply equipment. It mainly completes the transmission of real-time differential information of the detected receiver and the return task of the differential location data, so as to ensure the real-time measurement accuracy of the receiver reaches the best state.

The real-time data processing and display system mainly complete the real-time receiving and processing data, comparing the receiver real-time measurement of coordinates with calculated the "true value", to finish display, statistics and storage real-time measurement error. Finally the measurement accuracy is given in the different state of motion receiver.

3 Verification of Dynamic Measurement Accuracy Under Receiver Dynamic Condition

3.1 Verification of Dynamic Measurement Accuracy Under Horizontal Linear Motion Condition

3.1.1 The Setting and Verification Principle of Horizontal Linear Motion Conditions

The construction of a horizontal straight line rail and rail cars, respectively in both sides of the track at intervals in each vertical bar, for placing a shot of the photoelectric sensor being installed detection receiver, antenna and antenna rod car on the track, when placing receiver rail car to set speed in orbit, it will touch the photoelectric receiver antenna rod the sensor of rays, and sensor receiver will produce a pulse, up or down time. This can be recorded through precise trigger circuit of pulse time recording. That is to say, the recording time is the time to touch the receiver antenna through precise point, Fig. 2 shows GNSS receiver linear dynamic accuracy test chart.

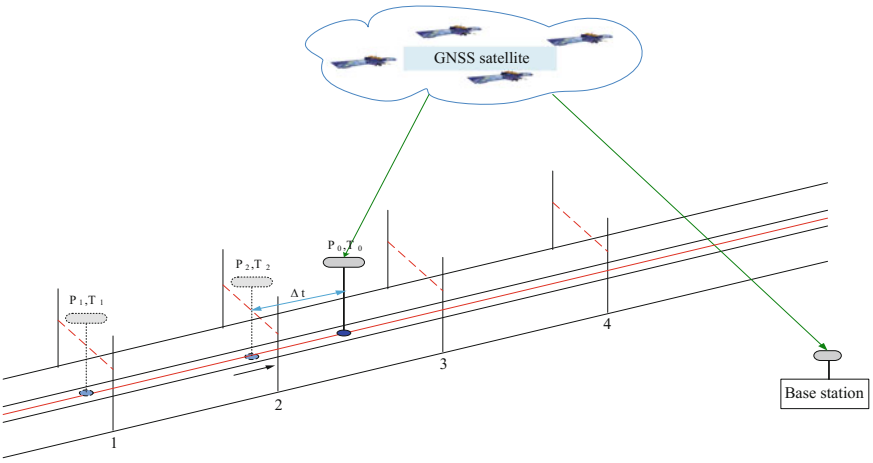


Fig. 2 GNSS receiver horizontal linear dynamic accuracy test

3.1.2 Measure the "Truth Value" of Position

The exact coordinates of the antenna touch position can be obtained by static real time carrier phase difference or geodetic coordinate measurement, and the time of the receiver's output data is very difficult to coincide with the antenna touch time under the limitation of the receiver's data update rate. Therefore, the "real value" of measurement position is obtained through touch position coordinates and two position records the time difference ΔT and X, Y direction velocity multiplied. When the antenna is moving uniform linear motion in a straight line, from position P1 over position P2 to P0 point, the change rate V_X, V_Y, V_Z are fixed in the X, Y, Z three directions, as shown in Fig. 3 [4, 5], among:

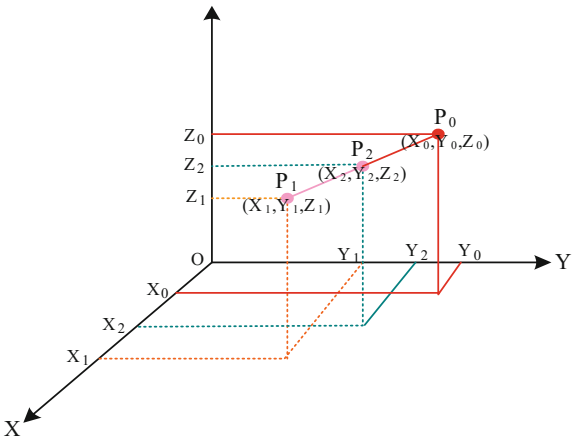


Fig. 3 Coordinate change analysis diagram of Receiver horizontal linear motion

$$\begin{cases} V_X = \frac{X_2 - X_1}{t_2 - t_1} \\ V_Y = \frac{Y_2 - Y_1}{t_2 - t_1} \\ V_Z = \frac{Z_2 - Z_1}{t_2 - t_1} \end{cases} \quad (1)$$

It is calculated by the P_2 point coordinates (X_2, Y_2, Z_2) that the P_0 point true value coordinates (X_0, Y_0, Z_0) is:

$$\begin{cases} X_0 = X_2 + V_X \times (t_0 - t_2) = X_2 + \frac{(t_0 - t_2)}{(t_2 - t_1)} \times (X_2 - X_1) \\ Y_0 = Y_2 + V_Y \times (t_0 - t_2) = Y_2 + \frac{(t_0 - t_2)}{(t_2 - t_1)} \times (Y_2 - Y_1) \\ Z_0 = Z_2 + V_Z \times (t_0 - t_2) = Z_2 + \frac{(t_0 - t_2)}{(t_2 - t_1)} \times (Z_2 - Z_1) \end{cases} \quad (2)$$

3.2 *Verification of Dynamic Measurement Accuracy Under Two—Dimensional Motion Conditions*

3.2.1 The Setting and Verification Principle of Horizontal 2D Motion Conditions

The horizontal two-dimensional motion environment is mainly composed of the grating matrix, the receiver and the motion carrier. Grating matrix crossbar multiple photoelectric correlation sensor, with the level of linear motion is similar in principle. when the receiver antenna rod contact movement bump against the rays, the precise time record system records the trigger pulse of the photoelectric sensor rising or falling along the precise time. The receiver antenna rod can be calculated from the exact position in the matrix by the precise location of sensor being triggered. Figure 4 shows the test principle of GNSS receiver level two-dimensional dynamic accuracy.

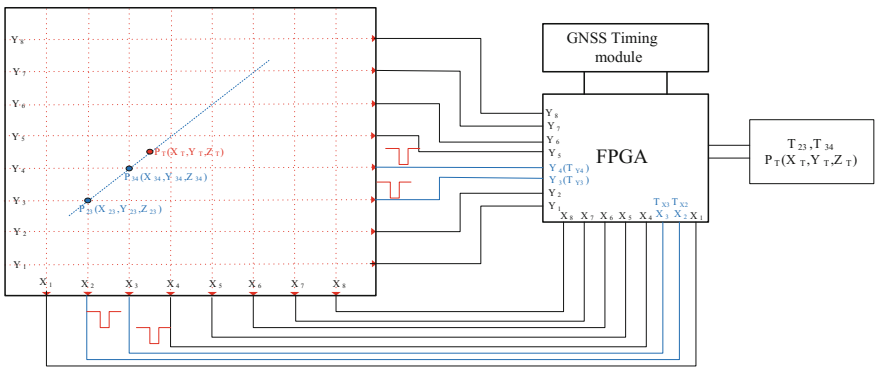


Fig. 4 Test principle of GNSS receiver horizontal 2D dynamic accuracy

3.2.2 Measure the "Truth Value" of Position

The exact coordinates of the antenna touch position can be obtained by the static real time carrier phase difference or the position coordinates of the corresponding sensor. The time of the receiver's output data is very difficult to coincide with the antenna touch time under the limitation of the receiver's data update rate. Therefore, the "true value" at the measurement location is achieved by the coordinate value of the touch position plus the time difference between the two location records and the speed of X, Y and Z directions. When the antenna matrix moves uniformly at a straight line, the antenna goes from position P_{23} to P_{34} point to PT point, and the rate V_X, V_Y, V_Z of change in X, Y, Z of three directions is fixed. the analysis is shown in Figs. 3 and 5. Figure 5 shows the coordinate change analysis diagram of receiver Horizontal two-dimensional motion, among:

$$\begin{cases} V_X = \frac{X_{34} - X_{23}}{t_{34} - t_{23}} \\ V_Y = \frac{Y_{34} - Y_{23}}{t_{34} - t_{23}} \\ V_Z = \frac{Z_{34} - Z_{23}}{t_{34} - t_{23}} \end{cases} \tag{3}$$

The true value coordinates (X_T, Y_T, Z_T) of P_T point are calculated by the P_{34} point coordinates (X_{34}, Y_{34}, Z_{34})

$$\begin{cases} X_T = X_{34} + V_X \times (t_T - t_{34}) = X_{34} + \frac{(t_T - t_{34})}{(t_{34} - t_{23})} \times (X_{34} - X_{23}) \\ Y_T = Y_{34} + V_Y \times (t_T - t_{34}) = Y_{34} + \frac{(t_T - t_{34})}{(t_{34} - t_{23})} \times (Y_{34} - Y_{23}) \\ Z_T = Z_{34} + V_Z \times (t_T - t_{34}) = Z_{34} + \frac{(t_T - t_{34})}{(t_{34} - t_{23})} \times (Z_{34} - Z_{23}) \end{cases} \tag{4}$$

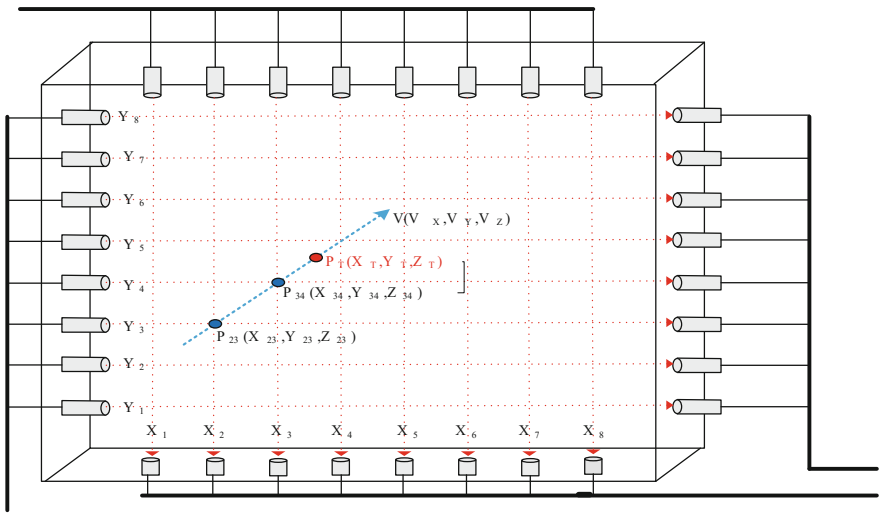


Fig. 5 The coordinate change analysis of receiver horizontal 2D motion

3.3 *Verification of Dynamic Measurement Accuracy Under Three—Dimensional Conditions*

The construction of an inclined orbit, respectively in a certain interval at both sides of the track in each vertical bar, is placed on the radio for photoelectric sensor and detection receiver and antenna are mounted on the vehicle of track. when the rail car on the track moves accelerately from up to down, the receiver antenna rod will touch the photoelectric sensor to shoot the light. In the same way, the precise time record system records the trigger pulse of the photoelectric sensor rising or falling along the exact time. The accurate coordinate of antenna touching position through the static time difference carrier phase or the earth coordinate measurement can be got. The measurement position of the "truth value" through touching position coordinates and the time difference of the two positions recording is multiplied to get. The calculation method of "truth value" in the position of measurement is the same to Sect. 3.2.2. Figure 6 is a three-dimensional dynamic precision test chart for GNSS receiver.

4 Evaluation of the Accuracy of Receive Dynamic Measurement Verification

The measurement error of test system including trigger position error, the time synchronization error, the trigger time record of the time error, selection of the sensor response delay time and the calculation error of spatial position measurement time of "true value", how to minimize the error of these factors is the key to ensure the accuracy of dynamic verification system.

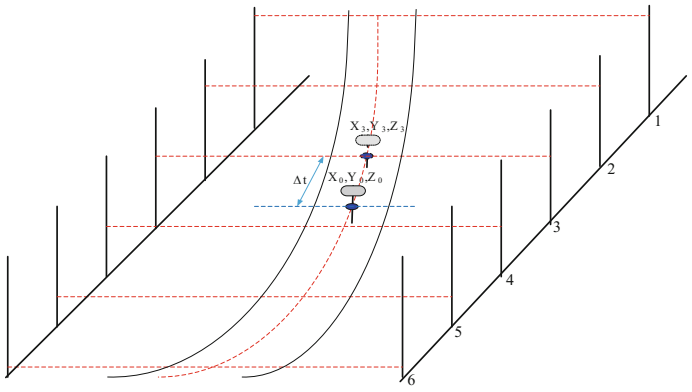


Fig. 6 GNSS receiver space 3D dynamic accuracy test

4.1 Measurement and Analysis of Time Measurement Error

4.1.1 Time Synchronization Signal Error Measurement and Accuracy Evaluation

The measured receiver output 1PPS and precise time record output circuit 1PPS time synchronization signal will inevitably exist certain errors. Two probes measurement error will be measured at the output of the 1PPS receiver and 1PPS precision time record output circuit in the end respectively connected to the oscilloscope. It shows the filter is the synchronization accuracy between two machine time. Figure 7 shows the 1PPS error test chart for NovAtel RT2 receiver 1PPS synchronization signal and precision time recording circuit, which is output by UM220 timing card. From Fig. 7, we can see that the error of two synchronization signals is 476 ns.

4.1.2 Time Delay Error Measurement and Accuracy Evaluation

The key of GNSS receiver antenna rod laser light sensor trigger accurate time is calculated to measure the position of the receiver "true value". The time error measurement including response time, sensor trigger signal rising or falling, the accuracy of signal transmission along the time delay, the precise time recording circuit in the signal processing and recording delay etc. Only, precise calculation and evaluating the error can influence the degree to ensure that the system is in the use of precision.

For the measurement of time delay, it can use the oscilloscope comparison method. The signal output terminal of the sensor is connected with the input end of

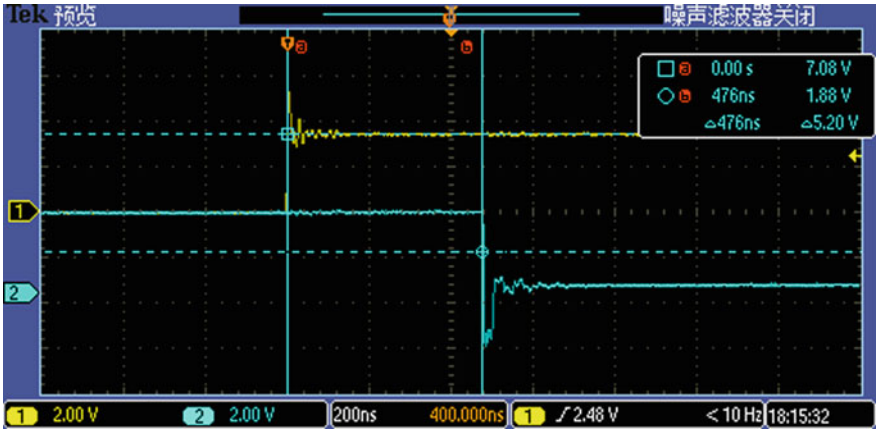


Fig. 7 UM220 OEM plate 1PPS (yellow) and RT2 receiver 1PPS (blue) signal delay

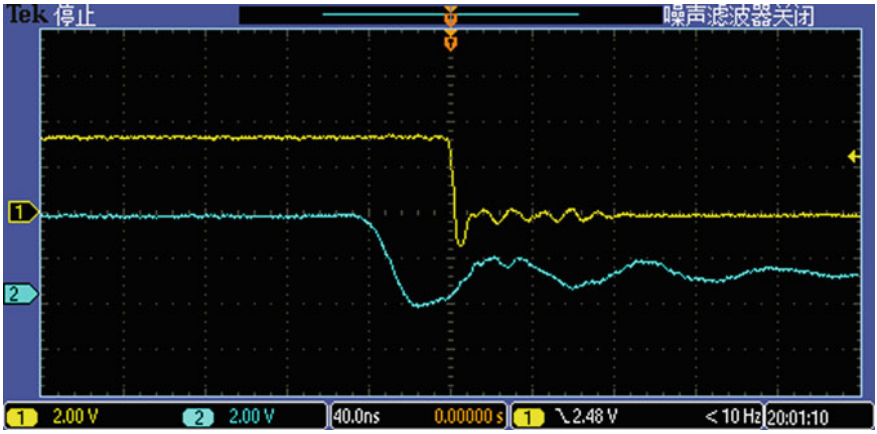


Fig. 8 The sensor (blue) triggers the pulse and the FPGA pulse (yellow) signal delay

the trigger time precision recording system. The signal output and precise time recording system FPGA circuit two are respectively connected to the sensor probe of oscilloscope. The oscilloscope display is the trigger signal in precise trigger time record delay circuit signal processing. As shown in Fig. 8, the time delay is about 40 ns; the output signal will be the output end of the 1PPS and the laser sensor receiver terminal simultaneously accessing the oscilloscope two probes. Photoelectric trigger sensor in the receiver integer seconds time displayed on an oscilloscope is the photoelectric sensor triggering time delay error. By the measured time delay error multiplying the motion speed of the equipment the system error and the degree of influence of the timing can be obtained.

4.2 Calculation Error and Evaluation of "True Value" of Point Position Measured by Dynamic Receiver

When measuring, precise coordinates of antenna touching position is through the static time difference carrier phase or the earth coordinate measurement to get, existing the measurement errors; Because of data update rate restrictions, it is difficult to achieve real-time measurement position and touch position receiver match. Therefore, the "true value" of measurement position is through the touch location coordinates plus two position record time difference ΔT and motion speed to the multiplication to get, existing error of speed and time on the measurement of motion.

4.3 Correction Method of System Measurement Error

In order to further reduce the synchronization error, the precision time recording circuit timing module using as far as possible to choose the timing module and receiver synchronization accuracy to be tested; The correction of measuring time delay error, makes the signal accurately calculated the delay time by post-processing to eliminate and improves the system verification accuracy; the "true value" static error of dynamic receiver measurement's point can be corrected when repeated measurements. Ensuring the stable motion state and constant speed is one method to ensure the accuracy of methods.

5 Conclusions

With the continuous improvement of the Beidou measurement system and the diversity and complexity of the application of the measurement receiver, the accuracy of the receiver under the moving condition will also draw more and more attention from users. The verification system completed, will further improve the performance of GNSS receiver, detection methods and means, improve the user at the receiver in the field of application in the dynamic reliability and credibility. With the continuous improvement of detection technology and means, the accurate verification techniques and methods of dynamic precision of air high dynamic receiver will also be expected to be realized gradually.

References

1. Jiao H (2007) The test and analysis of GPS survey precision. Electronic Science and Engineering, Degree of Master, National University of Defence Technology, Changsha
2. Du J, Ma H, Yao F (2012) Testing and analyzing for accuracy of high precision compatible receiver of BDS. J Electron Measur Technol 36(5):97–100
3. Qiu T (2015) Design of large-screen interactive embedded system based on infrared sensor. J Mod Comput. https://doi.org/10.3969/j.issn.1007-1423
4. Li T (2008) GPS principle and application, 2nd edn. Science Press, Beijing, pp 16–19
5. Li M, feng B (2006) GPS positioning technology and application, 2nd edn. Science Press, Beijing, pp 18–33

Research on High-Precision Test Method of the Navigation Satellite Inner Multipath Effect

Songtao Huangfu, Bin Liu, Xin Nie, Hailong Xu and Tianxiong Liu

Abstract A similar elevation-dependent pseudoranging anomaly was appeared on BeiDou and GPS navigation system. Furthermore, anomalies seem to be present on signals from GPS SVNs 41, 43, 44, 49, 55, 58, 59, 60, and 61. This effect is caused by signal reflections coming from the impedance mismatch of a special auxiliary port on SVN 49. It is suspected that other satellites which show a pseudoranging anomaly have the same reason. In view of this situation, a high-precision test method is present to evaluate the inner multipath effect on navigation satellite in this article. The test method builds the relationship between the satellite antenna array's inner-outer ring delay difference (which caused by the multipath effect) and the beam marginal-subastral point pseudorange measurement divergence. The method is applied to the BeiDou satellite test, and the test result is given.

Keywords Navigation satellite · Multipath effect · High-precision Test method

1 Introduction

Article [1] describes the phenomenon about the code pseudorange variations of the BeiDou satellite system which result in code phase divergences of more than 1 m. And these code variations are elevation-dependent. A similar elevation-dependent pseudoranging anomaly appears on signals from GPS SVNs 41, 43, 44, 55, 58, 59, 60, 61 which show an anomaly at 1 m level in observation [2]. The most representative is SVN 49, the code pseudorange variations which is elevation-dependent reached a level of about 4 m [3].

S. Huangfu (✉) · B. Liu · X. Nie · T. Liu
Beijing Institute of Spacecraft System Engineering, Youyi Road, Haidian District, Beijing 100094, China
e-mail: huangfust@126.com

H. Xu
Tsinghua University, Beijing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_7

The GPS Wing and its contractors find out the cause of the problem on SVN 49, which brought by the L5 demo payload. The L1 and L2 signal is reflect from the J2 input port which is used to feed the L5 signal, then the receiver see both the direct and the multipath signals.

In order to finding the cause of the problem described by article [1], the mathematical model of multipath signal is built and the simulation results are given. These methods are applied to the BeiDou satellite ground test, and the test result is given.

2 The GNSS Multipath Effect Model

The main effect from multipath signal on the navigation receivers is the tracking error bias leading to the code pseudorange bias. This is highly consistent with the problem described in article [1]. The autocorrelation peak distortion caused by multipath effect is shown in Fig. 1.

The distortion will result in code phase deviation, which will cause the pseudo-range measurement error.

If τ_d is the transmission delay between the main path signal and the multipath signal, A_d is the amplitude attenuation factor of the reflect signal, θ_d is the phase lag. The code phase output with multipath effect of the navigation receiver code tracking is written as follows:

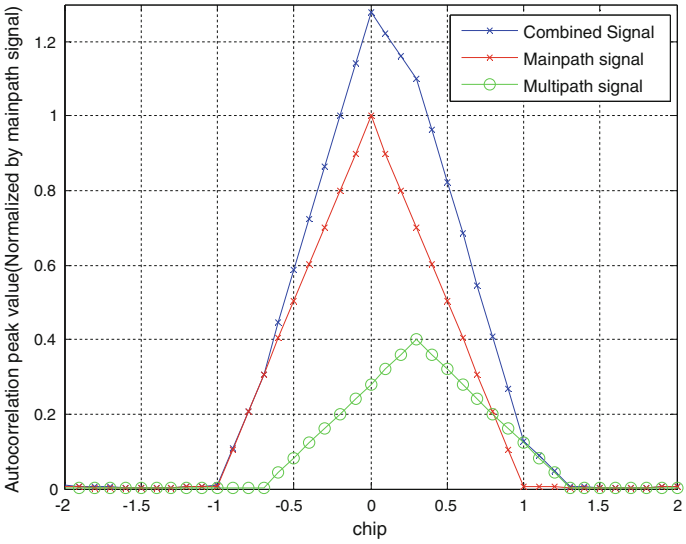


Fig. 1 Autocorrelation peak distortion caused by multipath effect

$$I_E = \frac{A}{\sqrt{2}}[\cos(\psi)R(\tau + d/2) + A_d\cos(\psi + \theta_d)R(\tau + d/2 - \tau_d)] \tag{1}$$

$$I_P = \frac{A}{\sqrt{2}}[\cos(\psi)R(\tau) + A_d\cos(\psi + \theta_d)R(\tau - \tau_d)] \tag{2}$$

$$I_L = \frac{A}{\sqrt{2}}[\cos(\psi)R(\tau - d/2) + A_d\cos(\psi + \theta_d)R(\tau - d/2 - \tau_d)] \tag{3}$$

$$Q_E = \frac{A}{\sqrt{2}}[\sin(\psi)R(\tau + d/2) + A_d\sin(\psi + \theta_d)R(\tau + d/2 - \tau_d)] \tag{4}$$

$$Q_P = \frac{A}{\sqrt{2}}[\sin(\psi)R(\tau) + A_d\sin(\psi + \theta_d)R(\tau - \tau_d)] \tag{5}$$

$$Q_L = \frac{A}{\sqrt{2}}[\sin(\psi)R(\tau - d/2) + A_d\sin(\psi + \theta_d)R(\tau - d/2 - \tau_d)] \tag{6}$$

There are coherent and non-coherent code phase discriminators for delay lock loop architectures of the navigation receivers. A coherent discriminator requires the navigation receiver is phase-locked to the signal. In positioning applications this is unrealistic because of low signal to noise ratio [4], but this is different in the satellite ground test scene because the measured navigation signal's power is much higher than in-orbit test, and there is no Doppler frequency shift. Furthermore, coherent code phase discriminator is more sensitive to the multipath effect. In summary, coherent code phase discriminator is suitable for satellite ground test.

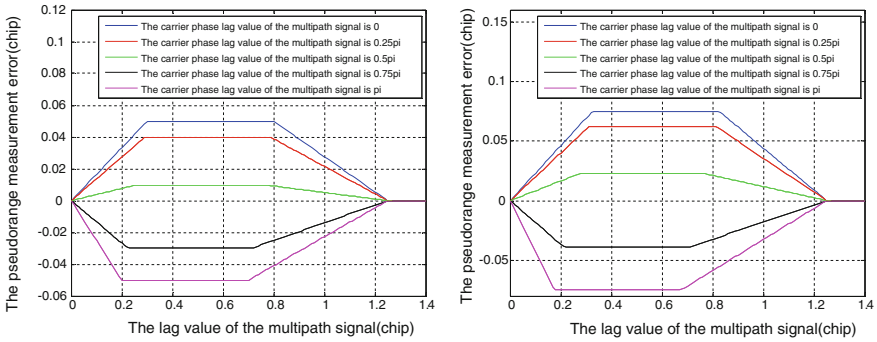
The expression for coherent code phase discriminator is described as follows:

$$S_c = I_E - I_L \tag{7}$$

Referring to Eqs. 1–7, the pseudorange measurement error is solving as Eq. 8 [5].

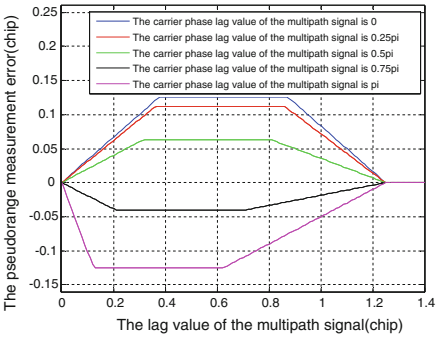
$$\Delta = \begin{cases} \dfrac{A_d \cdot [A_d + \cos(\theta_d)] \cdot \tau_d}{1 + 2A_d \cdot \cos(\theta_d) + A_d^2} & 0 \le \tau_d \le \dfrac{[1 + 2A_d \cdot \cos(\theta_d) + A_d^2] \cdot d}{2[1 + A_d \cdot \cos(\theta_d)]} \\ \dfrac{A_d \cdot [A_d + \cos(\theta_d)] \cdot d}{2[1 + A_d \cdot \cos(\theta_d)]} & \dfrac{[1 + 2A_d \cdot \cos(\theta_d) + A_d^2] \cdot d}{2[1 + A_d \cdot \cos(\theta_d)]} \le \tau_d \le 1 - \dfrac{(1 - A_d^2) \cdot d}{2[1 + A_d \cdot \cos(\theta_d)]} \\ \dfrac{A_d \cdot [A_d + \cos(\theta_d)] \cdot (d/2 + 1 - \tau_d)}{2 + A_d \cdot \cos(\theta_d) - A_d^2} & 1 - \dfrac{(1 - A_d^2) \cdot d}{2[1 + A_d \cdot \cos(\theta_d)]} \le \tau_d \le 1 + \dfrac{d}{2} \\ 0 & \tau_d > 1 + \dfrac{d}{2} \end{cases} \tag{8}$$

If the multipath signal's amplitude attenuation factor A_d = 0.2, 0.3 or 0.5, the pseudorange measurement error curve is shown as Fig. 2.



(a) Amplitude attenuation factor 0.2

(b) Amplitude attenuation factor 0.3



(c) Amplitude attenuation factor 0.5

Fig. 2 The pseudorange measurement error curve

When the chip rate is 2.046 Mc/s, if A_d = 0.2, there will be 0.05 chip error caused by multipath effect, the pseudorange measurement error will be 24.4 ns. If A_d = 0.3, there will be 0.075 chip error caused by multipath effect, the pseudorange measurement error will be 36.7 ns. If A_d = 0.5, there will be 0.125 chip error caused by multipath effect, the pseudorange measurement error will be 61.1 ns.

GNSS satellite uses an array of some antenna elements, a weaker signal with different phase may appear if the impedance mismatch happened, this may cause channel delay deviation between each antenna element channel.

An analysis model is built refer to GPS Block I which is constituted with 12 helical antenna elements. The array contains an inner ring of four elements (radius 16.24 cm) and an outer ring of eight elements (radius 43.82 cm), fed by an antenna coupler network. The inner ring of elements transmits 90% of the power, while the outer ring transmits 10%. The carrier phase difference between inner ring and outer ring is 180°.

Taking GPS L1(1575.42 MHz) as an example [6] ,The orbit height of the GPS satellite is 20,187 km. Referring to the earth radius 6371 km,the distance from the satellite to the beam marginal point is 25,783 km. The power of the wave beam marginal point relative to the subastral point is 2.1 dB, then the angle of the signal direction deviating from the array surface normal is 13.88°. In order to make the

surface received signal intensity basically consistent, the gain of the antenna array in the direction of 13.88 degrees should be 2.1 dB higher than that of the normal direction. The carrier phase is the same in all directions after beam forming. The simulation result of the antenna pattern is shown in Fig. 3.

Time delay difference between array elements may have an affect on the measurement result of coherent phase discriminator. Selecting GPS L1C/A as the signal source, the pseudorange measurement difference (caused by multipath) between marginal point and the subastral point is shown in Fig. 4. In the figure, the x axis is the delay difference between inner ring and outer ring of the satellite antenna array, the y axis is the pseudorange measurement difference between the beam marginal point and the subastral point.

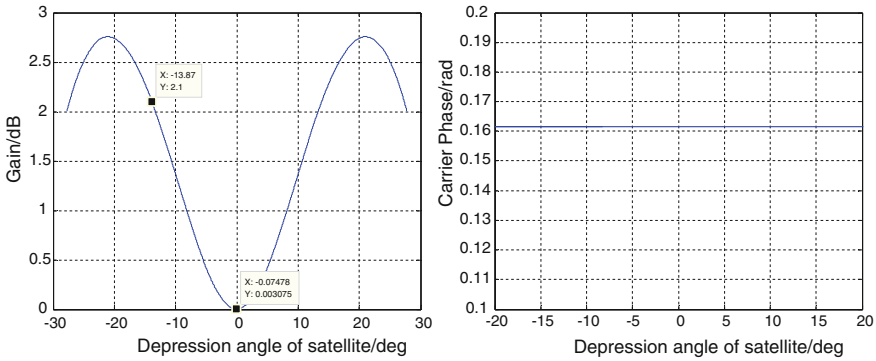


Fig. 3 GPS block I antenna pattern (L1 signal)

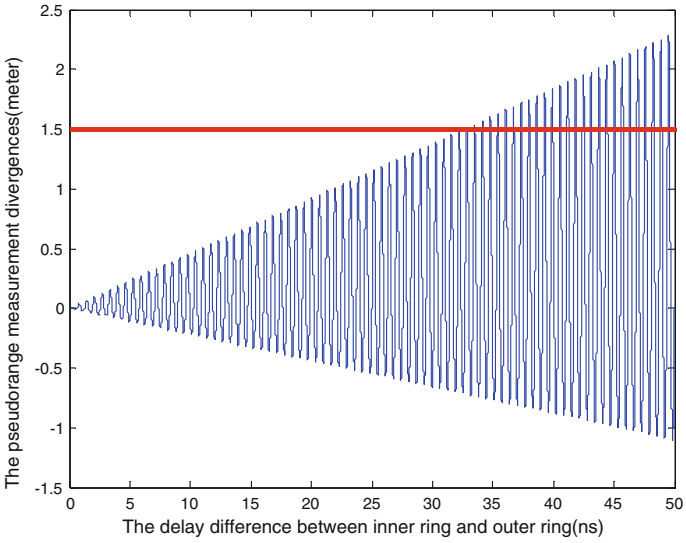


Fig. 4 The relationship between pseudorange measurement divergences and the inner-outer ring delay difference

Based on the simulation results, when the delay deviation between inner ring and outer ring is over 10 ns, there would be an obvious pseudorange measurement difference between the beam marginal point and the subastral point. When the delay deviation between inner ring and outer ring is over 30 ns, the phenomenon described by article [1] would appear.

3 Ground Test System for the Satellite Inner Multipath Effect

According to the analysis of the Chap. 2, the pseudorange measurement difference between the beam marginal point and the subastral point can be estimated by measuring the time delay of each antenna element through coherent code phase discriminator.

The Ground test system diagram for the satellite inner multipath effect is shown as Fig. 5.

In this satellite test system, the test port is the satellite antenna array coupling port, and the BeiDou 1PPS (1 Pulse per Second) signal test port. Connecting the antenna array coupling port with the ground radio frequency switch matrix through the RF coaxial cable, the system uses the microwave switch in the switch matrix to select the RNSS (Radio Navigation Satellite System) downlink signal from each antenna array coupling port. 1PPS rising edge triggered signal acquisition is conducted by the DAE (Data Acquisition Equipment), the acquisition data is transmitted to the computer to analyze, the processing results are obtained.

Satellite 10 MHz frequency standard is used as an external reference for the data acquisition equipment. The time delay result of the signal from coupling port 1 is evaluated, and then subtracted from the time delay results from each coupling port.

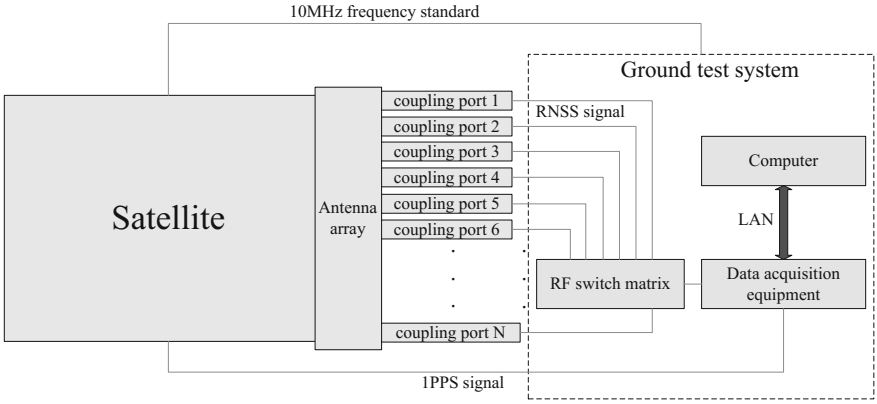


Fig. 5 The ground test system diagram for the satellite inner multipath effect

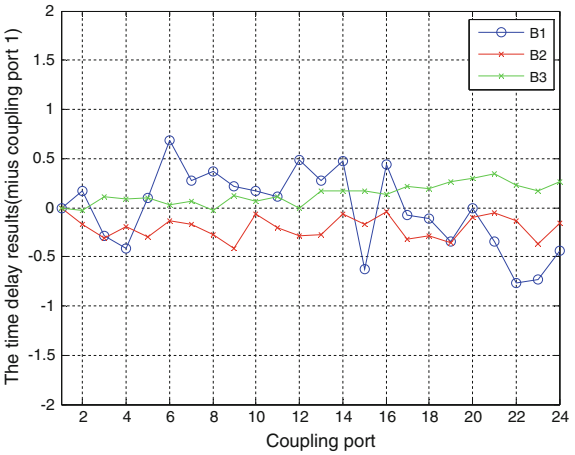


Fig. 6 The time delay of the signal from each coupling port minus coupling port 1

If the maximum value of the results is more than 30 ns, the satellite inner multipath effect may be the main cause of the description of the phenomenon in the article [1]. Otherwise, the fault branch can be excluded.

4 Test Result

According to the BeiDou ICD2.1, selecting the B1, B2 and B3 band signal as the ground test object, the time delay result of the signal from coupling port 1 is evaluated, and then subtracted from the time delay results from each coupling port, the test result of a BeiDou satellite is shown as Fig. 6.

Through the result, the delay deviations of the signal from each coupling port are within 1 ns, far from the 30 ns, therefore, the fault branch can be excluded.

5 Conclusion

Article [1] describes the phenomenon about the code pseudorange variations of the BeiDou satellite system which result in code phase divergences of more than 1 m. The inner multipath effect of the satellite is one of the most important inducements which happened on SVN49. In this article, the mathematical model of the relationship between the satellite antenna array's inner-outer ring delay difference (which caused by the multipath effect) and the beam marginal-subastral point pseudorange measurement divergence is built. This can be used to evaluating the impact of the inner multipath effect on navigation satellite. The Ground test system using the method has been applied to the BeiDou satellite test, and the test result is given. Through the result, the satellite inner multipath effect fault branch can be excluded.

References

1. Wanninger L, Beer S (2015) BeiDou satellite-induced code pseudorange variations: diagnosis and therapy. GPS Solution 19:639–648
2. Langley R (2009) The SVN49 pseudorange error. GPS World 20:8–12
3. Springer T, Dilssner F (2009) SVN49 and other GPS anomalies. Inside GNSS 4:32–36
4. Matilde S, Miguel A, Ana G (2007) Performance analysis and parameter optimization of DLL and MEDLL in fading multipath environments for next generation navigation receivers. IEEE Trans Consum Electron 53(4):1302–1308
5. Zhan Y, Strojwas A, Li X (2005) Correlation-aware statistical timing analysis with non-Gaussian delay distributions. In: Proceeding of ACM/IEEE design automatic conference, 2005, pp 77–82
6. Hegarty C, Chatre E (2008) Evolution of the global navigation satellite system (GNSS). In: Proceedings of the IEEE, vol 96, December 2008

Positioning and Velocity Performance Evaluation of Transponding System

Yuxi Liu, Xiaolin Jia and Rengui Ruan

Abstract This paper introduces the work principle and positioning performance of the transponding navigation system, and gives the navigation and positioning measurement equations. Using two consecutive days of data from the transponding system, the paper analize the performance of User Equivalent Range Error, static positioning accuracy, kinematic positioning and velocity accuracy. The trail show that: the UERE is better than 1.15 m and different satellites shows hierarchical phenomenon. In the constellation state that one IGSO satellite with four GEO satellites, the accuracy of static positioning is better than 4.85 m in horizontal, and 6.3 m in vertical. The kinematic positioning accuracy is better than 12.5 m in horizontal and 13.1 m in vertical, the accuracy of velocity is better than 0.08 m/s in horizontal and 0.12 m/s in vertical, and the repeatability between days is better than 0.01 m/s.

Keywords Transponding system · Positioning · User equivalent range error Velocity

1 Introduction

In 2002, Academician Ai Guoxiang, National Astronomical Observatory of Chinese Academy of Sciences, put forward the concept of "transposing satellite navigation and positioning" for the first time, Proposed the use of commercial communications satellite transponding forwarding navigation messages and other information to reduce the number of launch dedicated navigation satellite to achieve navigation and

Y. Liu (✉) · X. Jia · R. Ruan
State Key Laboratory of Geo—Information Engineering, Xi'an 710054, China
e-mail: eryunudt@163.com

Y. Liu · X. Jia · R. Ruan
Xi'an Research Institute of Surveying and Mapping, Xi'an 710054, China

R. Ruan
Information Engineering University, Zhengzhou 450001, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_8

positioning, speed and timing [1]. Because the time signal required for navigation and positioning is sent by the ground control station's atomic clock, it is not necessary to create a space-borne atomic clock which is more difficult to maintain. The atomic clock in the ground-based station superior to maintain high accuracy and stability, which reduce the technical difficulty and improve navigation and positioning accuracy. June 2005, Chinese Area Positioning System (CAPS) was successfully developed,when there is no IGSO communications satellites at the time, the CAPS validation system rented transponders on four Geostationary Earth Orbit (GEO) communications satellites, which formed the constellation of verification systems. At the same time, using the altimeter to determine the user's height as an innovative method of virtual satellites, it solves the key problem that the signals can not be three-dimensionally located using only the GEO satellite signals [2].

On March 30, 2015, China's first Beidou global pilot satellite was successfully launched. PRN No. C31 belongs to the Inclined Geosynchronous Satellite Orbit (IGSO), It is also the only non-GEO satellite in orbit currently equipped with repeaters in Beidou satellite navigation system. This experimental satellite has become an important part of the development and experimental verification of the transponding system, enabling the transponding system to realize the constellation configuration of the IGSO + GEO satellite. In 2016, the National Time Service Center of the Chinese Academy of Sciences developed transponding orbit determination system consisting of a main control station (Xi'an) and six railway stations (located in Changchun, Xi'an and Kunming, etc.). After verification, the three-dimensional RMS values of the orbital overlapping arcs of the GEO and IGSO satellites reached 2 and 0.9 m [3–5].

This article is based on the construction status of transponding system and the PVT test conducted by the National Time Center of Chinese Academy of Sciences in Beijing and Sanya from September 4 to September 5, 2017, using the measured data to analyze the current performance of the transponding system from the aspects of space signal accuracy, static/dynamic positioning accuracy and speed accuracy.

2 Positioning Principle of Transponding System

2.1 Theory

C-band is used in transponding system, advantage in ranging accuracy, small weather conditions and more GEO satellite [6]. The system transmits the navigation message through the transponding on the satellite and completes the ranging, satellite transponding as a spatial reference, but it does not carry a high-precision atomic clock, so the necessary time and frequency reference for the repeater system navigation are generated by the ground master station, broadcast by the ground station and transmitted by the satellite transponding broadcast to the user. Figure 1 is the composition of the principle of the transponding system, dotted line stands for

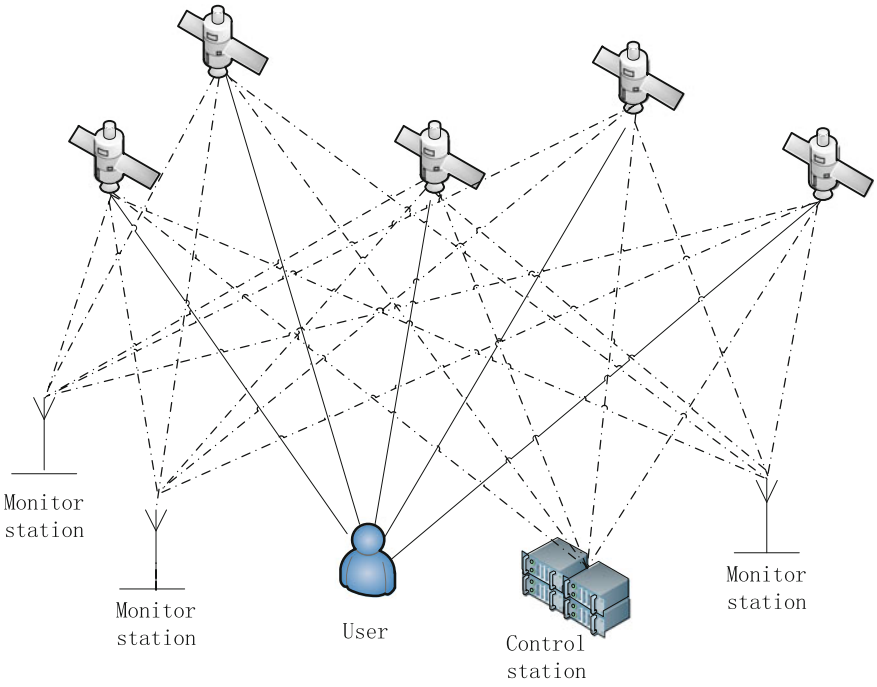


Fig. 1 Consitute of Transfer System, with one control station located in Xi'an and six monitor station diffusely located around inland

two-way communication, solid line stands for one-way communication. Both ground control station and monitor station broadcast and receive signals in both directions at the same time. The ground station serves as a known point and meets the measured satellite's space position, and then transmits the signal to the user through the master control station to perform navigation, positioning and timing services. Therefore, different from GPS, BDS and other direct-turn navigation systems, the starting point of the observation of the transponding system is transferred from the satellite to the ground. The starting site of the time reference and the spatial reference are all at the ground control station. In order to make the work of the communication satellite have Similar to a dedicated navigation satellite-like features, In the transposing navigation message, it is necessary to give the navigation signal at the time when the phase of the satellite antenna phase center is consistent with the time precision of the system. For this purpose, CPAS develops the virtual atomic clock system, and the satellite terminal uses the "virtual atomic clock" method to accurately predict the satellite signal Spread the time, The concrete realization principle of virtual clock is detailed in [7, 8].

The ephemeris parameters broadcast by the transponding system consist of a set of satellite coordinates, the ephemeris reference time, and the ephemeris data age parameters, which are represented by the Chebyshev polynomial coefficients in six

order. Ephemeris update in 15 min, with the effective duration of 30 min. Using the Chebyshev polynomial fitting to calculate the satellite position at any time.

2.2 Navigation Equation

As the ranging and time base are transferred to the ground, the measurement equation of C-band transponding system is different from the straight-line system. Similar to the SLR data, it adopts the bidirectional ranging mode that includes the master station to satellite and satellite to the user two sections ranging [9]. The measurement model as shown below:

$$\rho = \rho_{up}(t_m, t_s) + \rho_{down}(t_s, t_u) \tag{1}$$

$$\begin{aligned}\rho_{up}(t_m, t_s) = {} & cdt_m + c(t_s - t_m) + ion_{up} + \\ & tro_{up} + rel_{up} + ant_{up} + d_s + \varepsilon\end{aligned} \tag{2}$$

$$\begin{aligned}\rho_{down}(t_m, t_s) = {} & c(t_u - t_s) + ion_{down} + tro_{down} + \\ & rel_{down} + ant_{down} + \varepsilon\end{aligned} \tag{3}$$

Therefore, $\rho_{up}(t_m, t_s)$ and $\rho_{down}(t_s, t_u)$ indicates the uplink and downlink C-band ranging values, C is the speed of light, t_m, t_s, t_u respectively stands for ground control station, satellite downlink signal and the user receives the signal at the moment of the clock time. dt_m is the ground station atomic clock time deviation, the ground station usually with high stability atomic clock, so dt_m can be ignored for small, ion_{up}, tro_{up}, rel_{up}, ant_{up}, d_s respectively represent the delay of the ionosphere, the tropospheric delay, the relativistic, the antenna correction, and the time delay of the transponding between the satellite receiving signal and the transponding in the uplink signal. ion_{down}, tro_{down}, rel_{down}, ant_{down} respectively denote the ionospheric delay, tropospheric delay, relativity, antenna correction, measurement noise and multipath in the downlink signal.

3 Experimental Analysis

The PVT test is based on the constellation of one IGSO satellite (C2-Compass I1-S) and four GEO satellites (C6, C7, C8, C5), the transponding satellite distribution shown in Fig. 2, among them, C5 belongs to a small incline GEO satellite, so there is a north-south movement during the observation period. Test scenarios are divided into static experiments and dynamic sports car experiments, including static tests in Beijing and Sanya simultaneously, dynamic sports car test in Beijing. The test

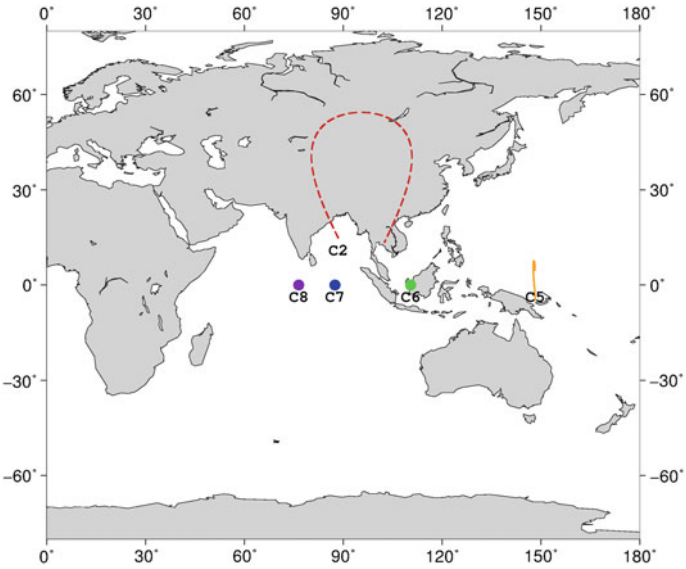


Fig. 2 Satellites distribution, line and point in different color stands for different satellites, note that C2 is IGSO satellite and C5 is GEO satellite with small incline, others are GEO satellites

period is from September 4, 2017 to September 5, 2017. To ensure that five satellites are observed. The test was conducted after the IGSO satellites entered in and covered most of the transit arc of the IGSO satellite.

3.1 Data Processing Strategy

(1) The position solution uses a weighted least-squares estimate, weighting each satellite based on the elevation angle [weight = sin(ele)].
(2) UERE calculation using single epoch all satellite to eliminate the receiver clock.
(3) Static test use the known coordinates, dynamic test using synchronization output positioning and speed measurement results by the vehicle inertial navigation device.
(4) Statistics of positioning and speed error only when five satellites all visible, UERE statistical all satellites arc, the data processing did not set the mask angle, based on the observation data whether the judge visual or not. (Analysis of measured data shows that all the satellite within the observation arc show an angle of 25° or more).
(5) The results of statistics using the median to the gross error method removing gross error [10].

3.2 Static Test

(1) UERE evaluation

The User Equivalent Range Error (UERE) refers to the residual error of the user ranging signal, that is, the difference between the observed value of the pseudorange and the geometric distance of the satellite station. Among them, the pseudo-range observation needs satellite antenna phase center, antenna phase center station, broadcast satellite ephemeris, satellite receiver clock error, ionospheric delay, tropospheric delay, relativity and earth rotation error correction. In satellite navigation, orbit and clock accuracy will be directly reflected in the UERE, UERE and DOP values can also be used to roughly calculate the accuracy of navigation and positioning. For transponding systems, the UERE calculation formula can be calculated by Eq. (4).

$$UERE = \mathrm{P} - c|t_u - t_m| - clk_{rec} - ion_{up/down} - tro_{up/down} - rel_{up/down} - ant_{up/down} - \delta_{etc} \quad (4)$$

Table 1 and Fig. 3 give the UERE statistical values and error sequences of the satellites in the static test arc respectively. In the figure, different colors represent different satellites. Mean results of UERE multi-satellites statistical Beijing and Sanya observation are less than 1.15 m, the difference between days and days within 0.15 m, single-satellite UERE better than 1.7 m. The error sequence shows that the UERE stratification of Beijing observation site is more obvious than that of Sanya observation site. The UERE calculation results of different satellites are different due to different observation directions. The UERE of Beijing observation site is larger than that of Sanya observation site, while the observation of Sanya observation site C7 satellite UERE large, September 5, the latter part of the observation arc C2 satellite UERE also become larger trend.

(2) positioning accuracy

Positioning accuracy is determined by the spatial signal accuracy and constellation geometry configuration. Due to the limited conditions of the current transponding system, only five satellites are used and four satellites belong to the approximately collinear GEO satellites. Therefore, the constellation geometry is

Table 1 Statistics of UERE for all satellites in the test arc (95%, m)

Site	Date	C2	C5	C6	C7	C8	Average
Beijing	9.4	0.90	1.44	1.21	0.74	0.90	1.04
	9.5	1.09	1.51	1.27	0.79	1.03	1.14
Sanya	9.4	0.98	0.90	0.58	1.50	1.03	0.99
	9.5	1.66	0.83	0.77	1.42	0.94	1.12

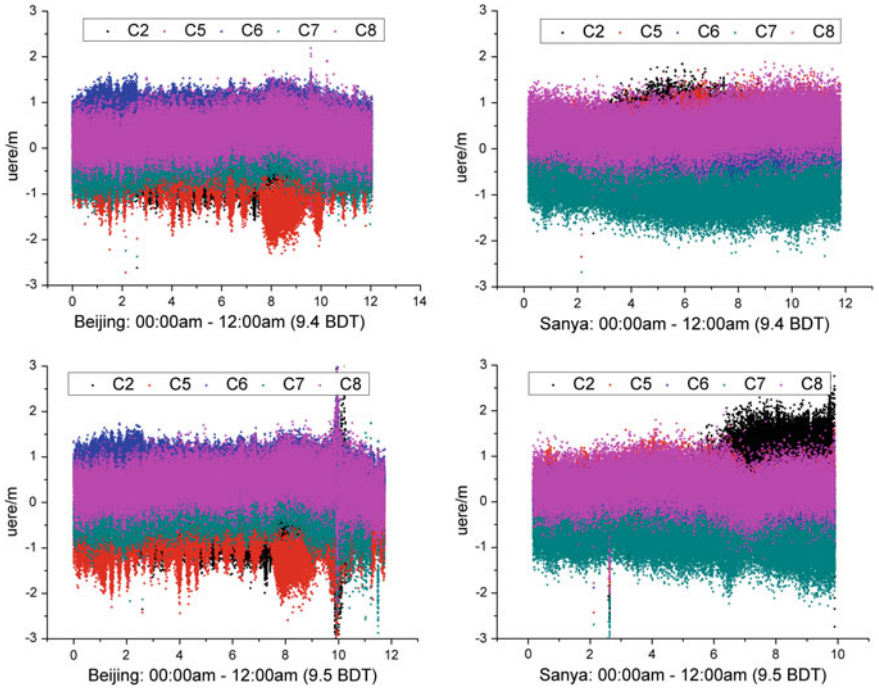


Fig. 3 Time sequence of UERE, different satellite in different color and the figures upload gives the results in September 4 and figures download give the results in September 5 (left for Beijing and right for Sanya)

still far from the ideal state. Based on the measured data, the Position Dilution of Precision (PDOP) value of the test arc of the test was calculated and the variation is shown in Fig. 4. During the BDT period of September 4, the PDOP values of

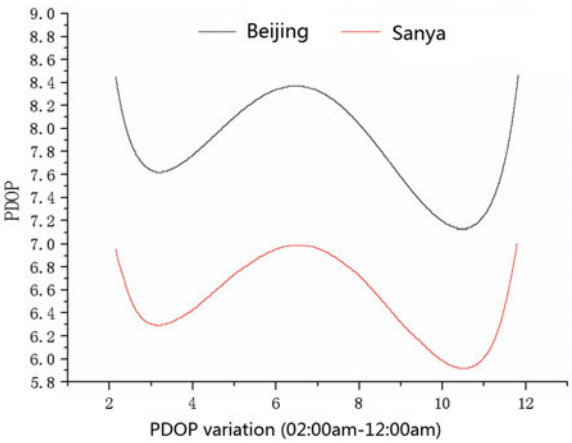


Fig. 4 Variation of PDOP, black line stands for the pdop variation of Beijing and red line stands for the pdop variation of Sanya

Beijing observation site (black curve) ranged from 7 to 8.6, and the Sanya observation site (red curve) slightly outperformed Beijing observation conditions. The PDOP values ranged between 6 and 7.

Figures 5, 6 and Table 2 show the static positioning error sequences of Beijing observation site and Sanya observation site under the condition that all five forward-type satellites are visible. To ensure coverage of the visible arc of the IGSO satellite, the daily testing time is from 2 a.m. to 12 noon. Under 95% quantile statistics, the horizontal positioning accuracy of Beijing observation site is better than 4.85 m, the vertical accuracy is better than 6.3 m, the accuracy of three-dimensional is better than 7.4 m, and the observation geometry of Sanya observation site is better than that of Beijing. At 2.95 m, vertical accuracy is better than 5.95 m, and 3D positioning accuracy is better than 6.1 m.

In addition, it can be found from the error sequence map of Beijing observation point that there is obviously a high-frequency fluctuation period with the most significant N direction. The analysis may be due to the periodic change caused by

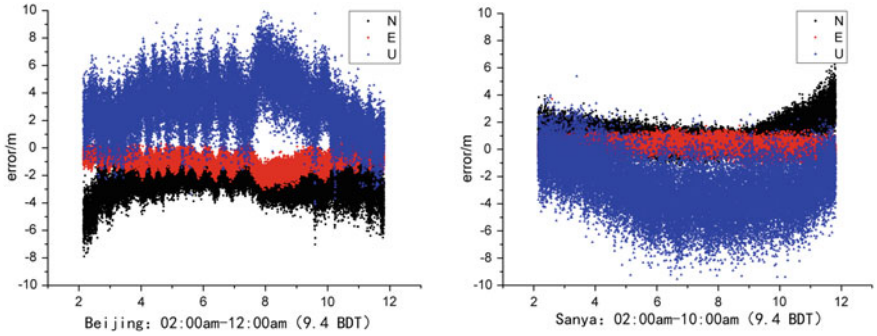


Fig. 5 Positioning error for static, black, red, blue points stand for the north, east, vertical positioning error respectively in September 4, left for Beijing site and right for Sanya site

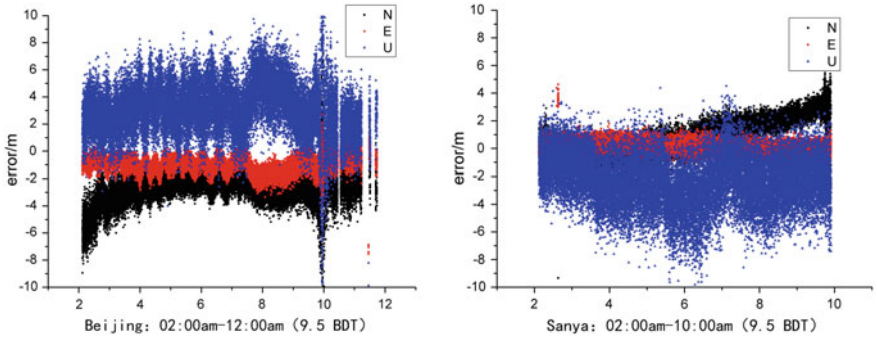


Fig. 6 Positioning error for static, black, red, blue points stand for the north, east, vertical positioning error respectively in September 5, left for Beijing site and right for Sanya site

Table 2 Statistics of positoning error in test arc (95%, m)

Site	Date	N	E	U	3D
Beijing	9.4	4.30	1.91	6.28	7.32
	9.5	4.66	1.94	6.02	7.24
Sanya	9.4	2.73	0.87	5.91	6.06
	9.5	2.89	0.85	5.31	5.50

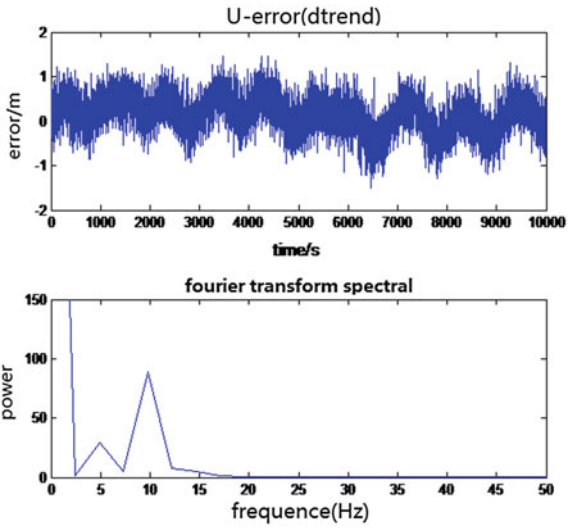


Fig. 7 U-error and spectrum, figure upload is the positioning error of vertical orientation after detrended, figure download is the Fourier transform spectral of U-error

the retransmission period of the message system in the forwarding system. In order to analyze its cycle size, spectrum analysis of the localization error N in Beijing measuring point is carried out. Taking 10,000 epoch data, polynomial fitting is firstly used to remove the trend term in the error sequence, followed by fast Fourier transform to remove the trend. The error sequence and spectrum of the term are shown in Fig. 7. It can be seen that after the spectral change, there is a clear peak around the spectrum 10 Hz. Corresponding to the data sampling frequency, there is a fluctuation period close to 1000 s in the N-directional error sequence, which is close to the 15-min navigation message refresh rate in the transponding system.

3.3 Kinematic Testing

Kinematic test using the inertial output as reference, the receiver and inertial navigation system homology locate on the car, the test route selection Beijing

Fig. 8 Sports route, the figure shows the full route of the car in the test

Institute of Optoelectronics—G7 Jingxin Expressway—North 6 Central—East 6 Central, shown in Fig. 8.

Figures 9, 10 and Table 3 show the dynamic positioning and speed evaluation results under the condition that all five satellites are visible. As a result, a relatively stable piece of data during the test is selected. Under the condition of 95% quantile statistics, the 3D positioning accuracy is better than 16.6 m, the 3D velocity accuracy is better than 0.13 m/s, and the difference between days is better than 0.01 m/s. It can see from the figure that the positioning error sequence consistent with the speed error sequence.

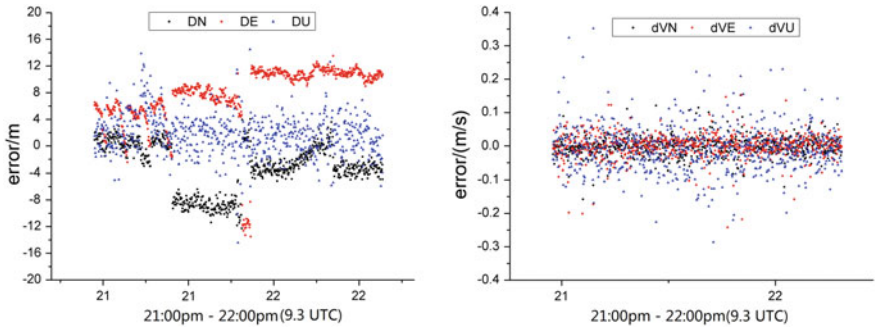


Fig. 9 Positioning and velocity error in kinematic, black, red, blue points stand for the north, east, vertical positioning error respectively in September 5, left for positioning error and right for velocity error

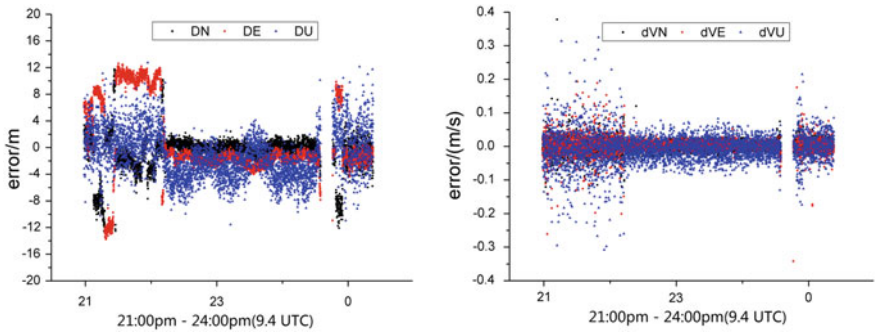


Fig. 10 Positioning and velocity error in kinematic, black, red, blue points stand for the north, east, vertical positioning error respectively in September 5, left for positioning error and right for velocity error

Table 3 Statistics of positioning and velocity error (95%, m)

Type	Date	N	E	U	3D
Positioning	9.4	7.83	9.47	13.10	16.55
	9.5	8.55	12.10	12.94	16.36
Velocity	9.4	0.06	0.07	0.12	0.13
	9.5	0.04	0.04	0.11	0.12

4 Conclusion

This paper introduces the composition of the transponding system and the principle of navigation and positioning, and then uses the measured data of the transponding test to evaluate the SIS accuracy, positioning accuracy and speed accuracy. The following conclusions are obtained:

(1) UERE accuracy of transponding system is better than 1.15 m, with straight-line satellite navigation system, UERE there is a certain degree of stratification, but also from the side reflects the UERE layered phenomenon more by the receiver segment design.
(2) Under one IGSO and four GEO condition, the transponding system can achieve the static positioning accuracy of 4.85 m in horizon and 6.3 m in vertical, dynamic speed accuracy 0.08 m/s in horizon and 0.12 m/s in vertical, verifying that the transponding system can be used for high-precision navigation and positioning.

As can be seen from the results, as a pilot system, the transposing navigation system already possesses the basic conditions of high-precision navigation and positioning. Under the ideal condition of the constellation geometry, its navigation positioning could be better than the straight-haul satellite navigation system in the area navigation. At the same time, it can provide reference for the construction and

development of Beidou satellite navigation system in the future in terms of high-precision determination of satellite orbit and time service.

Acknowledgements This work was supported National key Research Program of China "Collaborative Precision Positioning Project" (No. 2016YFB0501900).

References

1. Editor's note (2008) Innovation and exploration of tansponding satellite navigation system. Sci China, Ser G 38(12):1613–1614
2. Ai GX, Shi HL, Wu HT et al (2008) Positioning system based satellite communication and Chinese area positioning system (CAPS). Chin J Astron Astrophys 8(6):611–635
3. Yang X, Ding S, Lei H et al (2016) Research progress of the technology of orbit observation and determination via transfer(OODT). J Time Frequence 39(3):216–224
4. Lei H, Li Z, Yang X et al (2015) Precise orbit determination experiment of compass-GEO based on transponder ranging. Acta Geodaetical et Cartographica Sinica 40(5):31–33
5. Wang J, Huang G, Zang N et al (2015) Simulation experiment and precision analysis of CAPS. J Navig Positioning 3(3):110–116
6. Lu X, Wu H, Bian J et al (2008) Signal of Chinese area positioning system editor's note, innovation and exploration of tansponding satellite navigation system. Sci China G 38 (12):1634–1647
7. Li X, Wu H, Bian Y et al (2008) Virtual satellite atomic clock with pseudo-difference function. Sci China G 38(12):1723–1730
8. Wu H, Bian Y, Lu X et al (2008) CAPS navigation signal ground launch time synchronization and carrier frequency control. Sci China G 38(12):1660–1670
9. Shi H, Pei J (2008) Solution of measurement system of transponding satellite navigation and positioning system. Sci China G 38(12):1687–1701
10. Huang Y (1990) Data detection and robust estimation. Survey and Mapping Press, Beijing

Design of Distributed Test System for Satellites

Xuejia Li, Jianwei Yang and Jiahao Li

Abstract The satellite test mode is experiencing the change from local test for single target to distributed test for multiple satellites. Distributed Test System (DTS) can coordinate geographically dispersal test nodes to achieve integrated test, improve the utilization rate of test resources and the efficiency of satellite test. In this paper, the system framework of the satellites DTS is discussed. The requirement of the system hardware is analyzed and module design of software is presented. The key technologies of DTS are also elaborated in detail. All the works have set a basis for the further implementation of satellites distributed test system and pushed the construction of Beidou-3 a step forward.

Keywords Distributed test system · Satellite test · Auto test system

1 Introduction

With the on-going construction of Beidou-3 navigation satellite system, China's space industry has further developed, thus raising higher requirements for satellites test. At present in China, single star testing and centralized testing domain the auto test field, whose resource utilization efficiency and openness are relatively low that hardly meet the needs of constellation test [1]. While China's satellite ground stations are widely distributed, it is necessary to build a general distributed auto test system that makes full use of existing resources and testing personnel, shortens the

X. Li (✉) · J. Yang · J. Li
College of Intelligent System, National University of Defense Technology, Changsha, China
e-mail: lixj12@outlook.com

J. Yang
e-mail: yjw94@sina.com

J. Li
e-mail: lijiahao554579@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_9

satellite test cycle and reduces the test cost to guarantee the normal operation of the satellite in orbit.

DTS connects measurement instrumentations and computers equipped in dispersal test nodes through specific communication links, exchanging data collected and processed. Under the management of the system control center, all the nodes cooperate to complete the test tasks, which can achieve resource sharing, decentralized operation, centralized management, measurement process monitoring and fault diagnosis [2]. Research on DTS abroad has developed more mature. The Test and Training Enabled Architecture (TENA) adopted by the joint military mission and joint mission environment test capacity (JMETC), enabling range interoperability and range resource reuse across the DoD range community, has played an important role in the joint military missions [3]. The US company SpaceX uses plenty automatic test and remote test technology to cut down the test time in the launch site so that reduces launch time significantly compared with other satellites. The distributed test technology in China is still in the initial stage. A scheme of spacecraft integration test based on remote test proposed by China Academy of Space Technology realized the coordination between front launch and rear analysis for the first time, optimizing test personnel distribution [4]. Spacecraft Integrated Test Information Management Platform proposed by Beijing Institute of Spacecraft System Engineering secured parallel test for multi-spacecraft [5].

Distributed test has become a trend in the development of automatic test technology. Based on the requirement of satellite test, this paper proposed the structure of distributed satellite automatic test system and designed the software system of test nodes. Integrated with other modules such as test scheduling, equipment management, fault diagnosis and other subsystems, the system could realize the overall automation in remote satellite test.

2 System Design

2.1 System Framework

The satellite DTS is composed of the control center, the remote ground test station and the front test equipment set, as shown in Fig. 1. As the core of the system, the control center receives test requests initiated by remote clients equipped with client software, provides corresponding services to them, controls and manages all test nodes. The test node should respond to the instructions from the control center in time. Following the specific instruction information, the nodes measure satellites parameters based on a set of test process and configuration information, then analyze the test results and send back to the control center. In the non-cooperative mode, the nodes can also perform the test mission independently.

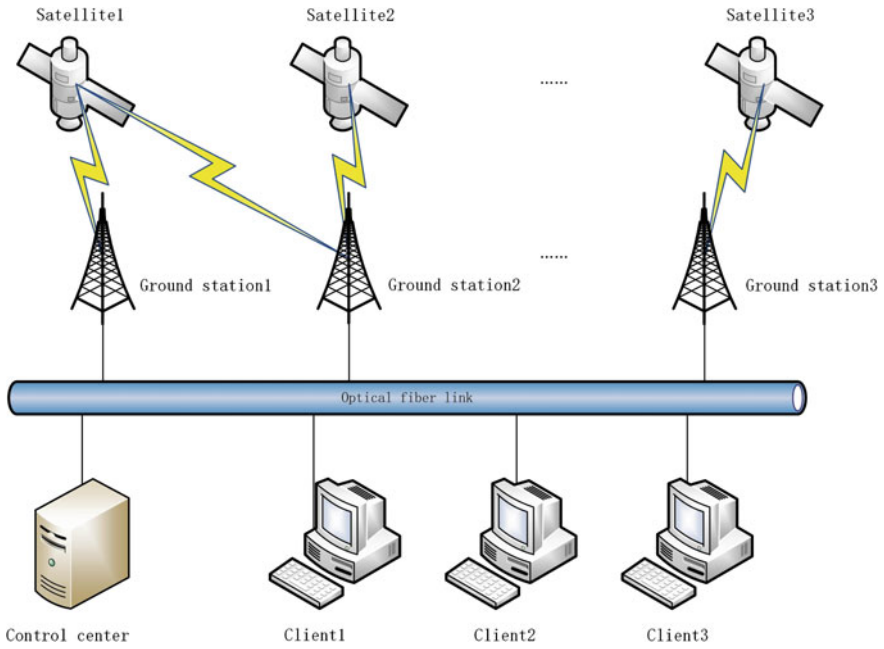


Fig. 1 Framework of distributed satellites test system

2.2 *Hardware Requirements*

To enhance the flexibility of the system and the exchangeability of the instruments, the distributed system hardware should adopt an open architecture and regulate general key interfaces. The system hardware interface includes the internal interface and the external interface. The internal interface is to define the communication path between the computer of the control center and the test instrumentations, as well as the communication connection between the instrument controller and the test module. The external interface defines the host to the external environment interface, which is an effective communication with the remote system [6]. The regulation of the external interface contributes to the convenient interaction between the test system and the diagnostic framework, realizing the sharing and reusing of the diagnose information. This kind of interoperability makes integrated test diagnosis implementable.

The standardized basic module combined with LXI bus technology are utilized to realize the "plug and play" of the instrument, which prolongs the system life and reduces upgrading costs.

2.3 Software Design

The upper computer software adopted Microsoft Visual Studio 2010 as the development platform. Net Framework 4 as the development environment, and C# as program language.

According to the distribution characteristics of the nodes, the client/server (C/S) architecture is adopted in the software system, which has the advantages of fast response, high security performance and full isolation of the functional components of the system. Considering the application scenario that each node is oriented to multiple test objects at the same time, the system should support multi-channel parallel test, which can be achieved by multithread technology in the software perspective to realize dynamic allocation and optimal scheduling of test resources [7]. Besides, test and data processing are time-consuming. To keep friendly interaction of software user interface, the UI process should be separated from test operations through multithreading while delegate and event in C# is applied to achieve inter-thread data exchange. Then the test results could be displayed timely.

The test client is divided into configuration management module, data processing module and display module according to the function, and the three modules can be further subdivided into several sub-parts. The modules remain relatively independent and have a clear flow into the outflow interface for invoking and reconstitution. The software components of the client are shown in Fig. 2.

The test process management module can customize the test project and organize the test process using standardized test language.

The state control module generates the control instructions for the test instrument and the test channel to manipulate the switch of the test equipment remotely.

The parameter configuration module mainly includes the configuration of remote control orders, telemetry parameters and measurement parameters. Remote control

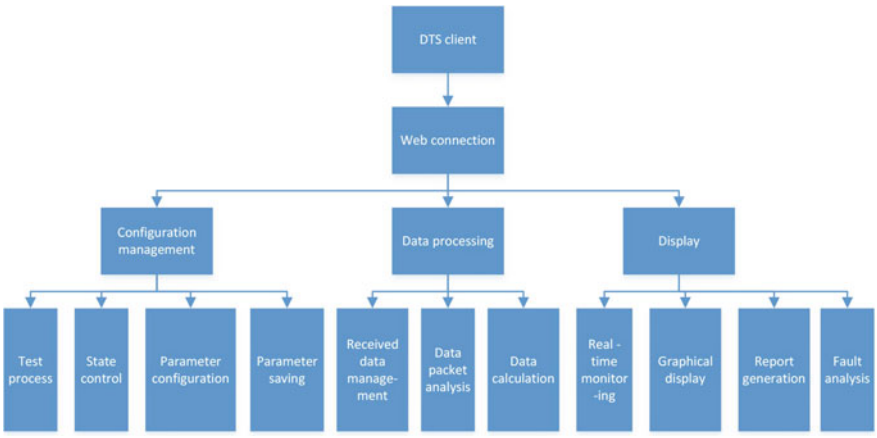


Fig. 2 Software components of remote client

module configures the orders injected into the satellites, which can not only load the pre-stored instruction document, but also input the injection instruction manually. The telemetry parameter configuration is used to analyze the parameters of the telemetry data. The remote control and telemetry data reflect the working state and the environmental parameters of systems in the satellite, so that it can be used as the basis for evaluating the performance of the satellite and analyzing the fault. In order to complete distance and speed measurement, the parameters of carrier ring bandwidth, pseudo-code ring bandwidth, information rate and code rate are needed through the measurement parameter configuration module. The configuration parameters can be stored in a local or shared database in text form through the parameter saving module. Figure 3 is the client parameter configuration interface.

The received data management module realizes the continuous reception of the data flow. To ensure the integrity of the received data, a separate receiving thread is

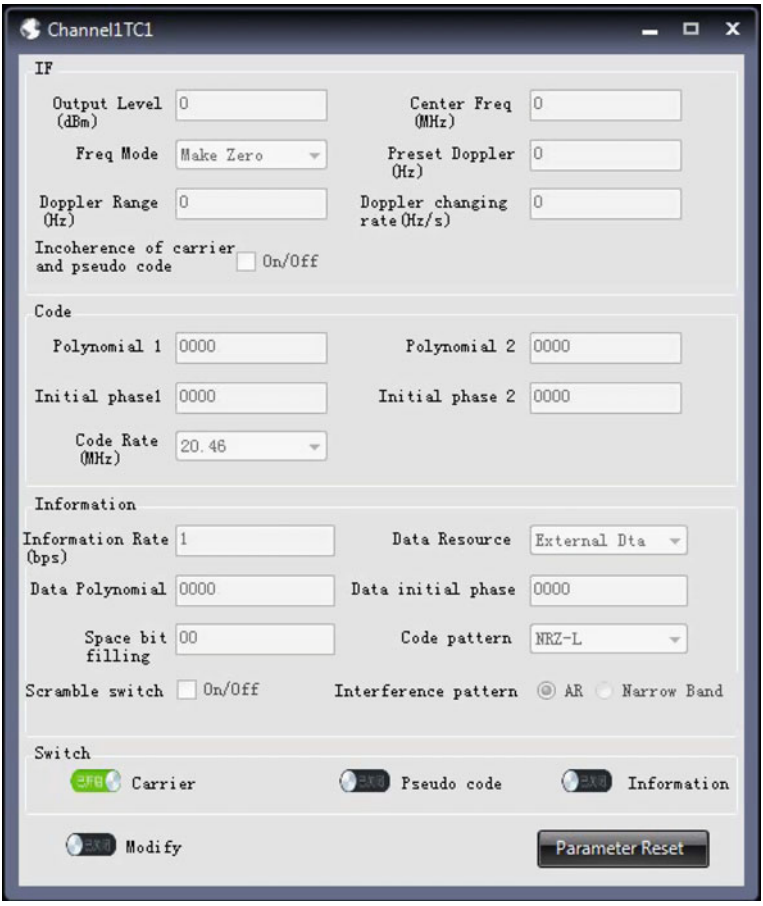


Fig. 3 Interface of parameter configuration

created and kept circulating. Meanwhile, a FIFO type receive buffer is set to store the received data, and then the packet is extracted from the buffer according to the pre-set frame information for subsequent analysis.

The data analysis module obtains the corresponding type of each packet by parsing the data at specific location, and then extracts the information in the packet in accordance with the data format of this type. When data packet format designed, the information requires characteristic clearness and integrity. It should carry more effective parameters with the least data quantity so as to improve the test efficiency.

The data calculation module completes the simple calculation of the parameters, including the statistical error rate, the distance, the speed calculation and so on.

The real-time monitoring module displays data received in time line. With warning line set, data surmounting over the line would be marked for latter fault analysis.

The graphical display module can dynamically draw the relevant parameter curve and reflect its changing trend. Since most satellite tests last for a long time and collect large amounts of data, if new point accumulated on curves over time, memory overflow could occur or even the whole program could die. Therefore, it is necessary to limit points number on the curve by interpolation method.

The report generation module presents the analysis results in the form of spreadsheets, which saves manual filling, improves testing efficiency and avoids errors caused by human transcribing.

The fault analysis module calls corresponding rules to complete the data failure analysis and state judging. Then the fault information is added into the center knowledge base, gradually enhance the level of fault diagnosis.

3 Research on Key Technologies

3.1 Time Synchronization

To ensure the effectiveness of distributed test, the satellite test data acquired by different nodes require to be rearranged on the same time axis. However, due to the independent clock in each system node, complexity of test environment and different time delay in the transmission process, there will be time derivation between nodes inevitably in practical occasions. Time synchronization is the key technology and research hotspot of distributed test system. Currently, the main way to secure time synchronization is satellite time transfer, network time synchronization protocol (NTP) and precise time protocol (PTP).

3.1.1 Satellites Time Transfer

Satellites time transfer is based on the hardware timestamp to achieve high precision time synchronization. Using the high precision time transfer module of the global navigation satellite system, the time signal broadcast by satellites transmit accurately. This signal can synchronize the clock of each node. Satellites time transfer technology shows high precision and can unify the time standard among widely distributed nodes. However, the high application cost such as the requirement for expensive chip assembly and the sensitivity to complex environment have limited its application scenarios [8].

3.1.2 NTP Synchronization

NTP time synchronization was originally proposed in 1991 to synchronize the clock of the computer network [9]. It is based on the client/server mode. The synchronization packet transmission time delay between the local clock and clock server is calculated firstly, then the deviation between the local clock and the reference clock time server is deducted, and finally the local clock synchronization with the reference time server can be realized. The protocol has low cost and achieves millisecond synchronization. It is suitable for the scene with relative low precision.

3.1.3 PTP Synchronization

PTP time synchronization is defined by IEEE 1588 standard, which is suitable for precise time synchronization in measurement and automation systems [10]. The PTP protocol adopts the Master/Slave mode, and the nodes are synchronized through the exchange of time packets. In a distributed test system, the central server, as the main node of the system, provides a global clock. Synchronization is divided into time offset synchronization and information transmission delay calibration between master and slave nodes. The master node sends the synchronous message packet (Sync) and the following message packet (Follow_Up), where the following message packet carries the exact time of the Sync packet leaving the main node, thus the clock bias of the master slave node can be calibrated. Then, a delay request packet (Delay_Req) is sent from the slave node to the master node. Upon the master node receiving Delay_Req packet, the time is recorded and then sent back to the slave node through the delayed response packet (Delay_Resp) to complete the delay measurement [11]. The synchronization model is as shown in Fig. 4.

From the synchronization process above and model diagram, the deviation between the slave node and the master node can be expressed as:

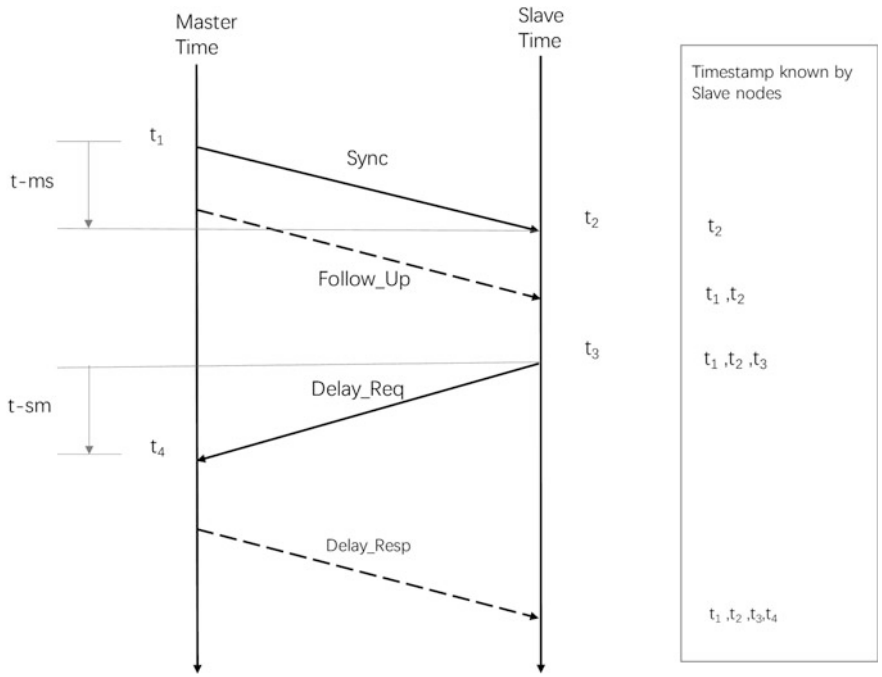


Fig. 4 Model of PTP time synchronization

$$\Delta\tau = \frac{(t_2 - t_1) + (t_4 - t_3)}{2} \tag{1}$$

Compared with the NTP protocol, the mechanism of the two message packets of the PTP protocol improves the accuracy of time measurement, and the synchronization accuracy reaches sub-microsecond level.

3.2 Interoperability of the System

Realizing interoperability of test system can make full use of the existing ATS, integrate the geographically distributed and functionally separated test resources and form a comprehensive environment, thus reducing the construction cost of future integrated test system. Middleware technology is designed to solve the distribution isomerism and realize real-time data exchange between systems.

Middleware is a general service between the hardware platform and the application, and its structure is shown in Fig. 5. It can shield the difference between application software and operation system (OS). The interface provided by the program defines a relatively stable high-level application environment. When

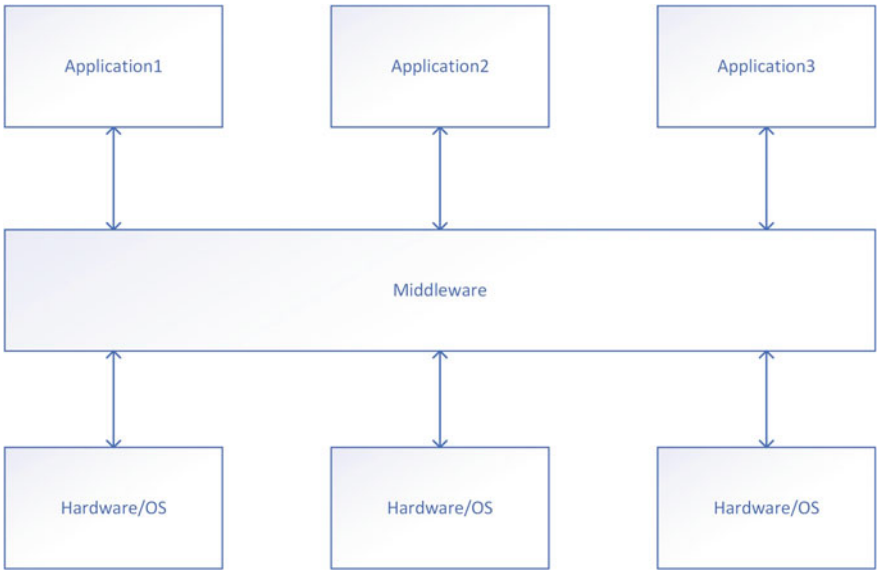


Fig. 5 Middleware construction

upgrading the underlying hardware, it only needs upgrading middleware, which reduces the cost of upgrading the application software [12].

TENA is exactly an example of how middleware provides real-time software system interoperability and interfaces with joint environments. The TENA Middleware combines distributed shared memory, anonymous publish-subscribe, and model-driven distributed object-oriented programming paradigms into a single distributed middleware system. The unique combination of highly abstract programming yields a powerful middleware system that enables the middleware users to rapidly develop complex yet reliable distributed applications [13].

4 Conclusion

Based on the growing demand of satellite test, a distributed satellite test system is proposed in this paper. The system architecture is described, the test module division of client software is introduced in detail and key technologies of the distributed test system is analyzed. It has set the basis for the further system development and implementation. The purpose of this system is to complete the collaborative test of multiple test nodes on the same task. In this mode, the equipment and personnel of each node can be adequately used and the satellite test efficiency can be improved, thus further promoting the construction of Beidou-3 system.

References

1. Fu X, Wang H, Yan J, Lu C, Wang Q (2017) Current situation overview of spacecraft system level test. Space Eng 26(1):120–126
2. Shen Y (2004) Research on the cooperation theory of distributed test system and the real-time performance of communication platform. Doctoral dissertation, University of Electronic Science and Technology of China, Chengdu
3. Home—TENA (2008) Introduction—TRMC Website [Internet]. Tena-sda.org. 2018 [cited 1 November 2017]. Available from: https://www.tena-sda.org/display/TENAintro/Home
4. Pan S, Zhang M, Li H, He Y, Xu Z (2015) Design and application of spacecraft remote test system. Spacecraft Eng 24(5):113–118
5. Chu H, He X, Song H, Bai S (2015) Design and application of spacecraft integrated test information management platform. Spacecraft Eng 24(6):123–128
6. Liu F, Cai D, Meng C, Yin X, Tang G, Liang Q (2015) Research progress of next generation automatic test system architecture. Comput Meas Control 23(2):339–341
7. Chen P (2011) The design of GPS clock synchronization system in distributed text system. Doctoral dissertation, North University of China, Taiyuan
8. Mou Z, Song P (2017) Research on time svnchronizatjon technology in distributed test system. Comput Meas Control 25(5):21–25
9. Mills D (1991) Internet time synchronization: the network time protocol. IEEE Trans Commun 39(10):1482–1493
10. Wen Y, Wen B, Qi Y (2014) Study of time synchronization of distributed data acquisition system. Comput Meas Control 22(4):14
11. IEEE (2008) IEEE standard for a precision clock synchronization protocol for networked measurement and control systems. IEEE 11(2):1–300
12. Ma S, Yu D (2011) Language and system of spacecraft automatic test. National Defence Industry Press, Beijing
13. Hudgins G, Poch K, Secondine J (2012) TENA, JMETC, enabling integrated testing in distributed LVC environments. In: MILCOM 2011 military communications conference, 2012, pp 2182–2187

The Analysis of GPS/BDS Ionosphere-Weak Combination SPP Result

Yulong Kong, Hongzhou Chai, Zongpeng Pan, Rui Wang and Chunhe Liu

Abstract Pseudo-code noise amplification theory of ionosphere-free combination is studied in this article, In allusion to the disadvantage of iono-free combination, The balance between the noise reduction and the incomplete elimination of ionospheric delay is finded. A new combination is proposed considering both the noise reduction and the ionospheric correction, which is called Iono-weak combination. According to the experiment, we reach the conclusion that the SPP result of ionosphere-weak combination is better than ionosphere-free combination with GPS data. The positioning accuracy of optimal combination can be increased about 15%. In allusion to the phenomenon that BDS ionosphere-weak combination can only improve the horizontal precision, the assumption of Total Group Delay (TGD) effect in BDS broadcast ephemeris is raised. The experiment result of BDS is consistent with GPS using ionosphere-weak combination without TGD correction. The rationality of the assumption is proved.

Keywords Ionospheric delay · Ionosphere-free combination · Ionosphere-weak combination · Standard point positioning

1 Introduction

High-precision positioning of satellite navigation systems is subject to a variety of additional conditions. The relative positioning is restricted by the distribution of the reference base station and the length of the base line. Precision Point Positioning (PPP) is limited by the limitation of timeliness due to the dependence on corrective information such as precise ephemeris, clock error and other correction issued by IGS [1–3]. Standard Point Positioning (SPP) is widely used in daily life due to its advantages of simple operation and less restriction. Ionospheric delay is an important source of error in pseudorange positioning. The effects of ionospheric

Y. Kong (✉) · H. Chai · Z. Pan · R. Wang · C. Liu
Institute of Surveying and Mapping, Zhengzhou 450001, China
e-mail: 492327387@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_10

delay are usually weaken by constructing ionosphere free combination measurements in case of dual frequency observations. In this process, un-modelled errors such as pseudorange noise and multipath effects are amplified [4–6].

An ionosphere weak combination is proposed in this paper to avoid noise amplification problem. Through the formula derivation, the scope of the combinatorial coefficient screening is determined. Based on the SPP experiments of the BDS/GPS dual system data, the optimal combination mode is selected respectively. The ionosphere weak combination positioning results of BDS and GPS are statistically analyzed and compared. In view of the difference between BDS and GPS, the conjecture is proposed and the reasonable explanation is given.

2 Dual Frequency Ionosphere Free Combination Model

Code pseudorange observation equation can be written as follow:

$$P_i = \rho + c \cdot dt + \delta_{ion(i)} + B^s_{(P_i)} + B_{r(P_i)} + \delta_{trop} + \delta_{mult(P_i)} + \varepsilon_{P(i)} \tag{1}$$

where i represents a different frequency, ρ is recorded as the geometric distance between the station and the satellite. $c \cdot dt$ means receiver clock error, $\delta_{ion(i)}$ denotes as the ionosphere delay of different frequency code observations, $B^s_{(P)}$ and $B_{r(P)}$ are satellite and receiver code hardware delay respectively, δ_{trop} denotes as troposphere delay, δ_{mult} is multipath effect and ε_P represents code pseudorange observation noise [7, 8].

Ionosphere delay is frequency dependent. Thus, the linear combination of the dual frequency measurement can be used to eliminate the effect of the ionosphere, that is, to obtain the combined observation of the ionosphere.

$$\begin{aligned} P_{IF} &= \frac{f_1^2}{f_1^2 - f_2^2} P_1 - \frac{f_2^2}{f_1^2 - f_2^2} P_2 \\ &= \rho + c \cdot dt + \Delta B^s_{(P1,P2)} + \Delta B_{r(P1,P2)} + d_{trop} + d_{mult(P1,P2)} + \varepsilon_{P(1,2)} \end{aligned} \tag{2}$$

where ΔB is recorded as the combined hardware delay [9, 10].

$$\Delta B = \frac{f_1^2}{f_1^2 - f_2^2} B_{(1)} - \frac{f_2^2}{f_1^2 - f_2^2} B_{(2)} \tag{3}$$

There are three different frequency in the BDS2 regional navigation system, such as: B1I, B2I and B3I. When the navigation ephemeris file is used for SPP, the different frequency combination formulas have the following forms:

$$P_{IF} = \frac{f_1^2 P_{\mathrm{B1I}} - f_2^2 P_{\mathrm{B2I}}}{f_1^2 - f_2^2} - \frac{c \cdot (f_1^2 T_{\mathrm{GD1}} - f_2^2 T_{\mathrm{GD2}})}{f_1^2 - f_2^2} \tag{4}$$

$$P_{IF} = \frac{f_1^2 P_{\mathrm{B1I}} - f_3^2 P_{\mathrm{B3I}}}{f_1^2 - f_3^2} - \frac{c \cdot f_1^2 T_{\mathrm{GD1}}}{f_1^2 - f_3^2} \tag{5}$$

where T_{GD1} denotes as the satellite hardware delay bias between B1I and B3I, T_{GD2} means the satellite hardware delay bias between B2I and B3I. Because the BDS2 navigation ephemeris clock error is estimated by B3I observations, the navigation satellite clock error is consistent with the hardware delay of the B3I frequency. No hardware delay correction is required when B3I frequency is used.

3 Theory of Ionosphere Weak Combination

3.1 *Noise Amplification Principle of Dual Frequency Combination*

According to Eq. (2), the ionosphere free combination can completely eliminate the first order ionospheric delay in theory. However, in the process of using the linear combination to eliminate the ionospheric delay, un-modeled error terms such as code observation noise and multipath effects will be magnified to a certain extent. For the convenience of research, the noise and multipath effects can be assumed to be Gaussian white noise. Assuming observed noise standard deviation of frequency i denotes as σ_i. According to the variance covariance propagation law, the combined observation noise can be deduced [11]:

$$\sigma_{IF} = \sqrt{(\alpha\sigma_1)^2 + (\beta\sigma_2)^2} \tag{6}$$

where α and β are called ionosphere free combination coefficient, determined by:

$$\begin{cases} \alpha = \dfrac{f_1^2}{f_1^2 - f_2^2} \\ \beta = \dfrac{f_2^2}{f_1^2 - f_2^2} \end{cases} \tag{7}$$

Assuming different frequency are the same precision observations, that is $\sigma = \sigma_1 = \sigma_2$.

Table 1 Different noise amplification factor in BDS/GPS

System	Combined frequency	μ
BDS2	B1I (1561.098 MHz) B2I (1207.140 MHz)	2.90
BDS2	B1I (1561.098 MHz) B3I (1268.520 MHz)	3.53
BDS2	B3I (1268.520 MHz) B2I (1207.140 MHz)	14.29
GPS	L1 (1575.420 MHz) L2 (1227.600 MHz)	3.00
GPS	L1 (1575.420 MHz) L5 (1176.450 MHz)	2.59
GPS	L2 (1227.600 MHz) L5 (1176.450 MHz)	16.64

$$\begin{cases} \sigma_{IF} = \mu \cdot \sigma \\ \mu = \sqrt{\alpha^2 + \beta^2} \end{cases} \tag{8}$$

where μ is recorded as noise amplification factor.

Noise amplification factor of different combinations in BDS/GPS are shown in Table 1.

As shown in Table 1, the frequency of similar frequency combinations will seriously pollute the quality of observations. GPS L1/L5 dual-band combination has the minimum noise amplification factor. According to the newest BDS ICD document, two new public service signals of B1C (1575.42 MHz) and B2a (1176.45 MHz) will be added to the global system BDS3. At that time, BDS3 will also have the smallest noise amplification factor as GPS.

3.2 Noise Reduction Principle Based on Ionosphere Weak Coefficient

Suppose that the code observation noise is 1% of the code element width. For BDS, the code element width is about 150 m. Code measurement error can be regarded as 1.5 m, which is equivalent to 1/6–1/5 of the peak ionospheric impact in the middle latitudes during the solar activity peak. The noise amplification effect of the dual-frequency code IF combination is significant in SPP. In this paper, we consider the way to reduce the effect of noise amplification by adjusting the combination coefficient. This method with the consideration of the ionospheric correction and noise amplification is called ionosphere weak combination.

Define $\eta_{IF}=\left(\frac{f_1}{f_2}\right)^2$, Eq. (2) can be written as:

$$P_{IF}=\frac{\eta_{IF}}{\eta_{IF}-1}P_1-\frac{1}{\eta_{IF}-1}P_2 \tag{9}$$

Assume η as a variable. η_{IF} is a special value of η. A functional relationship between the noise amplification factor and the variable can be established:

$$\mu=\frac{\sqrt{\eta^2+1}}{\eta-1} \tag{10}$$

where η denotes as ionosphere weak combination regulator. μ changes with η as shown in Fig. 1.

According to Fig. 1, μ decreases as η increases. The rate of decrease is gradually slowing down and converging to σ. Noise amplification and the ionospheric delay correction should be considered at the same time. The ionospheric delay amplification equation can be deduced as follow:

$$\begin{aligned}\Delta_{ion}&=\frac{\eta}{\eta-1}\cdot\frac{A}{f_1^2}TEC-\frac{1}{\eta-1}\cdot\frac{A}{f_2^2}TEC\\&=\frac{\eta-\eta_{IF}}{\eta-1}\cdot\frac{A}{f_1^2}TEC\\&=(1-\frac{\eta_{IF}-1}{\eta-1})\delta_{ion1}\end{aligned} \tag{11}$$

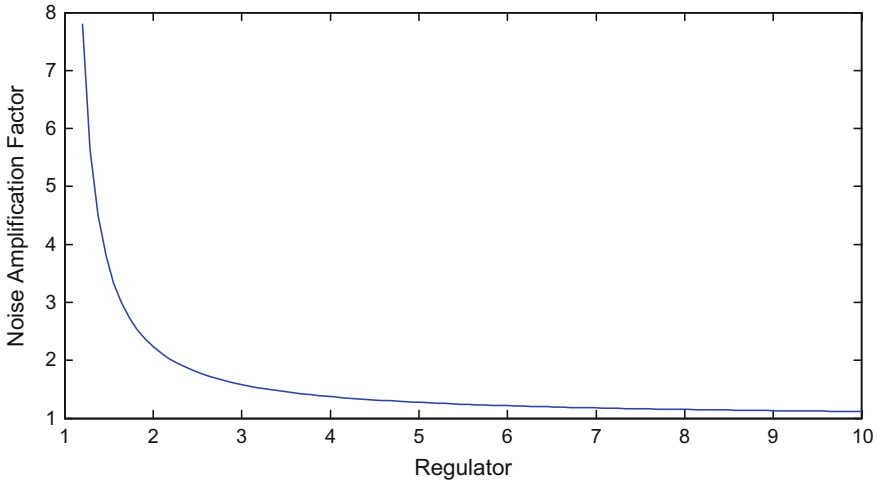


Fig. 1 Change of μ with η

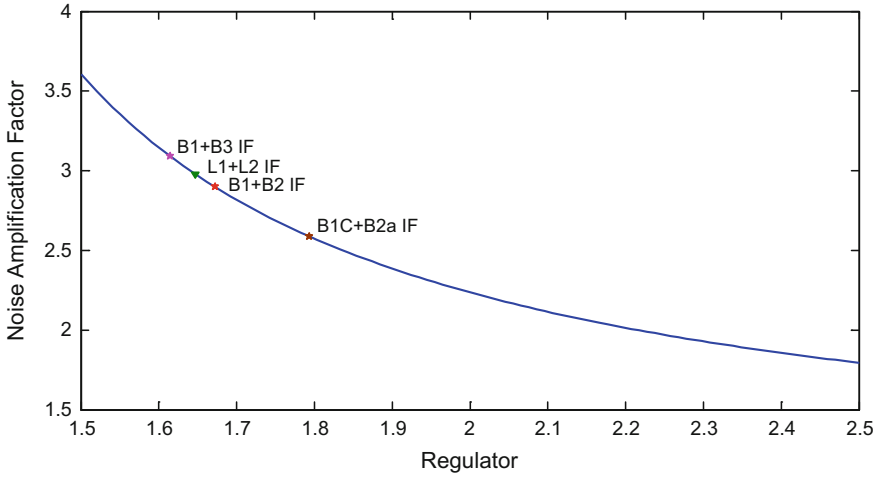


Fig. 2 Relation between BDS/GPS $\boldsymbol{\eta_{IF}}$ and $\boldsymbol{\mu}$

where Δ_{ion} is regarded as the ionospheric residual error in ionosphere weak combination, *TEC* means total electronic content, δ_{ion1} represents the ionospheric delay of frequency f_1. As shown in Fig. 2, the ionosphere weak combination with noise reduction should satisfy the condition $\eta > \eta_{IF} > 1$. The ionospheric residual error Δ_{ion} range is $(0, \delta_{ion1})$. In this paper, different ionosphere weak combination modes are obtained by changing η, and the SPP results of these modes are analyzed.

4 Statistics and Analysis of Positioning Result

For B1I + B2I dual-frequency combination in BDS, η range is $(1.47, 3.47)$, step variation $l = 0.2$. For L1+L2 dual-frequency combination in GPS, η range is $(1.45, 2.75)$, step variation $l = 0.1$. According to the multi-day observation data of CUT0 station of Curtin University in Australia, the SPP static positioning results under different combinations are obtained, shown in Tables 2, 3, Figs. 3 and 4.

According to the results from Tables 1 and 2, the accuracy of GPS SPP is significantly improved by ionosphere weak combination. Both horizontal and vertical positioning accuracy can be improved by about 15%. When $\eta = 1.95$, the accuracy increases to the extreme, and then the ionospheric delay residual become the main contradiction again, making the positioning accuracy began to decline. For BDS, ionosphere weak combination can improve the horizontal positioning accuracy of SPP, while it has the opposite effect of reducing the accuracy for vertical direction. By Eq. (4), BDS dual-frequency SPP need TGD correction. Thus, assuming that entirety TGD effect is opposite to ionosphere in BDS. If the broadcast TGD parameters have error, ionosphere weak combination will make the

Table 2 BDS B1I+B2I SPP positioning accuracy statistics of ionosphere weak combination

η	Positioning error RMS (m)				Improvement of positioning results (%)			
	N	E	U	2D	3D	U	2D	3D
1.47	3.13	3.38	5.37	4.61	7.07	6.4	−20	−2.5
IF	2.59	2.83	5.74	3.84	6.9	0	0	0
1.87	2.34	2.54	6.14	3.45	7.04	−7.0	10.2	−2.0
2.07	2.19	2.37	6.45	3.23	7.21	−12.4	15.9	−4.5
2.27	2.1	2.25	6.69	3.08	7.36	−16.6	19.8	−6.7
2.47	2.04	2.16	6.87	2.97	7.49	−19.7	22.7	−8.6
2.67	2	2.1	7.02	2.89	7.59	−22.3	24.7	−10.0
2.87	1.96	2.05	7.13	2.83	7.67	−24.2	26.3	−11.1
3.07	1.94	2.01	7.23	2.79	7.75	−26.0	27.3	−12.3
3.27	1.92	1.97	7.31	2.75	7.81	−27.4	28.4	−13.2
3.47	1.9	1.95	7.37	2.72	7.86	−28.4	29.2	−13.9

Table 3 GPS L1+L2 SPP positioning accuracy statistics of ionosphere weak combination

η	Positioning error RMS (m)				Improvement of positioning results (%)			
	N	E	U	2D	3D	U	2D	3D
1.45	2.82	1.82	6.7	3.36	7.49	−59.9	−54.8	−58.7
1.55	2.16	1.44	5.09	2.6	5.71	−21.5	−19.8	−20.1
IF	1.8	1.22	4.19	2.17	4.72	0	0	0
1.75	1.6	1.09	3.68	1.94	4.16	12.2	10.6	11.9
1.85	1.52	1.03	3.47	1.83	3.92	17.2	15.7	16.9
1.95	1.5	1	3.41	1.8	3.86	18.6	17.1	18.2
2.05	1.52	0.99	3.44	1.81	3.89	17.9	16.6	17.6
2.15	1.56	0.99	3.52	1.85	3.98	16.0	14.7	15.7

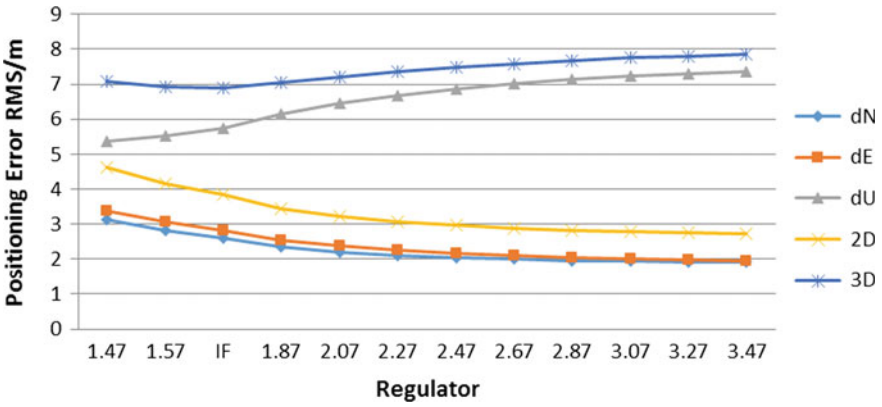


Fig. 3 BDS B1I+B2I ionosphere weak combination positioning result variation

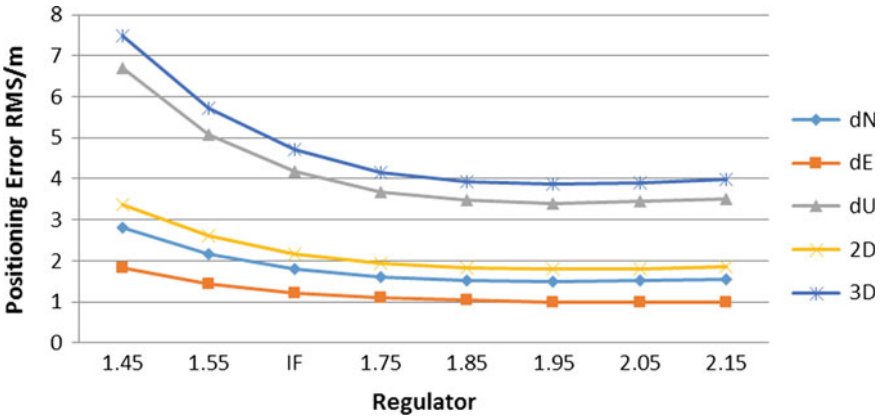


Fig. 4 GPS L1+L2 ionosphere weak combination positioning result variation

TGD error more significant. The ionospheric delay residual and satellite hardware delay residual have the same influence on the zenith direction, which leads to the reduction of the vertical positioning accuracy. Based on the above conjecture, the positioning results statistics are obtained for the same BDS data without TGD correction, as shown in Table 4 and Fig. 5.

The horizontal and vertical positioning accuracy can be improved by about 25% without TGD correction for BDS, even compared to dual frequency ionosphere free combination positioning results with TGD correction, there are nearly 15% of the increase effect. It verifies the existence of TGD correction error, and the entirety TGD is opposite to ionosphere delay.

Table 4 BDS B1I+B2I SPP positioning accuracy statistics of ionosphere weak combination

η	Positioning error RMS (m)				Improvement of positioning results (%)			
	N	E	U	2D	3D	U	2D	3D
1.47	5.4	5.68	6.8	7.83	10.37	−40.8	−25.7	−31.6
IF	4.13	4.66	4.83	6.23	7.88	0	0	0
1.87	3.5	4.13	4.2	5.42	6.86	13.0	13.0	12.9
2.07	3.13	3.8	4.04	4.93	6.37	16.4	20.9	19.2
2.27	2.89	3.57	4.05	4.6	6.13	16.1	26.2	22.2
2.47	2.73	3.41	4.13	4.37	6.01	14.5	29.9	23.7
2.67	2.61	3.29	4.22	4.2	5.95	12.6	32.6	24.5
2.87	2.52	3.19	4.31	4.07	5.93	10.8	34.7	24.7
3.07	2.46	3.12	4.4	3.97	5.92	8.9	36.3	24.9
3.27	2.4	3.05	4.47	3.89	5.93	7.5	37.6	24.7
3.47	2.36	3	4.54	3.82	5.93	6.0	38.9	24.7

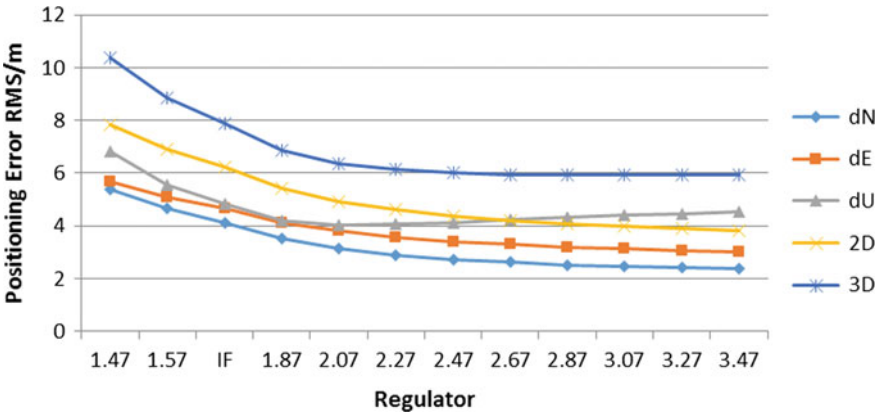


Fig. 5 BDS B1I+B2I (Un-TGD) ionosphere weak combination positioning result variation

5 Conclusions

(1) The noise amplification of dual-frequency ionosphere free combination SPP is analyzed in this paper. To solve this problem, the ionosphere weak combination mode is proposed for balancing the effect of noise amplification and ionospheric delay.
(2) Set the regulator factor η, The functional relationship between regulator and noise amplification factor is deduced, and the relationship between regulator and the ionospheric delay residual is also derived.
(3) According to BDS and GPS data, the effect of ionosphere weak combination SPP is analyzed. It is found that the optimal ionosphere weak combination can improve the horizontal and vertical positioning accuracy for GPS SPP. While in BDS system, the ionosphere weak combination can only improve the accuracy of the horizontal direction. Supposing that BDS broadcast TGD parameters have obvious error, and entirety TGD error have the same effect to ionosphere delay. BDS positioning results is similar to GPS without TGD correction. The accuracy of ionosphere weak combination SPP without TGD correction can be improved about 15% compared with ionosphere free SPP with TGD correction.

Acknowledgements This work is supported by China National Science Foundation of China (No: 41274045; 41374041; 41574010).

References

1. Liu JN, Chen JY, Zhang YP et al (1999) Wide area differential GPS theory and method. Surveying and Mapping Press, Beijing, China

2. Wübbena G, Schmitz M, Bagge A (2005) PPP-RTK: precise point positioning using state-space representation in RTK networks. In: Proceedings of international technical meeting of the Satellite Division of the Institute of Navigation, pp 2584–2594
3. Ye SR (2002) Theory and its realization of GPS precise point positioning using un-differenced phase observation. Ph.D. thesis, Wuhan University, Wuhan, China
4. Urquhart L (2009) An analysis of multi-frequency carrier phase linear combinations for GNSS. Senior technical report, Department of Geodesy and Geomatics Engineering Technical Report No. 263, University of New Brunswick, Fredericton, New Brunswick, Canada
5. Deng J, Pan SG, Hong ZZ (2014) A resolution method for ionospheric delay with optimal combination of three-frequency data. Geomatics Inf Sci Wuhan Univ 39(5):600–604
6. Chang ZQ, Hu XG, Su RR (2014) The positioning precision analysis with ionosphere free combination of regional satellite navigation system, Wuhan, China. J Geomatics Sci Technol 31(6):556–560
7. Li ZH, Huang JS (2010) GPS surveying and data processing. Wuhan University Press, Wuhan, Hubei
8. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solutions 5(2):12–28
9. Yuan YB (2002) Study on theories and methods of correcting ionospheric delay and monitoring ionosphere based on GPS. Ph.D. thesis, University of Chinese Academy of Sciences, Beijing, China
10. Sun ZM, Wang J, Gao JX (2008) Variation law of regional ionospheric TEC using dual-frequency GPS measurements. J Geomatics Sci Technol 25(3):199–201
11. Sui LF, Song LJ, Chai HZ (2010) Error theory and foundation of surveying adjustment. Surveying and Mapping Press, Beijing, China

An Investigation on Influence of Navigation Satellites Solar Panels on the RNSS Signal Propagation and Ranging Error

Xin Nie, Jun Xie, Tianxiong Liu, Songtao Huangfu, Shuguo Xie, Tian Jin and Xiaowei Cui

Abstract In order to analyze the influence of solar panels on the propagation of navigation signals, simulation analysis and ground far-field measurement campaign in the proving field were carried out. In this paper, the propagation of RNSS signal when solar panels rotated by different angles was simulated using the moment method. The effect of multipath caused by solar panels on the ranging errors was also studied. Simulation results are in agreement with the ground far-field measurement, which show that solar panels have little influence on the RNSS navigation signals. The influence on the ranging error is negligible.

Keywords RNSS signal · Solar panels · Ranging error

1 Introduction

There are many factors affecting the positioning accuracy of satellite navigation systems [1]. Most errors in satellite ranging, such as orbital error, satellite clock error, ionospheric and tropospheric delay, can be substantially corrected by models, while noise and interference can be reduced using spread-spectrum techniques, so multipath becomes one of the main sources of ranging error.

Due to the scarcity of satellite shape and layout data, the study of multipath around the satellite is very difficult, and the research results are few. In 2014, pseudo-range measurement of four satellite navigation systems was performed in [4]. The results show that the multipath (MP) value of the pseudo-range of BeiDou

X. Nie (✉) · J. Xie · T. Liu · S. Huangfu
Beijing Institute of Spacecraft System Engineering, Youyi Road, 100094 Haidian District, Beijing, China
e-mail: nxbuaa@126.com

S. Xie · T. Jin
Beihang University, Beijing, China

X. Cui
Tsinghua University, Beijing Shi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_11

systems is related to elevation and the time series of MP values of the GEO satellite is periodic. The MP values of navigation signal with different were compared in [5]. The variation of MP values with the time and elevation was also studied.

Equipments with large size installed on the surface of satellites, especially solar panels, reflector antennas, etc., may reflect the signal transmitted by satellites, resulting in multipath propagation. Multipath affects the performance of receivers, and the accuracy of ranging will degrade. In the meantime, due to the fact that the variation period of MP reported in the above thesis is about 24 h, which is the same as that of the solar panels rotating. In order to verify whether or not the change of MP values with the time is caused by solar panels, this paper carried out theoretical analysis and simulation of satellite multipath on one hand. According to the modeling and simulation analysis of the satellite, the multipath generated by the solar panel on-board is obtained by the method of moment. The amplitude ratio of multipath to line-of-sight (LOS) path, delay of multipath relative to the LOS path were obtained. On the other hand, one real-world GEO satellite was used in wireless far-field to carry out experiments on multipath to obtain the intensity and delay of reflected signal. Finally, the parameters of multipath were used to further obtain the ranging error.

2 Multipath Simulation Analysis

Moment method is mainly based on the integral equation of the electromagnetic field. The method can analyze the radiation pattern and the input impedance by analyzing the current distribution on the metal body and the antenna. It has the characteristics of high precision and is widely used in the analysis of such problems. The amplitude of multipath signal is smaller than that of the LOS signal, and the scatterers are located at the side lobe of antenna, so the moment method is chosen.

The body model used in the analysis can be found in [6]. After the parameters such as satellite size, solar panel size, reflection coefficient and antenna parameters are brought into the modeling software, the satellite's electromagnetic model can be obtained.

2.1 Simulation Results of Multipath Amplitude

According to the relationship between the satellite and the earth, the wavebeam need to cover the surface of the earth. The width of the beam is $\pm 8.7°$. Therefore, we will discuss only the multipath distributed within the main lobes. Antenna, astral shell, solar panels are set to PEC.

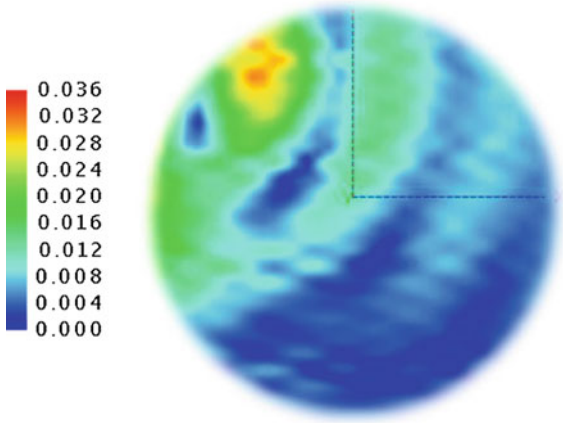


Fig. 1 The amplitude ratio of multipath to LOS in the main beam of GEO satellite

Figure 1 shows the multipath distribution when the earth is observed on the satellite. The upper and lower are the north and south directions of the earth respectively. The left and the right are respectively the west and east directions of the earth when observed in the space. It can be seen that the amplitude ratio of multipath to the LOS is not the same at different positions of the beam. The maximum ratio is 0.036. When moving in the vertical direction from lower to upper, the multipath amplitude is relatively small, and the amplitude ratio varies from 0 to 0.013 (which is 37 dB lower than the LOS). When moving along the horizontal direction from left to right, the amplitude ratio changes more obviously, and the amplitude ratio varies between 0.01 and 0.05 (which is 26 dB lower than the LOS).

2.2 Simulation Results of Multipath Delay

Based on the multipath simulation and extraction, the phase difference between the multipath and the LOS can be used to calculate the delay of the multipath relative to the LOS. According to the simulation result of the GEO satellite, the multipath delay is shown in Fig. 2.

According to this figure, it can be seen that the time delays of the multipath and the main path are not the same at different positions of the beam. The maximum time delay in this figure is 50 ns, corresponding to a distance of about 15 m. However, most of the time delay values are about 20 ns and the corresponding distance is 6 m, which is consistent with the distance from the antenna to the solar panel. The results prove the correctness of the simulation.

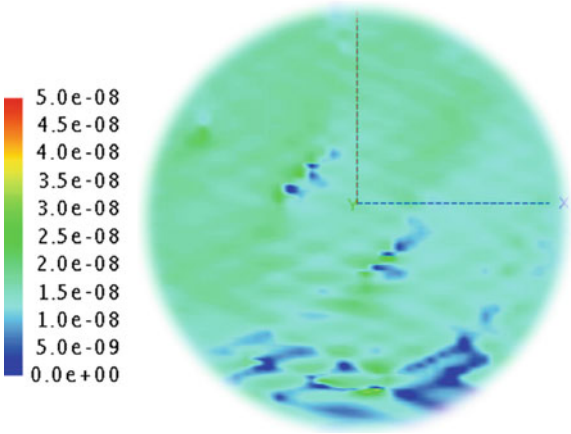


Fig. 2 The delay of multipath relative to LOS (unit: s)

3 Test Verification

Based on the simulation analysis, one real-world GEO satellite is used in wireless far-field to carry out experiments on multipath in order to obtain the reflected signal intensity and delay caused by solar panels. In this paper, sweep-frequency method and multi-frequency pseudo-range and carrier phase measurement combination method (Hereinafter referred to as MP method for simplicity) are used to evaluate the multipath.

The sweep-frequency method uses the sweep signal as a test signal. After being transmitted by the antenna, the LOS signal and the multipath signal are received by the receiving antenna. The amplitude and phase detection receiver will detect the amplitude and phase of the wideband swept frequency signal at different frequencies and then perform Fourier transform to distinguish the intensity and delay of the multipath signal and the LOS from the time domain. In theory, the wider the bandwidth of the sweep signal is, the higher the delay resolution is. The frequency sweep test results without solar panels were taken as the reference background. Then the background is subtracted from the frequency domain test results with solar panels, so the environment multipath is reduced. As a result, the resolution and sensitivity of multipath analysis in time domain is improved.

MP method is based on a widely used code-carrier phase MP value combination formula:

$$M(P_i, L_j, L_k) = MP_i - ML_j + \frac{\lambda_i^2 + \lambda_j^2}{\lambda_j^2 - \lambda_k^2}(ML_j - ML_k) + \zeta$$

The above equation can well evaluate the multipath effect of a certain frequency and realize the detection of code multipath. Without the impact multipath, the MP values should fluctuate at a constant value. The triple-frequency code-carrier combination can detect which frequency is affected by multipath. In the ground test environment, the Code minus carrier (CMC) value is considered to be the same as the MP value since there is no ionospheric influence. CMC is used to approximate pseudo-range multipath (ignoring very small carrier multipath and measurement noise).

$$CMC = \rho_{L1} - \phi_{L1} \cong 2I_{\rho L1} + MP_{\rho L1} + \varepsilon_{\rho L1} + \lambda_1 N_1$$

3.1 Wireless Far-Field Measurement

The GEO satellite was fixed in the semi-open microwave anechoic chamber, and the state of satellite is the same as working state on orbit. The measurement scenario is illustrated in Fig. 3. The solar panels are erected by a special tooling to simulate the effects of solar panels and their motion on signal reflection. The satellite transmitted the real navigation signal or the broadband sweep signal. The distance is so long that a plane wave was formed. The signal was received from the L-band antenna, and sent to the receiving equipment, which was synchronized with the transmitting equipment using optical fiber. The received signal was analysed to detect whether there is multipath. According to different test scenarios, the position and attitude of the satellite were adjusted by the tooling. To change the relative angle and distance relationship between the satellite and the receiving antenna, the attitude of the satellite, the rotation angle of the solar panels or the position of the receiving antenna was changed.

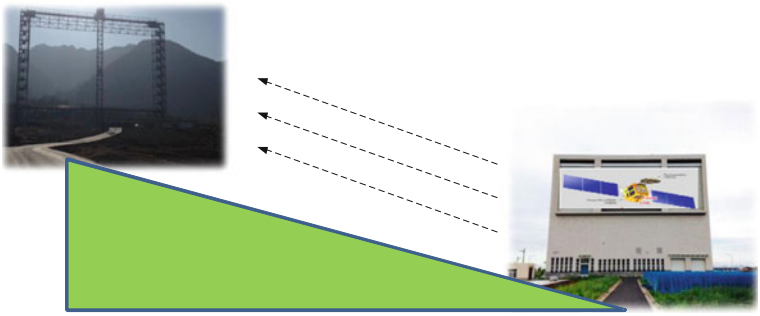


Fig. 3 Wireless far-field measurement scenario utilizing a real GEO satellite

3.2 Wireless Far-Field Test Results

Figure 4 shows test results using sweep-frequency method. Multipath introduced by the testing environment was cancelled in the frequency domain. The multipath due to solar panels is shown at ∇2 in the figure. As can be seen, the power of multipath was almost unchanged when solar panels were rotated to different directions. With respect to the LOS, the power of multipath caused by solar panels was −35 dB lower, which is consistent with the simulation results.

Test results using MP method is given in Fig. 5. The direction of the solar panels changes continuously downward by two cycles. The receiver was used to monitor the MP value when the solar panels were rotating. Test results show that the value of MP measured by the receiver shows no obvious change trend. The MP value is basically random jitter. Under the test condition, the standard deviation of the MP value is 0.046 m, while the standard deviation in the Gaussian white noise channel is about 0.03 m. The two deviations are basically the same. This shows that there is no correlation between the direction of the solar panel and the change of the MP value. The change of MP values on orbit is not caused by the solar panel rotation.

4 Ranging Error Analysis

Due to the presence of multipath signals, the received navigation signal can be represented as:

$$x(t) = \sum_{i=0}^{M} a_i(t) d(t) PN[t - \tau_i(t)] \cos[\omega_c t + \theta_i(t)]$$

The ranging error caused by multipath can be expressed as:

$$\varepsilon_0 = \begin{cases} \frac{\tau_{d1} a_1 \cos(\theta_1 - \hat{\theta}_0)}{a_0 + a_1 \cos(\theta_1 - \hat{\theta}_0)}, & 0 < \tau_{d1} \le \tau_p \\ \frac{a_1 d \cos(\theta_1 - \hat{\theta}_0)}{2a_0}, & \tau_p < \tau_{d1} \le \tau_q \\ \frac{a_1 \cos(\theta_1 - \hat{\theta}_0)}{2a_0 - a_1 \cos(\theta_1 - \hat{\theta}_0)} (T_c + \frac{d}{2} - \tau_{d1}), & \tau_p < \tau_{d1} \le T_c + \frac{d}{2} \\ 0, & \tau_{d1} > T_c + \frac{d}{2} \end{cases}$$

in which ε_0 multiplied by the speed of light c is the ranging error caused by multipath, τ_{di} is the time delay of multipath relative to the main path, a_0 is the main path amplitude, a_1 is the multipath amplitude, $\hat{\theta}_0$ is the carrier phase estimation of the direct signal, θ_1 is the phase of the multipath signal, and d is the interval,

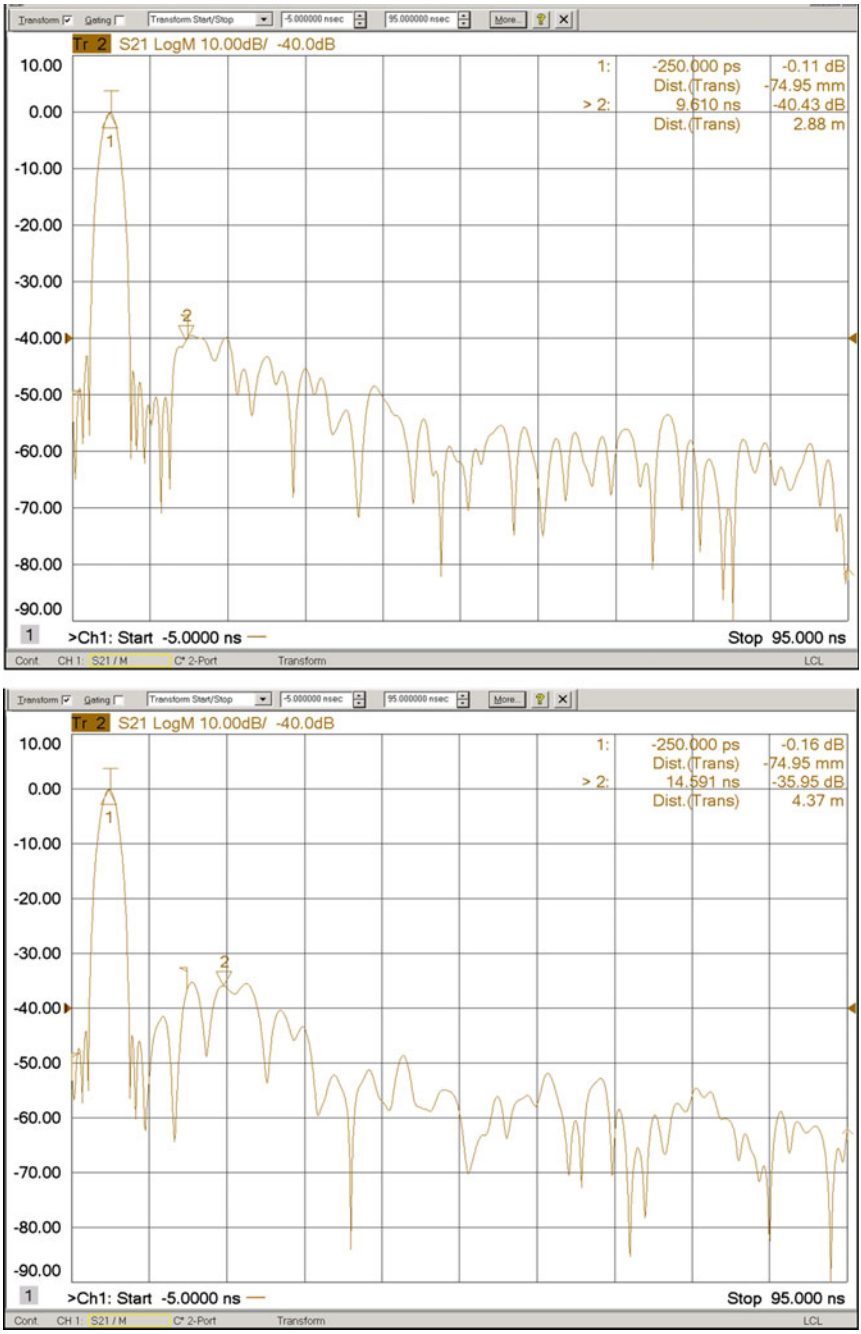


Fig. 4 Test results using sweep-frequency method (upper: solar panels forward and lower: solar panels down)

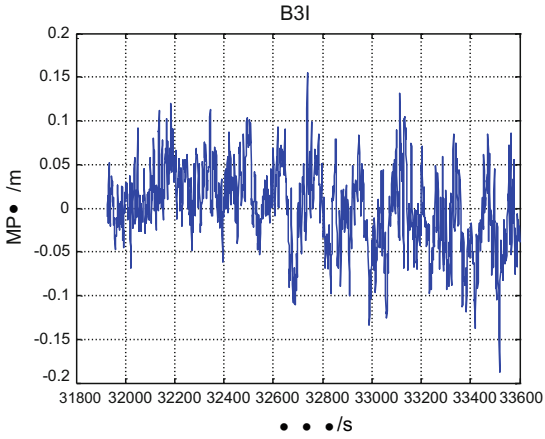


Fig. 5 Random jitter of MP values when solar panels change direction

$$\tau_p = \left[a_0 + a_1 \cos\left(\theta_1 - \hat{\theta}_0\right)\right] d/2a_0 \quad \text{and} \quad \tau_q = T_c - d/2 + a_1 d \cos\left(\theta_1 - \hat{\theta}_0\right)/2a_0.$$

The amplitude, phase and time delay of multipath signal were obtained through the simulation. These three parameters decide the ranging error of receivers. According to the formula, the ranging error caused by multipath can be obtained. The effects of solar panel rotation on the multipath were simulated and analysed. As the solar panel rotates, the ranging error of the satellites change with the rotation of the solar panel. Figure 6 is the ranging error map when the solar panel is rotated by 0°, 30°, 60°, 90°, 120° and 150°.

Beijing was chosen as a typical area. When the position of Beijing in the beam is brought in, it can be found that the ranging error changes with the rotation of the solar panel at this position. As shown in Fig. 7, it can be seen that when the solar panel rotates with time, the variation of ranging error in Beijing is about 1.5 cm.

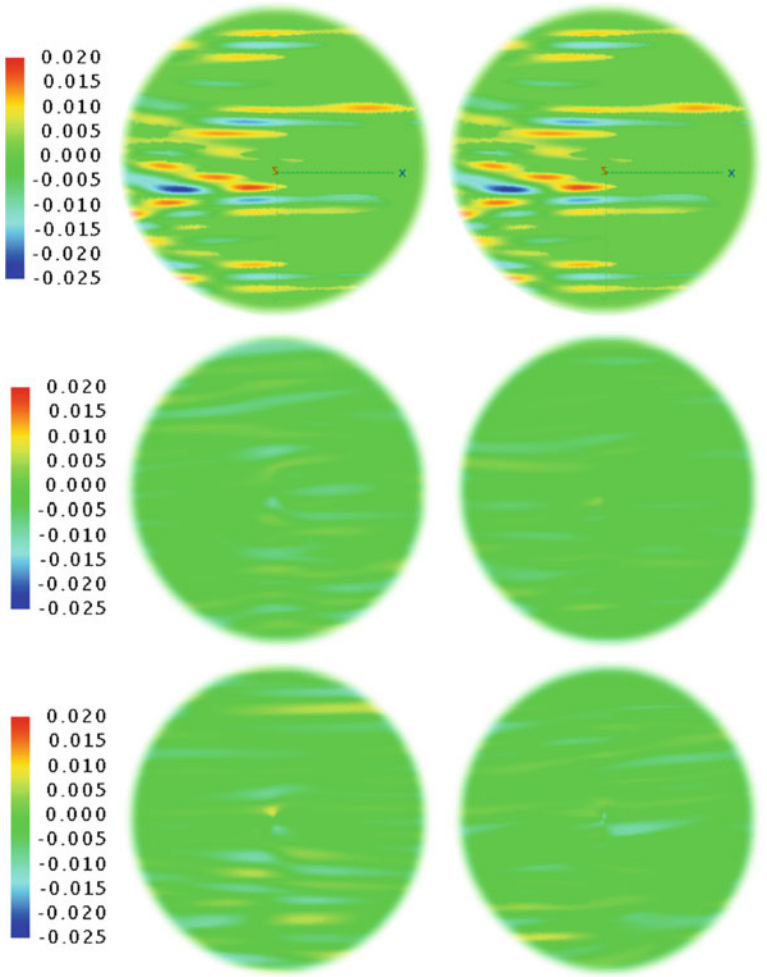


Fig. 6 Ranging error plot for solar panels rotated by a variety of angles

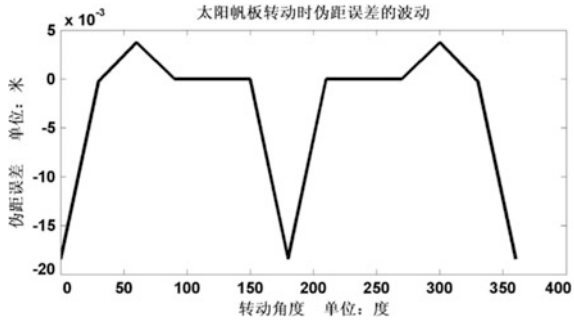


Fig. 7 Fluctuations in ranging error when the solar panel is rotating

5 Conclusion

In order to analyze the influence of satellite solar panels on the transmission of RNSS navigation signals, simulation analysis and the far-field measurement of a real satellite are respectively carried out. The simulation results are consistent with the far-field test results on the ground. Both the simulation and measurement results show that the power of multipath brought by solar panels is −35 dB below to the LOS. When the direction of satellite solar panels changes, the MP value is basically random jitter. The standard deviation of MP value is 0.046 m, which is basically equivalent to the standard deviation in Gaussian white noise channel. It is calculated that when the solar panel rotates with time, the variation of ranging error in Beijing is about 1.5 cm, which shows that the solar panels have less influence on the RNSS navigation signals and have negligible impact on the ranging error.

References

1. Xu G, Xu Y (2016) GPS: theory, algorithms and applications. Springer, Berlin
2. Kaplan ED, Hegarty C (2005) Understanding GPS: principles and applications. Artech House
3. Ng Y, Gao GX (2016) Direct position estimation utilizing non-line-of-sight (NLOS) GPS signals, ION GNSS+, 2016, Portland, Oregon, September 2016, pp 1279–1284
4. Wanninger L, Beer S (2015) BeiDou satellite-induced code pseudo range variations: diagnosis and therapy. GPS Solutions 19(4):639–648
5. Zhao Q, Wang G, Liu Z, et al (2016) Analysis of BeiDou satellite measurements with code multipath and geometry-free ionosphere-free combinations. Sensors 16(1):123
6. Montenbruck O, Schmid R, Mercier F et al (2015) GNSS satellite geometry and attitude models. Adv Space Res 56(6):1015–1029

Research on the Availability Analysis Method of Navigation Satellite Based on Petri Nets

Haisheng Li, Heng Zheng, Haitao Zhao and Zixia Zheng

Abstract Availability is an important index for measuring the service performance of navigation satellite. The traditional availability analysis methods only take into account the satellite failure factors, but not consider the satellite performance status. An availability analysis method of navigation satellite based on Petri nets is proposed in this paper, considering both the two factors of failure and performance status. With the consideration of three key navigation tasks (satellite uplink signal injection, navigation information processing, and navigation signal broadcasting), the satellite availability model, which describes the health of the single machine and the quality of the navigation signal, is firstly established by the colored Petri nets (CPN) on the basis of information flow. Then, the results of simulating analysis of satellite availability are obtained using Monte Carlo (MC) simulation algorithm.

Keywords Navigation satellite · Information flow · CPN · Outages Availability analysis · MC

1 Introduction

Navigation satellite is a core part of space segment of satellite navigation system. It is also the output terminal of the system. It broadcasts the navigation signal directly to users and provides navigation and timing information. The satellite health is directly related to the quality of the navigation signal. Availability of space signals, the ratio of the working time of the satellite to the expected working time, is an index specified in the BDS open service performance standard and is also an important index of the normal use of navigation satellites. Outage is a key factor that causes navigation satellite unusable. It refers to the state in which the satellite

H. Li (✉) · H. Zheng · Z. Zheng
China Academy of Aerospace Standardization and Product Assurance, Beijing, China
e-mail: leehaisheng@126.com

H. Zhao
Beijing Institute of Spacecraft System Engineering, Beijing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_12

cannot perform its required functions. The reason for the outage may be failure, maladjustment or planned and unplanned events. Outages are divided into four types: short term scheduled outages, short term unscheduled outages, long term scheduled outages, long term unscheduled outages. Short term unscheduled outages mean that the navigation signal is interrupted due to a sudden failure or other unscheduled events. Such outages cannot be informed to users in advance and are unpredictable and become a key factor restricting the availability of navigation satellites [1, 2].

The traditional satellite availability analysis methods only consider the satellite fault factors, not involving the performance of the satellite. This paper presents a Petri-based navigation satellite availability analysis method to analyze the availability of navigation satellites based on two factors of failure and performance. For the basic navigation services of RNSS and three key navigation tasks, firstly, the satellite availability model is established through the colored Petri net based on information flow. The model can be used to simulate the navigation data operation process and describe the single machine health status and navigation signal quality. Secondly, single machines are determined according to the satellite core tasks. Finally, Monte Carlo simulation algorithm is used to obtain satellite short term unscheduled outage simulation results and carry out satellite availability analysis, in which process the single machine performance parameters and reliability/maintainability data are as the input for the model.

2 Navigation Satellite Availability Model

The space segment is composed of multiple satellites, and each navigation satellite is established its own availability model. Based on the process of navigation satellite information flow related to the RNSS basic navigation services, a navigation satellite availability model is established by using Petri net method.

2.1 Brief Description of Petri Nets

Petri net is a kind of binary directed graph, which can be used for static structural analysis and dynamic behavior analysis. It is made up of place, transition, directed arc between place and transition, and token in place. The place represented by circles is static and used to record the status of the individuals that make up the system and the system itself. While the role of the transition (represented by rectangles) is to dynamically change the state of the system [3]. Basic Petri nets have simple definitions of token meanings and transition conditions, which are inconvenient to describe the system model. In comparison, the colored Petri net is an advanced net system defined on the basis of basic Petri net. It is more abstract

than basic Petri net: there are many resources available at each place, and each transition represents a variety of changes [4, 5].

2.2 Navigation Satellite Information Flow

For the basic services of RNSS, navigation satellite broadcasts space signals containing time, navigation information and the integrity information generated by the ground segment. Navigation satellite mainly completes three key tasks of uplink injection signal reception, navigation information processing and navigation signal transmission. The information flow process of navigation satellite is shown in Fig. 1.

2.3 Satellite System-Level Petri Net Model

The colored Petri net is used to establish the navigation satellite information flow model to simulate the flow process of navigation data. The short term unscheduled outage parameters are associated with the satellite key tasks' outage index.

As shown in Fig. 2, the Petri net model of the satellite system with basic navigation services includes three tasks: uplink injection signal reception, navigation information processing and navigation signal transmission. Satellite signal is connected with the monitoring station through a space signal, and connected with the injection station/master station through navigation data and integrity data.

The satellite-system-level Petri net model established by Petri net software TimeNET is shown in Fig. 3. The supplementary launch process in the model takes into account the time feature,including the ground test time, the transmitting time and the on-orbit test time. In addition, the ground resources needed for the

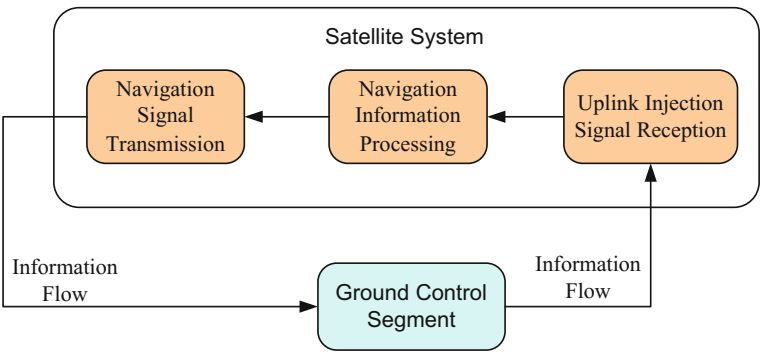


Fig. 1 The information flow of navigation satellite

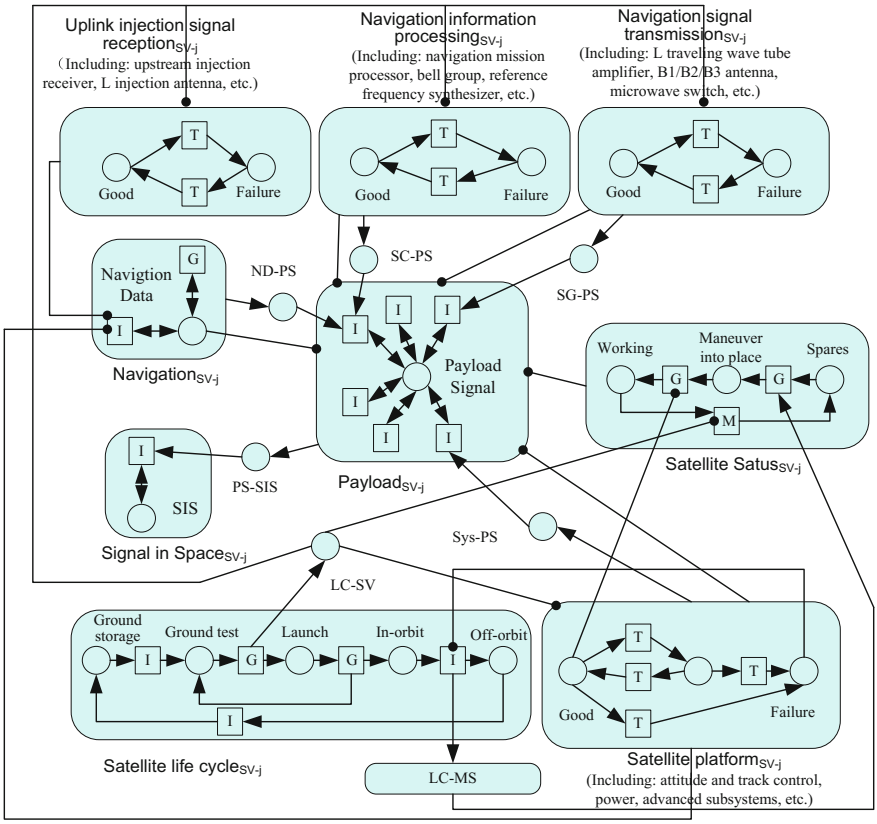


Fig. 2 The schematic diagram of satellite Petri net model

supplementary launch process are also considered, which is closely related to launch vehicle reliability (with launch reliability 0.96 as a constraint).

Table 1 shows the input parameters of the satellite Petri net model.

The success criteria for satellite's three key missions are as follows:

(1) Uplink injection signal reception

It can correctly receive the uplink radio signals and the navigation message parameters generated by the ground segment.

(2) Navigation information processing

It can store and process the navigation message parameters injected by the ground segment, integrate the satellite time reference information, and form the correct navigation message. In addition, it should have the functions of information encryption and decryption, uploading the injection program and automatically recovery from instantaneous logic failure.

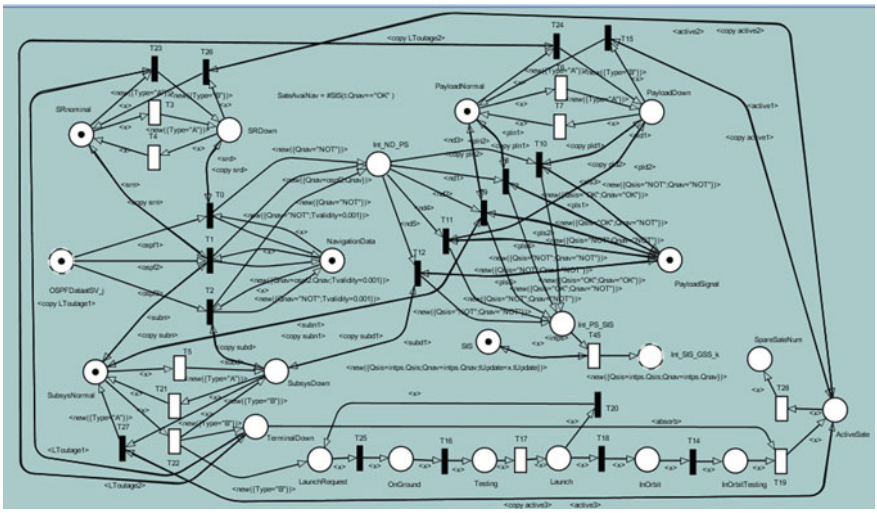


Fig. 3 The diagram of satellite system Petri net model

Table 1 Input parameters of satellite system Petri net model

Parameters	Description
$MTBO_{receiver}(sat)$	MTBO for uplink injection signal reception
$MTTR_{receiver}(sat)$	MTTR for uplink injection signal reception
$MTBO_{navdataproc}(sat)$	MTBO for navigation information processing
$MTTR_{navdataproc}(sat)$	MTTR for navigation information processing
$MTBO_{sisbrocast}(sat)$	MTBO for navigation signal transmission
$MTTR_{sisbrocast}(sat)$	MTTR for navigation signal transmission
$T_{testing}$	Ground test time
T_{launch}	Launching time
$T_{orbittesting}$	Satellite in-orbit test time
p_{launch}	Launch success rate

Note MTBO is the abbreviation of mean time between outages. MTTR is the abbreviation of mean time to repair

(3) Navigation signal transmission

It can broadcast a healthy space signal to users.

2.4 Satellite Functional Unit Petri Net Model

The main functional units used to accomplish the three tasks of navigation satellite are navigation task processing, navigation signal generation, navigation signal broadcasting, on-board time frequency maintenance, L band uplink reception and inter-satellite link.

For each functional unit, a corresponding Petri net model is established to describe the process of the unit transiting from normal working mode to failure mode. Figure 4 shows the Petri net model for the unit of navigation task processing. Others are similar.

The input parameters of the Petri net model for each functional unit of the satellite are shown in Table 2.

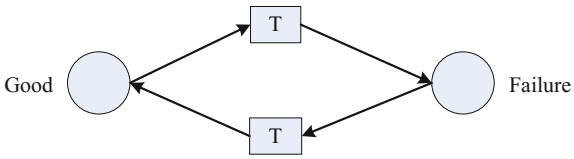


Fig. 4 Model for the state change of navigation task processing functional unit

Table 2 Input parameters of functional unit Petri net models

Parameters	Description
$MTBO_1(funit)$	MTBO for navigation task processing unit
$MTTR_1(funit)$	MTTR for navigation task processing unit
$MTBO_2(funit)$	MTBO for navigation signal generation unit
$MTTR_2(funit)$	MTTR for navigation signal generation unit
$MTBO_3(funit)$	MTBO for navigation signal broadcasting unit
$MTTR_3(funit)$	MTTR for navigation signal broadcasting unit
$MTBO_4(funit)$	MTBO for on-board time frequency maintenance unit
$MTTR_4(funit)$	MTTR for on-board time frequency maintenance unit
$MTBO_5(funit)$	MTBO for L band uplink reception unit
$MTTR_5(funit)$	MTTR for L band uplink reception on unit
$MTBO_6(funit)$	MTBO for inter-satellite link unit
$MTTR_6(funit)$	MTTR for inter-satellite link unit

3 Satellite Availability Simulation Analysis

The functional units and corresponding related single machines of the satellite basic navigation services are analyzed. The satellite short term unscheduled outages are simulated by using Monte Carlo algorithm based on the Petri net information flow model of the satellite system, which takes the reliability and maintainability parameters (including test data and simulation data) of the key single machines as inputs.

(1) Simulation conditions

The simulation time is 2,000,000 h, and the number of samples is 1000.

(2) Input conditions

Table 3 shows the outage parameters of the uplink injection signal reception, navigation information processing, and navigation signal transmission related to the basic navigation mission of the satellite. Other performance parameters refer to the functional performance indexes, which are omitted here.

Table 3 Outage parameters of satellite for RNSS (unit: hour)

No.	Uplink injection signal reception		Navigation information processing		Navigation signal transmission	
	MTBO	MTTR	MTBO	MTTR	MTBO	MTTR
1	8760	3	17,520	3	17,520	3
2	8760	2	17,520	2	17,520	2
3	8760	1	17,520	1	17,520	1
4	8760	0.5	17,520	0.5	17,520	0.5
5	5840	3	11,680	3	11,680	3
6	5840	2	11,680	2	11,680	2
7	5840	1	11,680	1	11,680	1
8	5840	0.5	11,680	0.5	11,680	0.5
9	4380	3	8760	3	8760	3
10	4380	2	8760	2	8760	2
11	4380	1	8760	1	8760	1
12	4380	0.5	8760	0.5	8760	0.5
13	3504	3	7008	3	7008	3
14	3504	2	7008	2	7008	2
15	3504	1	7008	1	7008	1
16	3504	0.5	7008	0.5	7008	0.5
17	2920	3	5840	3	5840	3
18	2920	2	5840	2	5840	2
19	2920	1	5840	1	5840	1

(continued)

Table 3 (continued)

No.	Uplink injection signal reception		Navigation information processing		Navigation signal transmission	
	MTBO	MTTR	MTBO	MTTR	MTBO	MTTR
20	2920	0.5	5840	0.5	5840	0.5
21	2503	3	5006	3	5006	3
22	2503	2	5006	2	5006	2
23	2503	1	5006	1	5006	1
24	2503	0.5	5006	0.5	5006	0.5
25	2190	3	4380	3	4380	3
26	2190	2	4380	2	4380	2
27	2190	1	4380	1	4380	1
28	2190	0.5	4380	0.5	4380	0.5
29	1947	3	3894	3	3894	3
30	1947	2	3894	2	3894	2
31	1947	1	3894	1	3894	1
32	1947	0.5	3894	0.5	3894	0.5
33	1752	3	3504	3	3504	3
34	1752	2	3504	2	3504	2
35	1752	1	3504	1	3504	1
36	1752	0.5	3504	0.5	3504	0.5
37	1593	3	3186	3	3186	3
38	1593	2	3186	2	3186	2
39	1593	1	3186	1	3186	1
40	1593	0.5	3186	0.5	3186	0.5
41	1460	3	2920	3	2920	3
42	1460	2	2920	2	2920	2
43	1460	1	2920	1	2920	1
44	1460	0.5	2920	0.5	2920	0.5

(3) Simulation results

On the basis of the satellite system-level Petri net model and Monte Carlo simulation algorithm, the outage parameters of the three functional units (uplink signal receiving, navigation information processing and navigation signal broadcasting) are considered, and then the short term unscheduled outage parameters and availability of a satellite are obtained. Simulation analysis results are shown in Table 4. The simulation results of satellite availability under different outages are shown in Fig. 5.

Table 4 Results of simulating analysis of satellite availability

No.	Availability	No.	Availability
1	0.9993	23	0.9993
2	0.9996	24	0.9995
3	0.9998	25	0.9972
4	0.9999	26	0.9983
5	0.9988	27	0.9991
6	0.9991	28	0.9995
7	0.9997	29	0.997
8	0.9998	30	0.9979
9	0.9986	31	0.9989
10	0.999	32	0.9995
11	0.9995	33	0.9965
12	0.9998	34	0.9977
13	0.9982	35	0.9989
14	0.9989	36	0.9993
15	0.9993	37	0.9962
16	0.9997	38	0.9974
17	0.9978	39	0.9988
18	0.9986	40	0.9993
19	0.9994	41	0.9959
20	0.9996	42	0.9969
21	0.9975	43	0.9988
22	0.9985	44	0.9992

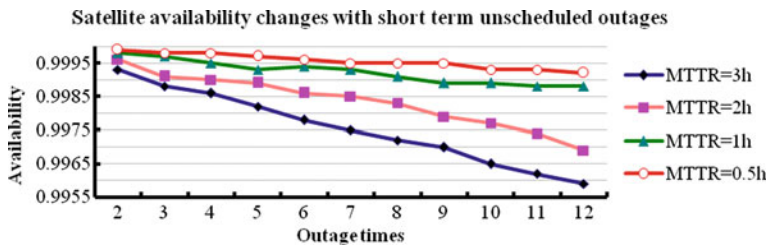


Fig. 5 The curve for results of simulating analysis of satellite availability

4 Conclusions

Navigation satellite is the output terminal of satellite navigation system, which broadcasts navigation signals directly to users. Satellite health is directly related to the quality of navigation signals. Availability is an important index for measuring the service performance of navigation satellite. The traditional availability analysis methods only take into account the satellite failure factors, but not consider the

satellite performance status. An availability analysis method of navigation satellite based on Petri nets is proposed in this paper, considering the two factors of failure and performance. With the consideration of three key navigation tasks (satellite uplink signal injection, navigation information processing, and navigation signal broadcasting), the satellite availability model, which describes the health of the single machines and the quality of the navigation signal, is firstly established by the colored Petri nets on the basis of information flow. Then, the results of simulating analysis of satellite availability are obtained using Monte Carlo simulation algorithm.

References

1. Zheng H, Li H, Yang Z (2011) Analysis of constellation availability of satellite navigation system. Acospace Contr
2. Li H (2013) Research on spare satellites strategy of navigation constellation based on system availability. In: China Satellite Navigation Conference (CSNC) 2013 Proceedings, Wuhan
3. Analysis techniques for dependability—Petri net techniques. IEC 62551.2012
4. DCPN-based models for dependability assessment. Volume 3, 2002
5. Assessment of Galileo key dependability and safety parameters: a simulation-based approach. In: Proceedings of ION national technical meeting, Long Beach, CA, September 2005

Tracking Performance Assessment of Tiered Code Effect on GNSS Signal in the Presence of CWI

Jianlei Yang, Hongjun Ye, Xiude Chen and Shenying Hui

Abstract In the signal structure design of new global navigation satellite system (GNSS), the ideas of secondary/tiered code was introduced. However, due to the inherent vulnerability of satellite navigation signal, frequency interference (RFI) has become a potential threat. The continuous wave interference (CWI) has become one of the most serious interference types for the characteristics of power concentration, multiple types and wide distribution of the interference source. The impact assessment of CWI on the performance of navigation signals has become a hot spot of research. However, the existing assessment method lacks the quantitative evaluation for the effect of the secondary code on signal performance in the presence of CWI. To solve the above problems, the impact of tiered codes with different length on signal power spectrum was analyzed firstly, and then influence of tiered codes on tracking performance by discrete spectrum carrier-to-noise ratio was analyzed. The experimental results show that the traditional analysis model based on continuous spectrum is no longer suitable for navigation signal performance analysis under the condition of CWI; the smaller the line distance is, the two times the code period is doubled and the line spacing is 0.1 times that of the original; and when coherent integration time doubles and the tracking performance deteriorates 2–3 dB; and when length of tiered code increase, the tracking performance can improve 2–3 dB; and short coherent integration time and long tiered code should be used in receiver design. A compromise needs to be made between these two parameters to ensure the tracking performance under CWI and the robustness of receivers. The proposed method can provide effective theoretical guidance for the development of the new GNSS signal receiver.

Keywords GNSS · Continuous wave interference · Tiered code
Secondary code · Tracking error · Interference tolerance

J. Yang (✉) · H. Ye · X. Chen · S. Hui
State Key Laboratory of Satellite Navigation System and Equipment Technology, Shijiazhuang, China
e-mail: yangjianlei555@163.com

J. Yang · H. Ye · X. Chen · S. Hui
The 54th Research Institute of CETC, Shijiazhuang, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_13

1 Introduction

With the advent of the new global navigation satellite system (GNSS), such as European Galileo, American modernized GPS and BeiDou Navigation Satellite System (BDS) of China, new signal design ideas have been introduced to meet the growing demand of navigation and positioning services. Among the several novel ideas, the tiered code and simultaneous distribution of the data and pilot channels are the most obvious characteristics. However, the vulnerability of new GNSS signals to radio frequency interference (RFI) is still one of the major issues of many applications, because satellite navigation signal has very low power, e.g. −155 to −160 dBW, which is under the noise floor level [1, 2]. Among different types of interferences, continuous wave interference (CWI) has serious adverse effects on the quality of the received GNSS signal [3, 4].

Many researches have been carried out to analyze the RFI effects on tracking performances. The theoretical model of effective carrier-to-noise ratio $(C/N_0)_{\text{Eff}}$ was derived in [5], and it was compared with precorrelation carrier-to-noise ratio (C/N_0). The code tracking error and $(C/N_0)_{\text{Eff}}$ was analyzed at different locations of narrowband interference (NBI) in [6]. All these analysis models were established on the assumption that power spectrum density (PSD) of the signal is continuous. However, in a real scenario, the spectrum consisted of a train of discrete spectral lines and CWI signal may not coincide with the spectral line of signal in some frequency positions. As a result, the above mentioned models are not reasonable when used to analyze effects of CWI on tracking performance [7].

Recently, effects of CWI on tracking performance were analyzed in detail with the consideration of line spectrum of pseudo random noise (PRN) code. Jian [1] and Balaei [3] proposed the effective C/N_0 model. Besides $(C/N_0)_{\text{Eff}}$, the model of carrier and code tracking error were also presented in [1]. However, all these analyses focused on GPS L1 C/A analysis without considering the characteristics of new GNSS, such as tiered code and simultaneous distribution of the data and pilot channels. Based on the above analysis, it can be seen that few studies have been done on analyzing tracking performance while taking the characteristics of new GNSS into consideration.

Focusing on the above problem, we derived expressions of interference tolerance, carrier and code tracking error under CWI. This paper provides the analysis of $(C/N_0)_{\text{Eff}}$ while considering the effect of the tiered code. The numerical results provide several possible ways for receiver and signal structure designs in the presence of CWI.

2 Signal Model Based on Discrete Spectrum in the Presence of CWI

The model of GNSS signals composed of data and pilot components but without CWI was given in [7]. In a similar way, after down conversion and digitization, the received intermediate frequency (IF) signal in the presence of CWI can be expressed as

$$\begin{aligned} S_{IF}(n) &= S_D[n] + S_P[n] + S_I[n] + \eta_{IF}[n] \\ &+ \sqrt{P'_s} e_D[n - f_s\tau] \cos(2\pi n(f_{IF} + f_{Do})T_s + \phi) \\ &+ \sqrt{P'_s} e_P[n - f_s\tau] \sin(2\pi n(f_{IF} + f_{Do})T_s + \phi) \\ &+ \sqrt{P'_I} \sin(2\pi n(f_{IF} + f_{Do} + f_I)T_s + \varphi) + \eta_{IF}[n] \end{aligned} \tag{1}$$

where P'_{s} and P'_{I} are the power of signal and CWI in Watt (W), with $P'_{\mathrm{s}} = 10^{P_{\mathrm{S}}/10}$ and $P'_{\mathrm{I}} = 10^{P_{\mathrm{I}}/10}$, while P_{S} and P_{I} are the power of signal and CWI in dBW; n is the epoch; τ and ϕ are the code delay and phase introduced by the transmission; f_s is the sampling rate [Hz]; $\eta_{\mathrm{IF}}[n]$ is the independent identical density (i.i.d) of white Gaussian noise (WGN) with PSD N'_0 in W/Hz with $N'_0 = 10^{N_0/10}$, and N_0 is the PSD of $\eta_{\mathrm{IF}}[n]$ in dBW/Hz; f_{IF}, f_{Do} and f_{I} are the receiver intermediate frequency, Doppler frequency and frequency offset of CWI from $f_{\mathrm{IF}} + f_{\mathrm{Do}}$ in Hz; φ is the phase of CWI; and $e_{\mathrm{D}}[n]$ and $e_{\mathrm{P}}[n]$ are data and pilot components, and can be expressed as

$$\begin{aligned} e_{\mathrm{D}}[n] &= d[n]sc[n]sec_{\mathrm{D}}[n]c_{\mathrm{D}}[n] \\ e_{\mathrm{P}}[n] &= sc[n]sec_{\mathrm{P}}[n]c_{\mathrm{P}}[n] \end{aligned} \tag{2}$$

where $d[n]$ is the navigation message; $sc[n]$ is the subcarrier; $sec_{\mathrm{D}}[n]$ and $sec_{\mathrm{P}}[n]$ are the secondary code sequence in tiered code construction; and $c_{\mathrm{D}}[n]$ and $c_{\mathrm{P}}[n]$ are the PRN code sequences. For the binary phase-shift keyed (BPSK) signals (such as GPS L1 C/A and L5C, Galileo E5, E6 CS and E1 OS, etc.) and GNSS signals without tiered code (such as GPS L1 C/A and L1C data and Galileo E1 OS data, etc.), $sc[n] = 1$ and $sec[n] = 1$, respectively.

Assume $f_{\mathrm{IF}} = f_{\mathrm{Do}} = 0$, and the PSD of GNSS signal $S_{\mathrm{D}}[n]$ or $S_{\mathrm{P}}[n]$ can be expressed as follows 1

$$G(f) = T_{\mathrm{Int}} \sum_{k=-\infty}^{\infty} \mathrm{sinc}^2((f - kf_{\mathrm{Dis}})T_{\mathrm{Int}})|C(k)|^2 S(k) \tag{3}$$

where T_{Int} is the coherent integration time [sec]; $C(k) = C_{\mathrm{PRN}}(k)/N$ is the normalized discrete Fourier transform (DFT) or PSD of PRN code, and $C_{\mathrm{PRN}}(k)$ is the N-point DFT or PSD of a finite PRN code; k is the spectral line index; f_{Dis} is the spectrum line repetition frequency in Hz; and $S(k)$ is the spectrum term related to a BPSK or binary offset carrier (BOC) [7] pulse shape, as shown in Eq. (4).

$$S(k)=\begin{cases}\frac{\sin^2\left(\pi\frac{k}{N}\right)}{\left(\pi\frac{k}{N}\right)^2} & \text{for BPSK}\\ \frac{\sin^2\left(\pi\frac{k}{N}\right)}{\left(\pi\frac{k}{N}\right)^2}\tan^2\left(\frac{\pi k}{mN}\right) m \text{ even} \quad \text{or} \quad \frac{\cos^2\left(\pi\frac{k}{N}\right)}{\left(\pi\frac{k}{N}\right)^2}\tan^2\left(\frac{\pi k}{mN}\right) m \text{ odd} & \text{for sine-BOC}\\ 4\frac{\sin^2\left(\pi\frac{k}{N}\right)}{\left(\pi\frac{k}{N}\right)^2}\left(\frac{\sin^2\left(\frac{\pi k}{2mN}\right)}{\cos\left(\frac{\pi k}{mN}\right)}\right)^2 m \text{ even} \quad \text{or} \quad 4\frac{\cos^2\left(\pi\frac{k}{N}\right)}{\left(\pi\frac{k}{N}\right)^2}\left(\frac{\sin^2\left(\frac{\pi k}{2mN}\right)}{\cos\left(\frac{\pi k}{mN}\right)}\right)^2 m \text{ odd} & \text{for cos-BOC}\end{cases} \tag{4}$$

where $m = f_{\mathrm{PRN}}/f_{\mathrm{SC}}$, f_{PRN} and f_{SC} are the rate of PRN code and subcarrier of BOC signal, respectively.

The PSD of signal based on continuous spectrum is also plotted as comparison, and the corresponding expression [8] is shown in Eq. (5).

$$G_{\mathrm{S}}(f) = \frac{1}{f_{\mathrm{PRN}}}\mathrm{sinc}^2\left(\frac{f}{f_{\mathrm{PRN}}}\right) \tag{5}$$

3 Effective Carrier-to-Noise Ratio Model Based on Continuous and Discrete Spectrum

The traditional effective carrier-to-noise ratio $(C/N_0)_{\mathrm{Eff,Inter,Trad}}$ in the presence of interference can be expressed as Eq. 5

$$(C/N_0)_{\mathrm{Eff,Trad}} = \frac{P_{\mathrm{S}}}{N_0}\left[1+\frac{P_{\mathrm{I}}}{N_0}\frac{\int_{-B/2}^{B/2} G_{\mathrm{I}}(f)G_{\mathrm{S}}(f)df}{\int_{-B/2}^{B/2} G_{\mathrm{S}}(f)df}\right]^{-1} \tag{6}$$

where, P_{S}/N_0 is the carrier to noise ratio without noise; P_{I} and P_{S} is the interference and useful signals respectively; $P_{\mathrm{I}}/P_{\mathrm{S}}$ is the ratio of interference to received signal; Q is the quality factor of interference, defined as

$$Q = \frac{\int_{-\infty}^{\infty}|H_{\mathrm{R}}(f)|^2 G_{\mathrm{S}}(f)df}{f_{\mathrm{PRN}}\int_{-\infty}^{\infty}|H_{\mathrm{R}}(f)|^2 G_{\mathrm{I}}(f)G_{\mathrm{S}}(f)df} = \frac{\int_{-\infty}^{\infty}|H_{\mathrm{R}}(f)|^2 G_{\mathrm{S}}(f)df}{f_{\mathrm{PRN}}\kappa_{\mathrm{IS}}} \tag{7}$$

where $H_{\mathrm{R}}(f)$ is the transmission function of the receiver with maximum amplitude 1; κ_{IS} is the Spectral Separation Coefficient (SSC) defined as

$$\kappa_{\mathrm{IS}} = \int_{-\infty}^{\infty}|H_{\mathrm{R}}(f)|^2 G_{\mathrm{I}}(f)G_{\mathrm{S}}(f)df \tag{8}$$

The effective carrier-to-noise ratio in the presence of CWI $(C/N_0)_{\text{Eff,CWI}}$ can be expressed as [1]

$$(C/N_0)_{\text{Eff,CWI}} = \frac{\int_{-B/2}^{B/2} G_{\text{S}}(f)df}{\frac{N_0}{P_{\text{S}}}\int_{-B/2}^{B/2} G_{\text{S}}(f)df + \frac{P_{\text{I}}}{P_{\text{S}}} T_{\text{Int}} \text{sin c}^2((f_{\text{I}} - kf_{\text{Dis}})T_{\text{Int}})|C(k)|^2 S(k)} \quad (9)$$

where B is the two-sided bandwidth of pre-correlation filter [Hz]; $G(f)$ is the continuous spectrum PSD.

4 Numerical Results and Analysis

The tracking performance of GPS L5 [10] signal with $N_{\text{PRN}} = 2$ is taken as an example to analyze in this section. In the experiments, $f_{\text{IF}} = f_{\text{Do}} = 0$ Hz, $B = 20.46$ MHz and $\Delta = 0.5$ chip. The PRN codes are generated at the rate of 10.23 MChip/s and the period of 1 ms. The data and pilot channels are modulated by 10 bit Neuman-Hofman (NH) codes ($\text{NH}_{10} = 0000110101$) and 20 bit NH codes ($\text{NH}_{20} = 00000100110101001110$), respectively. For pilot channel, the first 10 bit NH code of pilot channel is used when $N_{Sec} = 10$ bit in simulation.

By using the tiered code structure, the effective code length of data and pilot channels is extended to 10 and 20 ms, respectively. In the experiment, to make a fair comparison of tracking performance, the first ten bit of NH_{20} is used as the NH code when $T_{\text{Int}} = 10$ ms for pilot channel.

4.1 The Effect of Tiered Code on Signal PSD

For the old GNSS signal (i.e., the GPS L1C/A), there is no secondary code, and PRN code sequence repeats every PRN code period T_{PRN} [sec]. Therefore, the PRN code has a line spectrum with lines $1/T_{\text{PRN}}$ [Hz] apart. Tiered code is introduced in new GNSS signal structure to increase the effective code length so that the distance between the spectral line becomes $1/(T_{\text{PRN}}N_{Sec})$, where N_{Sec} is the length of secondary code. In Fig. 1, the PSD of GPS L5 with PRN number $PRN_{\text{Nu}} = 2$ and different N_{Sec} have been plotted. The GPS L5 is made up of data and pilot channels and has the tiered code structure. The PSD of signal based on continuous spectrum is also plotted as comparison.

From Fig. 1, it can be found that when $N_{\text{Sec}} = 10$ and 20 bit, the distance of two spectral line is reduced to 0.1 and 0.05 kHz from 1 kHz (no NH code). The traditional PSD of continuous spectrum model is not consistent with the real PSD of signal.

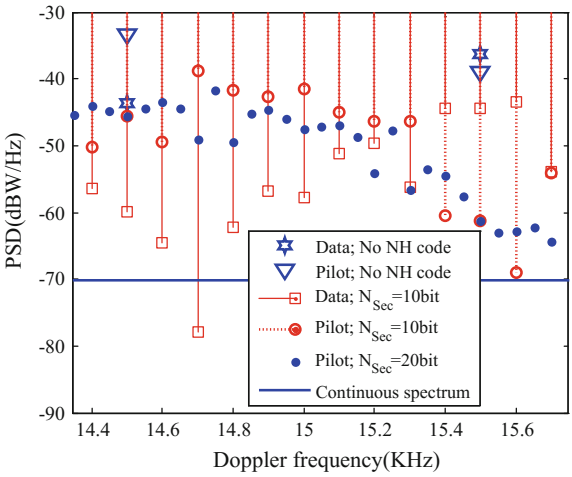


Fig. 1 PSD comparison of PRN code with or without secondary code

4.2 *Signal Performance at Different Spectral Line Positions*

To analyze the relationship between the tracking performance and spectral line position, $(C/N_0)_{\mathrm{Eff,CWI}}$, with different spectral line positions are plotted in Fig. 2. The corresponding continuous spectrum models $(C/N_0)_{\mathrm{Eff,Trad}}$ are also plotted for comparison. In the experiment, the GPS L5 pilot PRN code with P_S/N_0 = 44 dB Hz, P_I/P_S = 25 dB, B_{PLL} = 20 Hz, B_{DLL} = 5 Hz and T_{Int} = 10 ms is adopted. The CWI sweeps from 14.4 to 15.4 kHz.

From Fig. 2, it is clearly shown that the traditional continuous spectrum models are not suitable to analyze the tracking performance in the presence of CWI, because the real PSD of PRN code has 0.1 kHz intervals, and the bandwidth of

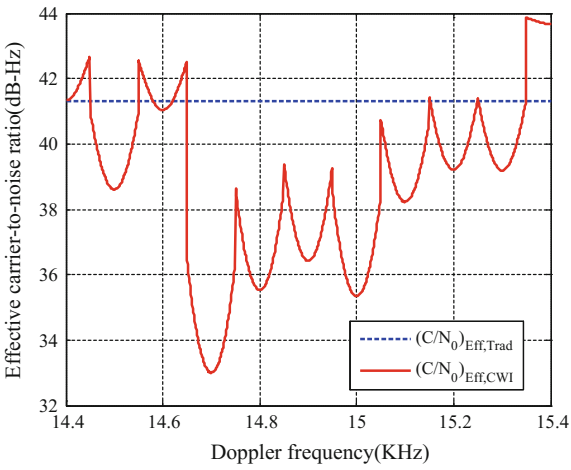


Fig. 2 Effective carrier-to-noise ratio at different spectral line positions

CWI can not cover the PRN code spectrum line. Therefore, the discrete spectrum PSD is not smooth into continuous spectrum [9].

From Figs. 1 and 2, we can see that when the position of CWI coincides with a PRN code spectrum line, the tracking performance will degrade rapidly, and that the larger power the spectrum line has, the worse the tracking performance is. As shown in Fig. 1, the maximum power of spectrum line is −38 dBW/Hz at 14.7 kHz for pilot when $N_{\text{sec}} = 10$ bit, the corresponding $(C/N_0)_{\text{Eff,CWI}}$ is the lowest in Fig. 2.

Similar with Ref. [11], in order to analyze the tracking performance when CWI sweeps through the whole bandwidth of signal, the normalized average tracking performance (NATP) was defined as Eq. (10), and plotted in Fig. 3 for different tracking performances.

$$NATP(k) = \frac{\frac{1}{k}\sum_{i=1}^{k} TP(i)}{\max\{NATP(1), NATP(2), \ldots, NATP(N)\}} \quad k = 1, \ldots, N \tag{10}$$

where $TP(i)$ and N is the ith tracking performance value and length of $(C/N_0)_{\text{Eff,CWI}}$ respectively.

From Fig. 3, we can see that the value of NATP is the lowest near 0, and the tracking performance is the worst. The NATP value is the largest at the positive center position of the signal spectrum, that is, when CWI is in the center of the signal power spectrum, it has little influence on the performance of the signal.

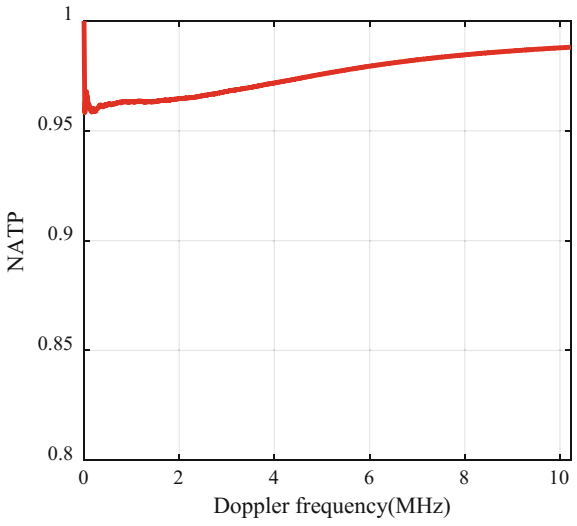


Fig. 3 The NATP of effective carrier-to-noise ratio

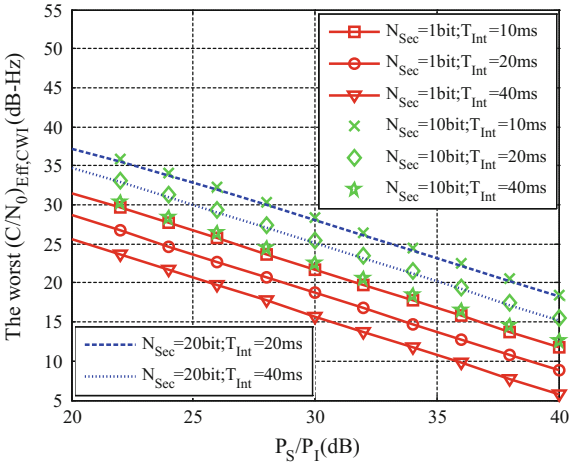


Fig. 4 $(C/N_0)_{\mathrm{Eff,CWI}}$ performance comparison under different N_{Sec} and T_{Int}

4.3 Effect of Secondary Code on Signal Performance

The PRN code modulated by different secondary codes can result in different PSD of spectral line (as shown in Fig. 1) and chip sequences. Different chip sequences have different tracking performances under the same CWI [11]. To analyze the tracking performance of different chip sequences of GPS L5 in Ref. 10, the worst relative code tracking error was used.

The $(C/N_0)_{\mathrm{Eff,CWI}}$ are analyzed for PRN code modulated by different NH codes, as shown in Fig. 4. In the experiment, the CWI sweeps from −10.23 to 10.23 MHz, P_S/N_0 = 44 dB Hz, B_{DLL} = 5 Hz, B_{PLL} = 10 Hz.

From Fig. 4, we can see that tracking performance becomes worse with the increase of T_{Int} at the same N_{Sec}. As shown in Fig. 4, under the same N_{Sec}, $(C/N_0)_{\mathrm{Eff,CWI}}$ deteriorates by 2–3 dB when T_{Int} increases from 10 to 20 ms or 20 to 40 ms.

Figure 4 also shows that the increase of N_{Sec} can improve the tracking performance. Under the same T_{Int}, the tracking performance can improve by 6–7 dB, 9–10 dB and 2–3 dB when N_{Sec} increases from 1 to 10 bit, 1 to 20 bit and 10 to 20 bit, respectively.

5 Conclusion

When the position of CWI coincides with a PRN code spectrum line, the tracking performance will degrade rapidly, and the larger power the spectrum line has, the worse the tracking performance is. The traditional continuous spectrum models are not suitable to analyze the tracking performance in the presence of CWI; Short T_{Int}

and long N_{Sec} can improve the tracking performance in the presence of CWI. In addition, too long N_{Sec} increases the acquisition complexity of weak signal case such as positioning in thick forests or indoor application; and short T_{Int} can degrade the robustness of receivers. Therefore, a balance needs to be made between tracking performances with N_{Sec} and T_{Int}.

In order to better analyze the impact of the two encoding on the performance of the navigation signal, the future research work will be carried out as follows: (1) the influence of secondary code on BOC, AltBOC and QMBOC; (2) The impact of Pilot channel introduced on anti CWI for GNSS signal; (3) the effect of secondary code on code tracking performance, carrier tracking performance and interference tolerance in the presence of CWI.

Acknowledgements This work was supported by National Key Research and Development Plan of China (Grant No. 2016YFB0502402).

References

1. Jang J, Paonni M, Eissfeller B (2012) CW Interference effects on tracking performance of GNSS receivers. IEEE Trans Aerosp Electron Syst 48(1):243–258
2. Ojeda OA, Grajal J, Lopez-Risueno G (2013) Analytical performance of GNSS receivers using interference mitigation techniques. IEEE Trans Aerosp Electron Syst 49(2):885–906
3. Balaei AT, Dempster AG, Lo Presti L (2009) Characterization of the effects of CW and pulse CW interference on the GPS signal quality. IEEE Trans Aerosp Electron Syst 45(4):1418–1431
4. Kim SY, Kang CH, Yang JH, Park CG, Joo JM, Heo MB (2012) A GNSS interference detection method based on multiple ground stations. J Korean GNSS Society 1(1):15–21
5. Ross JT, Leva JL, Yoder S (2001) Effect of partial-band interference on receiver estimation of C/N0: measurements. In: Proceedings of the 2001 national technical meeting of the institute of navigation, January 22–24, Long Beach, CA, USA; 2001, pp 829–838
6. Betz JW (2000) Effect of narrowband interference on GPS code tracking accuracy. In: Proceedings of the 2000 national technical meeting of the institute of navigation, January 26–28, Anaheim, CA, USA; 2000, pp 16–27
7. Borio D, O'Driscoll C, Lachapelle G (2009) Coherent, noncoherent, and differentially coherent combining techniques for acquisition of new composite GNSS signals. IEEE Trans Aerosp Electron Syst 45(3):1227–1240
8. Betz JW (2001) Binary offset carrier modulations for radionavigation. Navigation 48(4):227–246
9. Kaplan ED, Hegarty CJ (2005) Understanding GPS principles and applications, 3rd edn. Artech House, Norwood
10. Interface Control Working Group. IS-GPS-705 Revision A (2010) Navstar GPS space segment/user segment L5 interfaces [internet]. Available from http://www.gps.gov/technical/icwg/IS-GPS-705A.pdf
11. Liu Y, Ran Y, Ke T et al (2011) Code tracking performance analysis of GNSS signal in the presence of CW interference. Sig Proc 91(4):970–987

The Assessment of GNSS Measurements from Android Smartphones

Kaishi Zhang, Fangtan Jiao and Jianwen Li

Abstract Android, the most popular smart operating system with an approximately 86% market share, has been able to support devices in exporting GNSS raw measurement. Taking an example of Huawei P9 smartphone, the paper mainly illustrated the way to obtain GNSS raw measurements on Android OS platform. Moreover, the quality of the satellite signal, the accuracy of GNSS measurements obtained from Android smartphone were analyzed. It can be approved that under good observing conditions, the mean Carrier-to-noise Ratio (CNR) of the satellite signal received by smart mobile devices with Android operating system can usually reach above 25 dBHz, the pseudorange noise of each satellite is about 8–12 m, and the cycle clip ratio of carrier phase measurements is extremely high, the carrier phase range noise can be within 0.006 m when no cycle slip. Compared with common GNSS geodetic receivers, due to the smartphone's hardware, the CNR is lower, and the quality of measurements needs to be improved.

Keywords Android · Pseudorange · Carrier phase · Carrier-to-noise ratio

1 Introduction

"Android", the name of the humanoid robot in the science fiction L'Eve Future written by French author Auguste Villiers de L'lsle-Adam, which represents the most popular smart operating system now. In 2003, Andy Rubin founded Android

K. Zhang · J. Li (✉)
Zhengzhou Institute of Surveying and Mapping, Zhengzhou,
Henan Province, China
e-mail: zzljw@126.com

K. Zhang
e-mail: cashzhang1995@163.com

F. Jiao
TSINGHUA University High School, Beijing, China
e-mail: jiaowh0927@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_14

Inc. and began to research and develop the intelligent operating system. In 2005, Google Inc. acquired Android as well as its R&D team. In November 2007, Android operating system was roll-out, and a global Open Handset Alliance composed of 34 software, hardware and telecommunication enterprises was established to develop open standards for Android devices. In 2008, Google released the initial version of Android, approved by the US Federal Communications Commission. As the system continued updating and optimizing, Android has been widely accepted by customers. By the first quarter of 2017, the market share of Android has reached about 86% [1].

The early version Android can export the location such as longitude, latitude and altitude with GNSS controller, and provides the related application program interface (API) to obtained location. However, due to the commercial confidentiality requirement of the manufacturer, the operating system encapsulates the variables of GNSS measurements, so that it is impossible for developers and users to obtain the GNSS measurements. In 2016, the Android Nougat system was released by Google Inc. added the APIs to export the available GNSS raw measurements, which brought a chance to get the pseudo-range and carrier phase measurements to achieve precision positioning by the mobile phone.

At present, the Android smart mobile devices that support the output of GNSS raw observations include HUAWEI's Honor 8 and 9, HUAWEI's P9 and P10, HUAWEI's mate 9, Samsung's S8, Google's Pixel and Nexus 6P, 5X phones and Nexus 9 tablets. The main feature of those device is listed in Table 1 [2].

This paper briefly introduces the Android system library about GNSS raw measurements and APIs to obtain measurements. Moreover, it assessed the measurement data collected by Huawei P9 smartphone in Carrier-to-Noise Ratio(C/N0), Pseudorange Noise, Ranging Precision and Cycle Slip.

Table 1 Comparison between compatible devices

Model	Pseudorange rate	Accumulated delta range	Satellite constellation
Honor 9	√	√	GPS/GLONASS
S8 (Exynos)	√	√	GPS/GLONASS/Galileo
S8 (QCOM)	√	×	GPS
P10	√	√	GPS/GLONASS/Galileo/ BDS
P10 Lite	√	×	GPS
P9	√	√	GPS/GLONASS/BDS
Pixel	√	×	GPS
Nexus 6P	√	×	GPS
Nexus 9	√	√	GPS
Nexus 5X	√	√	GPS

2 Location Libraries and GNSS Raw Measurements

Compared with the earlier versions, new classes had been added into the Location library in Android Nougat, such as GnssMeasurement class, GnssNavigation class, GnssClock class and so on.

The GnssClock class, which represents a measurement of the receiver's clock, mainly contains the timestamp information. As Android is developed based on the Linux kernel, the clock set in Android devices can be classified into the Real Time Clock (RTC) and the System Clock. The System Clock is a "software" clock to users. The RTC is a "Hardware" Clock (HWClock), which runs independently from the operating system. The RTC is maintained by an integrated circuit in the smart devices. In the previous version of Android, developers and users cannot obtain HWClock timestamps as it is encapsulated. The GnssClock class provides methods to get HWClock, so the developers can compute the GPS time of Android devices.

GnssMeasurement class mainly contained raw measurements and computed information, such as Constellation Types, Carrier Frequency, Space Vehicle ID, Carrier-to-Noise Ratio, Received Satellite Time, Pseudorange Rate, Accumulated Delta Range, etc. And GnssNavigationMessage class contained a GNSS satellite Navigation Message. These classes is indispensable to compute pseudorange and carrier phase measurements so as to position precisely on Android devices.

3 Quality Assessment

With a NovAtel GNSS receiver set together, a Huawei P9 (EVA-AL10) smartphone was adopted to collect measurements on the rooftop station in Zhengzhou, Henan Province, without electromagnetic interference detectable and obstacles above 10° of elevation angle around. The data during 15:44:37–00:15:17, Nov. 11th, 2017 (GPST) was selected as the sample with 1 s interval, 30,642 valid epochs and 227,831 valid measurements.

3.1 Carrier-to-Noise Ratio

As shown in Fig. 1, in good observing condition, the Carrier-to-Noise Ratio of signal received by Huawei P9 mainly distributed within 20–35 dBHz.

As shown in Fig. 3, the CNR of signal from G30 satellite decreased as the elevation angle of satellite descend. Referring to the observation data, the mean CNR value is 28.88 dBHz. When the elevation angle was below 30°, the CNR was under 25 dBHz. And it reached above 30 dBHz only with a elevation angle over 40°. However, in the same condition, the mean CNR value of NovAtel DL-V3 receiver was 44.64 dBHz (Fig. 2).

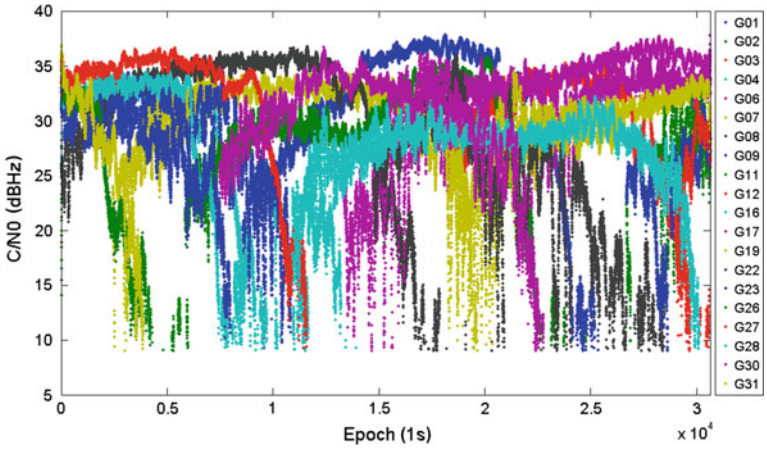


Fig. 1 Distribution of C/N0 on Huawei P9

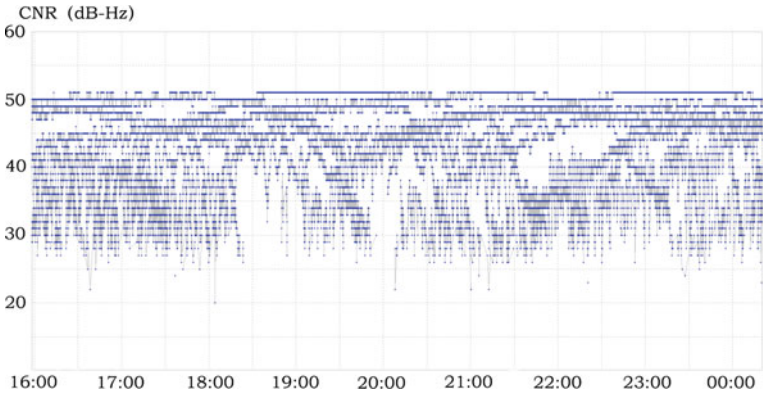


Fig. 2 Distribution of C/N0 on NovAtel receiver

3.2 Pseudorange Measurements

3.2.1 Theoretical Pseudorange Noise

Pseudorange noise is one of the key point of the quality assessment for navigation devices. Since PN code ranging is achieved by accurate tracking of PN code, the noise of code tracking is the decisive factor of range measurement. The principles were shown as follow.

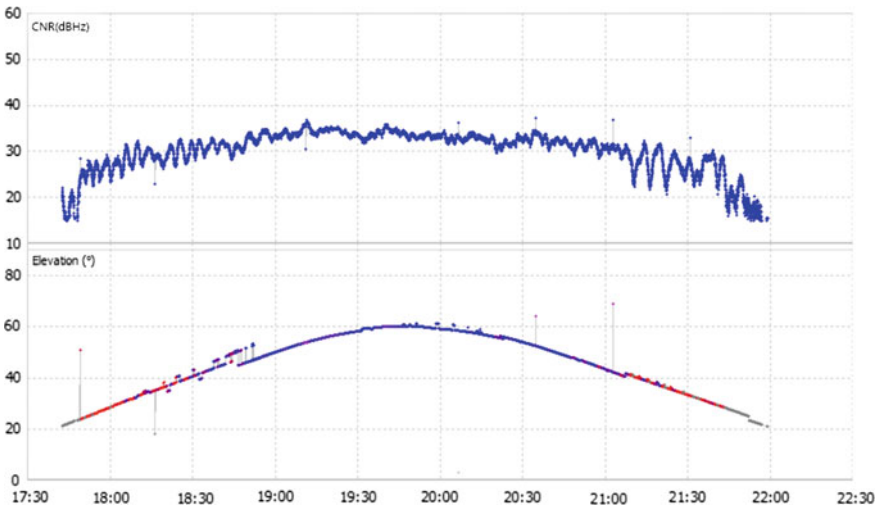


Fig. 3 C/N0 and elevation (G30)

$$\sigma_{CELP}^2 = \frac{K_1}{C_s/N_0} \tag{3.1}$$

$$\sigma_{NELP}^2 = \sigma_{CELP}^2 \left[1 + \frac{K_2}{C_s/N_0}\right] \tag{3.2}$$

$$K_1 = \frac{B_L(1 - 0.5B_L T)\int_{-\beta_r/2}^{\beta_r/2} G_s(f)\sin^2(\pi f\Delta)df}{(2\pi)^2\left(\int_{-\beta_r/2}^{\beta_r/2} fG_s(f)\sin(\pi f\Delta)df\right)^2} \tag{3.3}$$

$$K_2 = \frac{B_L(1 - 0.5B_L T)\int_{-\beta_r/2}^{\beta_r/2} G_s(f)\cos^2(\pi f\Delta)df}{T\left(\int_{-\beta_r/2}^{\beta_r/2} fG_s(f)\cos(\pi f\Delta)df\right)^2} \tag{3.4}$$

C_s/N_0 is Carrier-to-Noise Ratio; σ_{CELP}^2 is the tracking error of coherent delay-locked loop; σ_{NELP}^2 is that of noncoherent delay-locked loop. Most of GNSS receivers adopt the noncoherent delay-locked loop. And K_1, K_2 are the coefficients which is related to loop bandwidth B_L, integrate time T, filter bandwidth β_r, Code correlation interval Δ. $G_s(f)$ is the normalized power spectral density with infinite bandwidth, and $\int_{-\beta/2}^{\beta/2} G_s(f)df = 1$ [3]. Referring to the equation above, it is obvious that when CNR gets higher, the code tracking error increases as the range noise reducing.

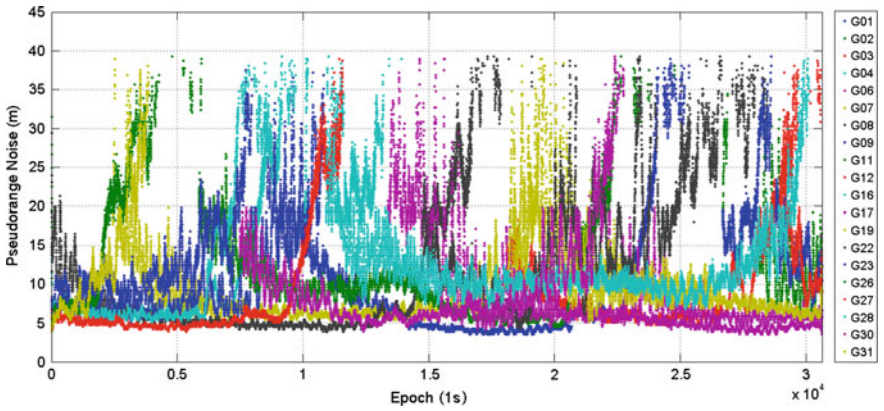


Fig. 4 Distribution of pseudorange noise

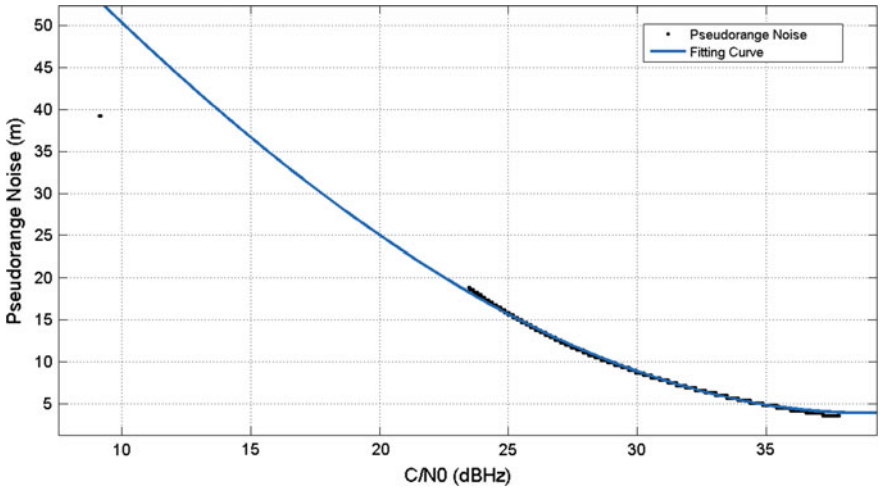


Fig. 5 Fitting curve: pseudorange noise and C/N0

As shown in Figs. 4 and 5, the range noise of Huawei P9 distribute from 5 to 40 m. The mean theoretical value of pseudorange noise is about 10.76 m. Comparing the theoretical pseudorange noise with CNR, it can be approved that the noise value reduces as the CNR values rises.

3.2.2 Pseudorange Noise in Practice

In terms of the pseudorange measurements computed, the polynomial fitting method was used to get the pseudorange noise with each 120 adjacent measurements divided into one segment to compute the fitting parameters [4, 5].

Figure 6 has shown the Huawei P9 pseudorange noise of each satellite. In general, the value decreases when the CNR gets higher. Most of the pseudorange noise values of each satellite were within 8–12 m and no higher than 20 m. The mean noise value of all is 10.10 m, close to the theoretical pseudorange noise while that of Novatel receiver was only 0.023 m, as shown in Fig. 7.

3.3 *Carrier Phase Measurements*

3.3.1 Carrier Phase Noise

As mentioned already, the tracking noise of carrier phase measurements is the main factor of carrier phase range noise. The principles were shown as follows.

$$\sigma^2_{coherent} = \frac{K_3}{C_s/N_0} \tag{3.5}$$

$$\sigma^2_{cost} = \sigma^2_{coherent}\left[1 + \frac{K_4}{C_s/N_0}\right] \tag{3.6}$$

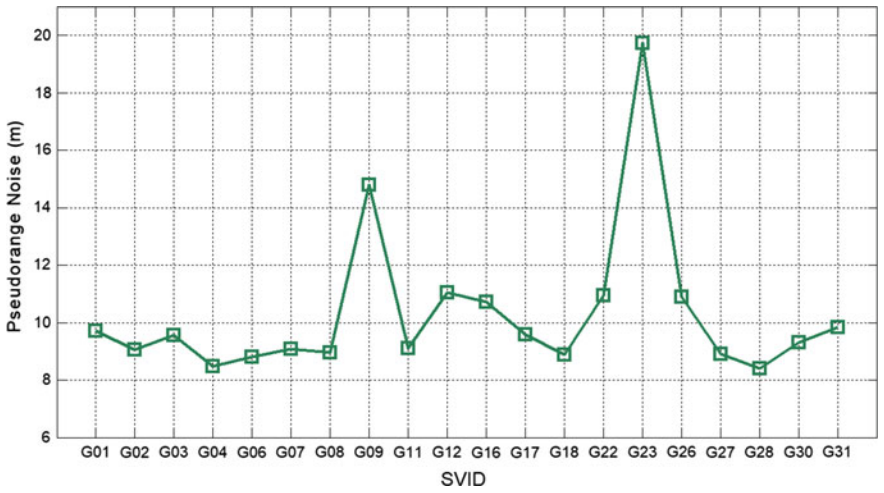


Fig. 6 Pseudorange noise of Huawei P9

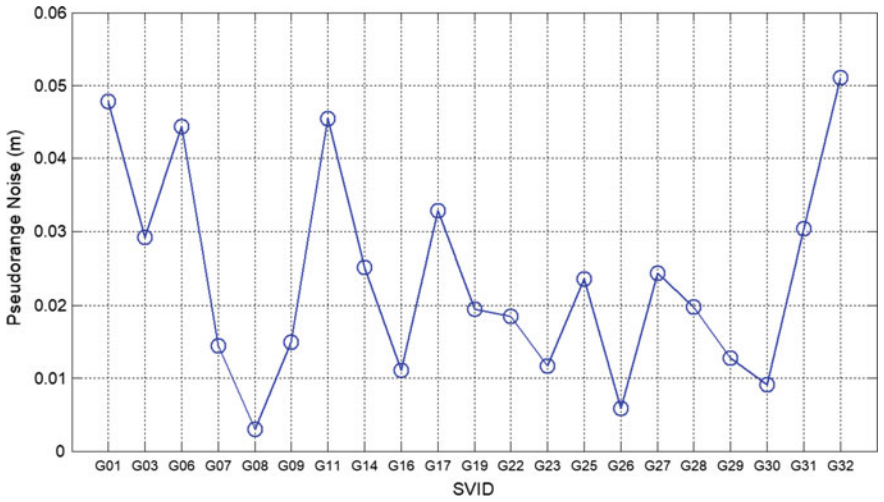


Fig. 7 Pseudorange noise of NovAtel receiver

$$K_3 = \frac{B_\theta(1 - 0.5B_\theta T)}{\int_{-\beta/2}^{\beta_r/2} G_s(f)df} \tag{3.7}$$

$$K_4 = \frac{1}{2T\int_{-\beta_r/2}^{\beta_r/2} G_s(f)df} \tag{3.8}$$

$\sigma^2_{coherent}$ is the coherent delay-locked loop tracking error; σ^2_{cost} is the noncoherent delay-locked loop tracking error; B_θ is loop bandwidth [5]. And also the tracking precision decreases as the CNR gets lower [3].

As is shown in Figs. 8 and 9, when cycle slip occurred, the carrier phase range noise overflows and without cycle slip, the noises mainly distributes within 0.001–0.006 m, reducing as the C/N0 gets higher.

3.3.2 Cycle Slip Detection

Cubic epoch differential method was adopted to detect cycle slip with carrier phase measurement.

$$\Delta\rho(t_i) = \rho(t_i) - \rho(t_{i-1}) \tag{3.9}$$

$$\nabla\Delta\rho(t_i) = \Delta\rho(t_i) - \Delta\rho(t_{i-1}) \tag{3.10}$$

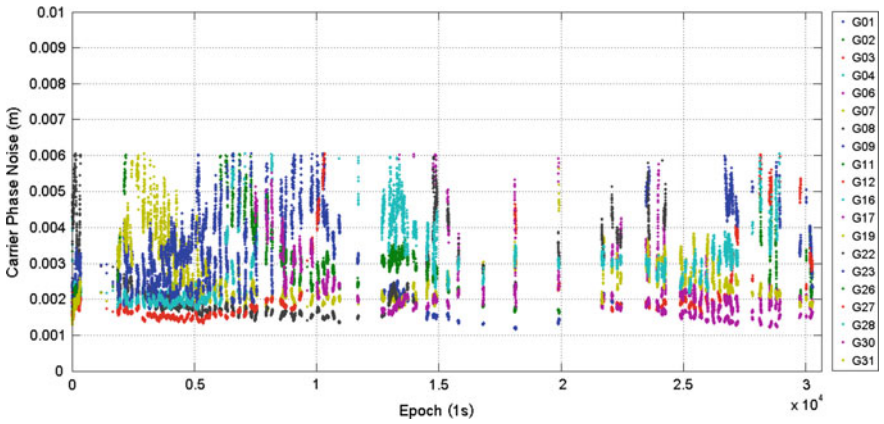


Fig. 8 Distribution of carrier phase noise

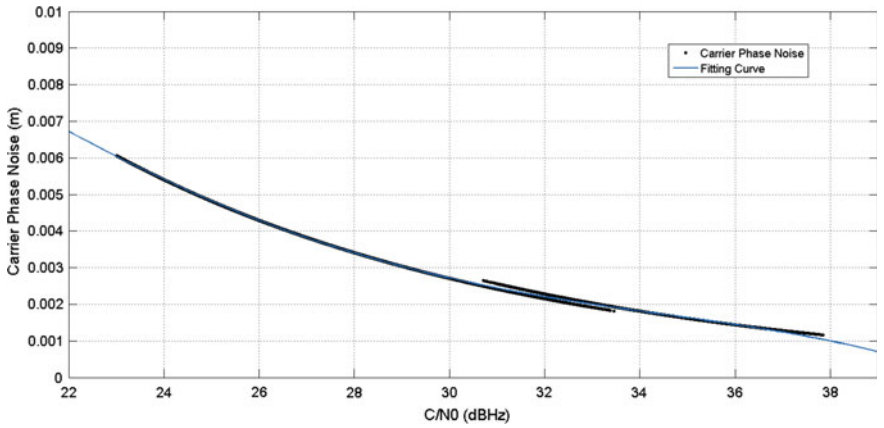


Fig. 9 Fitting curve: carrier phase noise and C/N0

Table 2 Principle of cycle slip detection

Epoch	Phase	Single difference	Double difference	Triple difference
1	0	0	0	0
2	0	0	0	0
3	ε	ε	ε	ε
4	ε	0	$-\varepsilon$	-2ε
5	ε	0	0	ε
6	ε	0	0	0
7	ε	0	0	0

Table 3 Cycle slip of Huawei P9

SVID	Cycle slip	Cycle slip ratio (%)
G01	8053	51.16
G02	1037	38.85
G03	6627	56.85
G04	2942	32.13
G06	5039	58.82
G07	11,060	66.66
G08	9929	58.29
G09	5514	43.71
G11	11,194	66.53
G12	728	43.31
G16	4505	43.97
G17	11,774	69.03
G18	7640	62.00
G22	8678	72.09
G23	3268	33.12
G26	1935	45.81
G27	4657	40.32
G28	12,585	65.69
G30	8710	57.82
G31	1480	39.30

Table 4 Cycle slip of Nexus 9

SVID	Cycle slip	Cycle slip ratio (%)
G03	3	0.84
G07	41	1.27
G08	34	0.88
G09	40	1.03
G11	29	0.98
G16	37	0.95
G23	36	0.93
G26	14	0.58
G27	38	0.98
G30	27	3.09
G31	11	0.44

$$\Delta\nabla\Delta\rho(t_i) = \nabla\Delta\rho(t_i) - \nabla\Delta\rho(t_{i-1}) \tag{3.11}$$

By means of differential method, the cycle slip of carrier phase can be amplified to be more detectable [6].

It was found after analysis that the cycle slip ratio of carrier phase measurements received from Huawei P9 smartphone is overtop. According to statistics, there are 120, 759 measurements with cycle slips in 227,831 carrier phase measurements (Table 2). Compared with Huawei P9, that of NovAtel is much lower with only 28 measurements containing cycle slip. However, not all the Android smart devices (Table 3) have this fault. In similar conditions, using the Nexus 9 tablet with better performance collecting 31,791 measurements during 22:32:21–23:36:59, Aug. 7th, 2017 (GPST), there was only 326 cycle slips detected (Table 4).

4 Conclusion

The report briefly introduced GNSS libraries and raw measurements of Android system. Moreover, it assessed the Huawei P9 GNSS measurements and compared it with that of NovAtel receivers. It was demonstrated that in the same observing environment, the quality of Huawei P9 GNSS measurements still need to be improved.

(1) The mean value of signals CNR from NovAtel GNSS receivers is 44.63 dBHz, much higher than that from Huawei P9, 28.88 dBHz.
(2) The mean pseudorange noise of NovAtel receiver is only 0.023 m, while that of Huawei P9 smartphone reaches above 8 m. Moreover, the cycle slip of carrier phase measurements occurs much too frequent on Huawei P9.
(3) Though the GNSS raw measurements are now available in the output of Android Nougat, the GNSS module inside smartphone still needs to be optimized.

References

1. Guo L (2016) The first line of code, 2nd edn. Post and Telecom Press, Beijing, China, pp 01–04
2. Simon B, Van Frank D (2016) Innovation: precise positioning using raw GPS measurements from Android smartphones. GPS World J 27:43–48
3. Chen W, Li X (2015) Accuracy assessment for carrier phase and code measurements of BDS B1/B3 frequency observation. Aerospace Shanghai J 32:62–66
4. Guan M, Jiao W, Jia X (2017) Accuracy analysis of BDS pseudorange observation. Geom Sci Eng J 37:21–25
5. Zhang X, Wang F (2011) Research and analysis on pseudo-range noise modeling method. Aerospace Controland Appl J 37:44–48
6. Zhao L (2012) The Application of Triple-frequency Data Combination in Cycle Slip Detection and Repairing. Liaoning Technical University

The Study on BDS Dynamic Positioning Performance Assessment

Chenghe Fang, Liang Chen, Changjiang Geng, Zhichao Ma and Qi Mao

Abstract We provided the key technologies for BeiDou navigation satellite system (BDS) standard positioning. Using real-time precise point position (RT-PPP) technology to get the precise reference coordinate, the dynamic positioning performance of BDS in service area was evaluated. Our results indicate that the 95% confidence level error is 2.77 and 4.27 m for horizontal and vertical components with B1I signal, 3.24 and 4.95 m for B2I signal, and 4.11 and 6.84 m for B1I and B2I ionosphere-free combination.

Keywords BDS · Dynamic positioning · Performance

1 Introduction

BDS has been providing regional Positioning, Navigation and Timing (PNT) service since 2012. There have a number of studies on BDS service performance. In signal-in-space performance assessment area, Chen et al. analyzed broadcast ephemeris accuracy in 2013 and results showed that accuracy of the GEO satellite broadcast orbit in radial direction was better than 1.5 m, and accuracy of the IGSO/MEO satellite was better than 1.0 m [1]. Montenbruck et al. evaluated the performance of signal noise ratio, broadcast orbit and clock offset about BDS, results showed that BDS had an extremely high stability of inter-frequency deviation, and the quality of satellite hydrogen clock were comparable with other GNSS. Moreover, the horizontal and vertical accuracy of BDS-only static positioning can reach up to 12 cm [2]. In 2014, Zhang et al. analyzed the

C. Fang (✉) · L. Chen · C. Geng · Z. Ma · Q. Mao
Test and Assessment Research Center, Satellite Navigation Office, 1 Fengyingdong Road, Beijing 100094, China
e-mail: fangchenghefch@163.com; 534871372@qq.com

C. Fang · L. Chen · C. Geng · Z. Ma · Q. Mao
GNSS System Engineering Center, China Academy of Aerospace Electronics Technology, 1 Fengyingdong Road, Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_15

signal-in-space error of BDS, results showed the average Signal-In-Space User Ranging Error (SISURE) of all satellites was 1.45 m [3]. Zhang et al. evaluated the accuracy of the BDS broadcast Klobuchar model, results showed that the correction percentage was generally above 70%, the correction accuracy was about 1.5 m in the northern hemisphere and 3.5 m in southern hemisphere [4]. Montenbruck et al. analyzed the accuracy of BDS broadcast ephemeris in 2013, the results showed that accuracy of BDS broadcast ephemeris was equivalent to the GPS Block IIA satellite and better than GLONASS and Galileo satellites [5]. In 2016, Zhang et al. analyzed BDS broadcast ephemeris biases using the IGS post-processed precise orbit and clock as references, results showed that the bias of BDS broadcast ephemeris can reach up to 2 m [6]. In positioning performance assessment area, Haibo et al. evaluated the dynamic positioning performance of BDS/GPS combination in 2014, results showed that the availability and reliability of the single-frequency BDS/GPS and dual-frequency BDS single-epoch kinematic positioning were comparable to those of the dual-frequency GPS [7]. In 2015, Cong studied the theory and key technologies of high dynamic positioning performance of BDS satellite navigation system [8]. In 2017, Wang et al. evaluated the service performance of BDS in the Antarctic using data collected at Zhongshan station, results showed that the accuracies of BDS-only single-frequency point positioning higher than 22, 9 and 25 m in the north, east and up component [9].

Compared with static positioning, performance of dynamic positioning is more useful, which is mainly limited to terminal performance and complex environment. So far, there is no detailed analysis of the dynamic positioning performance of BDS. In this paper, we evaluated the dynamic positioning performance of BDS in service area based on large-scale dynamic observation data.

2 Algorithm of BDS Standard Positioning

The equation of the BDS pseudo-range observation of 1 satellite-user pair is as follows:

$$P = \rho + c(\delta t_r - \delta t_s) + \delta_{tgd} + \delta_{ion} + \delta_{trop} + \delta_{rel} + M_p + \varepsilon_p \quad (1)$$

where P is pseudo-range, ρ is geometric distance between satellite and receiver; δt_r, δt_s and δ_{tgd} respectively are receiver clock offset, satellite clock offset and Time Group Delay (TGD); δ_{ion}, δ_{trop} and δ_{rel} respectively are ionospheric delay, tropospheric delay and relativistic effects; M_p and ε_P respectively are multipath effect and pseudo-range observation noise.

In the positioning solution, the needed parameters and above errors are corrected using methods as follows:

(1) The algorithm and specified constants for computing satellite orbit and the satellite clock offset using broadcast ephemeris according BDS-SIS-ICD [10].

(2) The algorithm and specified constants for computing ionospheric delay using Klobuchar model parameters from navigation message BDS-SIS-ICD [10].
(3) The TGD is corrected by using TGD parameters of broadcast ephemeris [10].
(4) The relativistic effects is corrected by using the correction formula $\Delta t_r = F \cdot e \cdot \sqrt{A} \cdot \sin E_k$ [10].
(5) Using Saastamoinen model to correct tropospheric delay.

Note the differences between BDS and GPS standard positioning: first, there are differences in using of BDS and GPS Klobuchar-8 model. The BDS broadcast Klobuchar parameters are hourly updated and both algorithm and parameters are different from GPS. Second, GPS has only one TGD parameter which is L1 C/A frequency relative to the ionosphere combination frequency while there are 2 TGD parameters of BDS, represents B1I and B2I signal relative to the B3 signal. Third, the earth's gravitational constant and rotation rate defined in BDS interface control document are different from GPS [10]. It would lead to additional error if using GPS defined ones.

3 Reference of Dynamic Positioning Performance Assessment

Getting the precise coordinate of the mobile testing platform is an important work of the dynamic positioning performance test. For example, Jing et al. used coordinates of GPS-only PPP as reference, and only the ambiguities fixed epochs are used [11]. Mao et al. used RTK results as reference for vehicle device to be tested [12]. Both of these works are implemented in a small-scale region, but not suite for large-scale dynamic testing. Therefore, the more mature RT-PPP technology, which can achieve decimetre to centimetre-level positioning accuracy in real-time after convergence. And it has globally uniform positioning accuracy [13–15], which is suit for reference in large-scale dynamic testing in this paper.

The RT-PPP solution provider broadcast precise orbit and clock offset correction to user terminal. The user terminal utilizes the phase observation combined with PPP technique to achieve centimetre-level accuracy once converged.

Figure 1 shows the results provided by the RT-PPP receiver. From the figure we could see that the horizontal accuracy is about 10 cm after 15 min and the vertical accuracy could be converged to 10 cm in about 25 min. The accuracy is about 2–5 cm after totally convergence. That would be fit for using as the reference of the meter lever standard positioning.

Due to the existence of receiver clock offset, the dynamic receiver and the reference receiver are not actually synchronized. In this paper, the receiver clock offset is tested to be in nanoseconds and milliseconds. And the average speed of the dynamic test is about 100 km/h, then the error of distance caused by the 1 ms time synchronization error is about 2.8 cm, which can be neglected in the meter lever standard positioning.

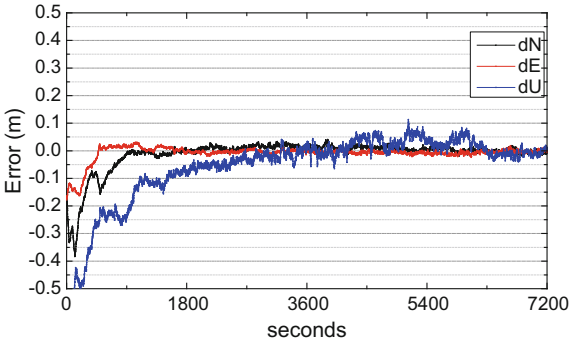


Fig. 1 Positioning result of RT-PPP receiver

4 Dynamic Positioning Service Performance Test

4.1 Number of Visible BDS Satellites and PDOP Along the Testing Route

Figure 2 shows the dynamic testing route during the period from August 7, 2017 to August 16th from Xi'an city, Shaanxi Province to Horgos Port of Xinjiang Province, mainly driving along the highway. The total length is about 3462 km.

Table 1 lists the number of BDS visible satellites and PDOP value along the testing route. From the table we could see that during the test period, the mean value of BDS visible satellites number are in the range of 8–11. And the mean value

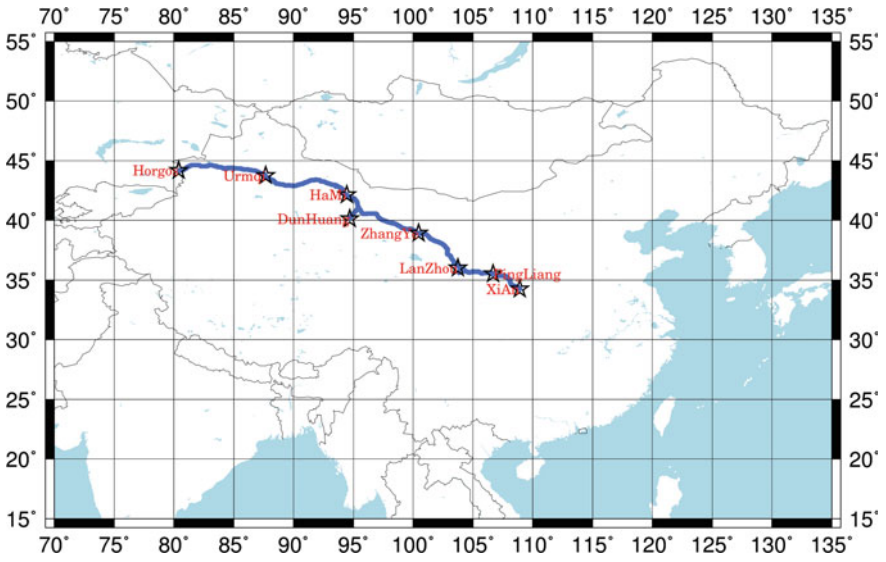


Fig. 2 Route of dynamic testing in China

Table 1 Number of visible satellites and PDOP along the testing route

Results	Doy							
	219	220	221	222	224	225	226	7-days
Mean (SatNum)	9	10	11	10	8	8	8	9
Mean (PDOP)	2.78	2.77	2.51	3.00	3.62	3.73	4.07	3.22
95% (PDOP)	4.45	5.77	3.45	4.70	7.78	5.26	6.68	5.69

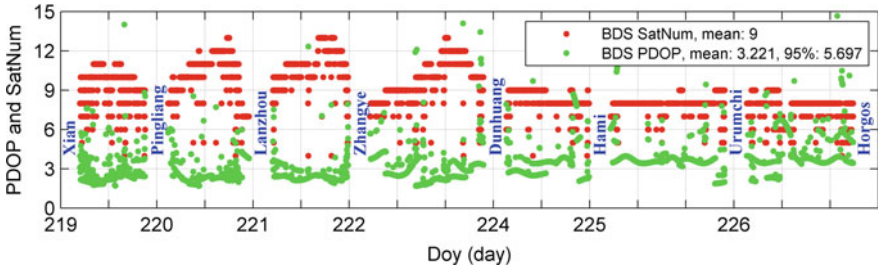


Fig. 3 Number of visible satellites and PDOP along the testing route

of PDOP are in the range of 2.51–4.07, and the 95% confidence level value are in the range of 3.45–4.07. Figure 3 presents the number of BDS visible satellites and PDOP along the testing route, from the figure we could see that with the advance of the test to the northwest border, the number of BDS visible satellites will be reduced, the corresponding PDOP value becomes larger. And in the entire testing, the average value of BDS visible satellites number is 9, and the mean value of PDOP is 3.22, the 95% confidence level value is 5.70.

4.2 Data Processing Strategy

Table 2 shows the error corrected model and processing strategy of the dynamic data processing, in which only BDS pseudo-range observation used. Specifically, the satellite orbit, satellite clock offset is calculated using broadcast ephemeris, as well as the other corrections considered are as follows: earth rotation correction, TGD correction, ionospheric delay correction, troposphere delay correction, and relativistic effects correction.

In addition, there is a convergence problem at the beginning of the measurement and over the tunnel for the RT-PPP receiver which is used as reference receiver. In order to reduce the influence, the reference receiver should be turned on for convergence before the normal measurement. After crossing tunnel, jungle and city, where navigation signals are blocked seriously, we would wait for the reference RT-PPP receiver to convergence and then start the measurement. Because the entire

Table 2 Error correction model and data processing strategy in data processing

Error Correction	B1I	B2I	B1I&B2I
Orbit	BDS broadcast ephemeris		
Satellite clock			
TGD correction	TGD parameters of broadcast ephemeris		
Ionosphere	BDS Klobuchar8		Ionosphere-Free Combination
Troposphere	Saastamoinen Model		
Relativistic effect	Model correction		
Earth Rotation	Model correction		

test route is along the highway in west of China, there is almost no other signal obstructions besides the bridge and tunnel. Therefore, the reference RT-PPP receiver and observation data of testing receivers will have a certain continuity and stability.

4.3 Analysis of BDS Dynamic Positioning Performance

Figure 4 shows horizontal accuracy of the BDS B1I positioning along the testing route, and from it we could see that the horizontal accuracy is within about 3 m, and most time it is in the range of 0.5–2 m. Figure 5 shows vertical accuracy of the BDS B1I positioning along the testing route, and from it we could see that the vertical accuracy is within 5 m, and most of the time, it is in the range of 1–3 m.

Table 3 lists the statistical result of horizontal and vertical accuracy of the BDS B1I positioning along the testing route. We could see that during the test period, the RMS of horizontal accuracy of BDS B1I positioning are in the range of 1.07–2.11 m, and the 95% confidence level value are in the range of 1.68–3.13 m. The RMS of vertical accuracy of BDS B1I positioning are in the range of 1.51–2.79 m, and the 95% confidence level value are in the range of 2.82–5.16 m. Figure 6 presents the results of BDS B1I positioning along the testing route, we could see that the positioning of BDS B1I has certain continuity and stability along the testing line. And in the entire test, the RMS of horizontal accuracy is 1.76 m, the 95% confidence level value is 2.77 m, and the RMS of vertical accuracy is 2.47 m, the 95% confidence level value is 4.27 m.

Table 4 lists the statistical results of horizontal and vertical accuracy of the BDS B2I positioning along the testing route. We could see that during the test period, the RMS of horizontal accuracy of BDS B2I positioning are in the range of 1.08–2.14 m, and the 95% confidence level value are in the range of 1.82–3.37 m. The RMS of vertical accuracy of BDS B2I positioning are in the range of 2.02–3.05 m, and the 95% confidence level value are in the range of 3.36–5.49 m. Figure 7 presents the results of BDS B2I positioning along the testing route, we

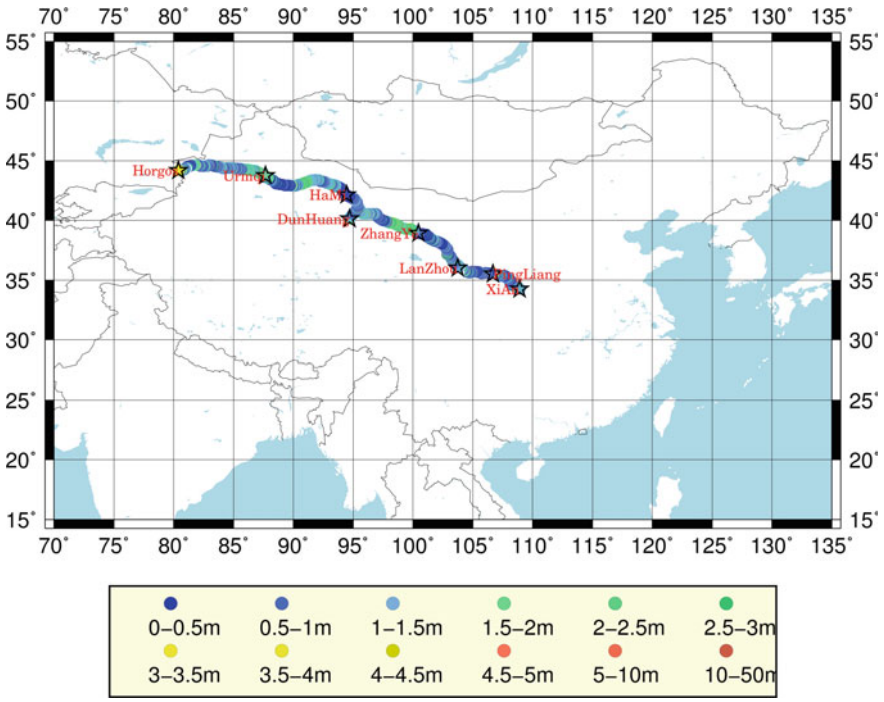


Fig. 4 Horizontal accuracy of the BDS B1I positioning along the testing route

could see that the positioning results of BDS B2I has certain continuity and stability along the testing line. Moreover, in the entire test, the RMS of horizontal accuracy is 1.89 m, the 95% confidence level value is 3.24 m, and the RMS of vertical accuracy is 2.74 m, the 95% confidence level value is 4.95 m.

Table 5 lists the statistical results of horizontal and vertical accuracy of the BDS B1I&B2I positioning along the testing route. We could see that during the test period, the RMS of horizontal accuracy of BDS B1I&B2I positioning are in the range of 1.52–3.24 m, and the 95% confidence level value are in the range of 2.48–4.50 m. The RMS of vertical accuracy of BDS B1I&B2I positioning are in the range of 2.27–3.64 m, and the 95% confidence level value are in the range of 4.29–7.63 m. Figure 8 presents the results of BDS B1I&B2I positioning along the testing route, we could see that the positioning results of BDS B1I&B2I has certain continuity and stability along the testing route. Moreover, in the entire test, the RMS of horizontal accuracy is 2.27 m, the 95% confidence level value is 4.11 m, and the RMS of vertical accuracy is 3.39 m, the 95% confidence level value is 6.84 m. We can see that the dual frequency is slightly worse than the single frequency results, it may be because the dual frequency could reduce ionospheric delay, but the amplify the noise which would be more than 3 times larger than the original one.

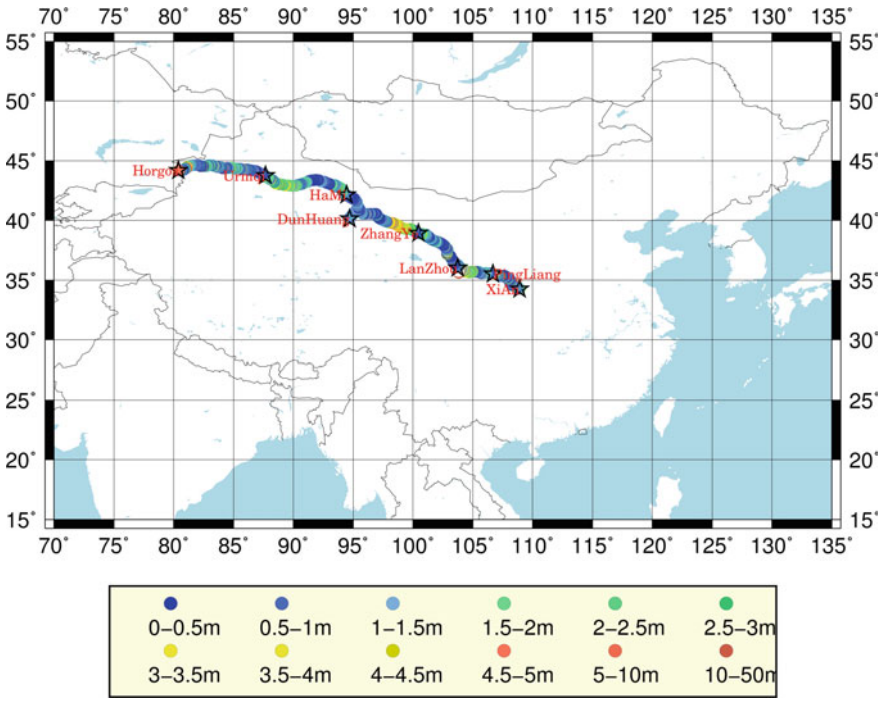


Fig. 5 Vertical accuracy of the BDS B1I standard positioning along the testing route

Table 3 Statistical result of horizontal and vertical accuracy of the BDS B1I positioning

Results (m)	Doy							
	219	220	221	222	224	225	226	7-days
RMS (dH)	2.11	1.31	1.07	1.98	1.77	1.75	1.58	1.76
RMS (dV)	2.41	2.12	1.51	2.79	1.95	1.92	1.91	2.47
95% (dH)	2.44	2.20	1.68	3.13	2.28	2.72	2.69	2.77
95% (dV)	4.92	4.00	2.82	5.16	3.23	3.47	3.99	4.27

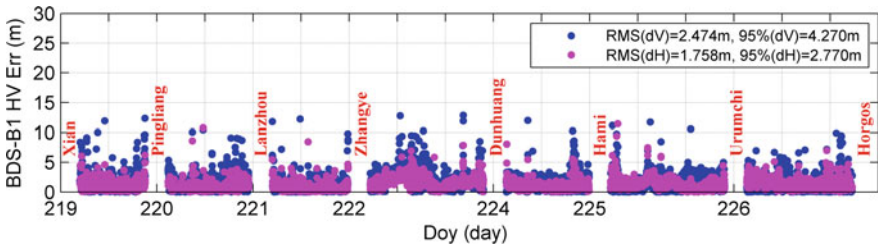


Fig. 6 Results of BDS B1I positioning along the testing route

Table 4 Statistical results of horizontal and vertical accuracy of the BDS B2I positioning

Results (m)	Doy							
	219	220	221	222	224	225	226	7-days
RMS (dH)	1.70	1.42	1.08	2.14	1.76	2.06	1.79	1.89
RMS (dV)	2.23	2.86	2.02	3.05	2.63	2.49	2.14	2.74
95% (dH)	2.52	2.00	1.82	3.33	3.32	3.37	3.34	3.24
95% (dV)	4.41	5.49	3.36	5.46	5.11	4.32	4.33	4.95

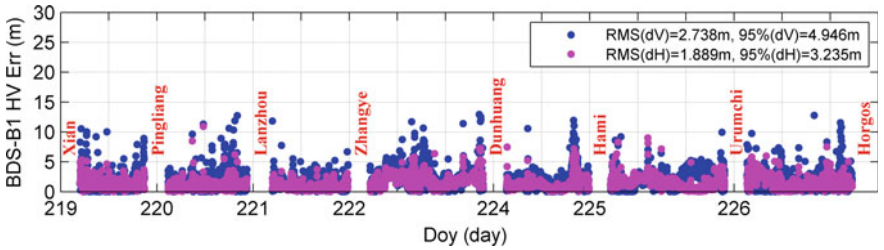


Fig. 7 Results of BDS B2I positioning along the testing route

Table 5 Statistical results of horizontal and vertical accuracy of the BDS B1I&B2I positioning

Results (m)	Doy							
	219	220	221	222	224	225	226	7-days
RMS (dH)	3.24	2.05	1.52	2.62	2.02	2.44	2.32	2.27
RMS (dV)	3.15	3.30	2.27	3.64	2.96	2.67	3.05	3.39
95% (dH)	3.07	3.65	2.48	4.50	3.85	3.78	4.20	4.11
95% (dV)	6.33	7.20	4.29	7.63	6.14	5.56	6.40	6.84

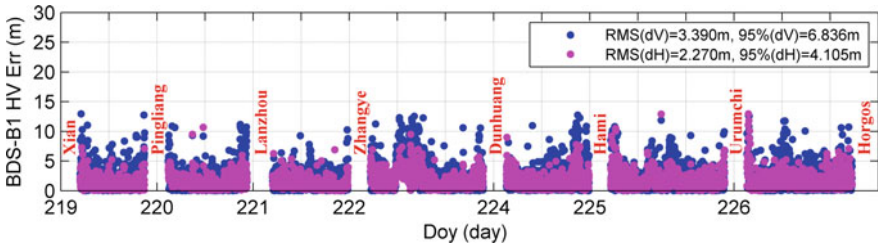


Fig. 8 Results of BDS B1I&B2I positioning along the testing route

5 Conclusions

From the dynamic positioning performance test, we could see that the positioning results of BDS B1I, B2I and B1I&B2I are stable and reliable. The positioning accuracy of BDS B1I is 2.77 (95%) and 4.27 m (95%) for horizontal and vertical components. The positioning accuracy of BDS B2I is 3.24 (95%) and 4.95 m (95%) for horizontal and vertical components. The positioning accuracy of BDS B1&B2I ionosphere-free combination is 4.11 (95%) and 6.84 m (95%) for the vertical and horizontal components.

References

1. Chen L, Jiao W, Huang X et al (2013) Study on signal-in-space errors calculation method and statistical characterization of BeiDou navigation satellite system. Springer, Berlin, pp 423–434
2. Montenbruck O, Hauschild A, Steigenberger P et al (2013) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. GPS Solutions 17(2):211–222
3. Zhang Q, Sui L, Jia X et al (2014) SIS error statistical analysis of BeiDou satellite navigation system. Geomatics Inform Sci Wuhan Univ 39(3):271–274
4. Zhang Q, Zhao Q, Zhang P et al (2014) Evaluation on the precision of Klobuchar model for BeiDou navigation satellite system. Geom Inform Sci Wuhan Univ 39(2):142–146
5. Montenbruck O, Steigenberger P, Hauschild A (2015) Broadcast versus precise ephemerides: a multi-GNSS perspective. GPS Solutions 19(2):321–333
6. Zhang Y, Chen J, Zhou J et al (2016) Analysis and application of BDS broadcast ephemeris bias. Acta Geod Cartogr Sin 45(S2):64–71
7. Haibo He, Jinlong Li, Yuanxi Yang et al (2014) Performance assessment of single- and dual-frequency BeiDou/GPS single-epoch kinematic positioning. GPS Solutions 18(3):393–403
8. Cong D (2015) Research on the theory and key technologies of BDS in high kinematic positioning accuracy calibration. Acta Geodaetica Cartogr Sin 44(12):1402
9. Wang Z, Du Y, Xiong Y, et al (2017) Preliminary assessment of positioning performance of BeiDou navigation satellite system at Zhongshan station, Antarctica. Geom Inform Sci Wuhan Univ 42(8):1027–1034
10. BDS-SIS-ICD (2016) Beidou navigation satellite system signal in space interface control document open service signal B1I (Version 2.1). China Satellite Navigation Office, November, 2016
11. Jing Y, Zeng A (2015) Test on performance of BDS/GPS fusion kinematic positioning. China Satell Navig Conf 34(6):114–118
12. Mao Q (2016) Research on real-time dynamic detection technology of vehicle GNSS positioning. Southeast University, Nanjing, China
13. Dixon K (2006) StarFire™: a global SBAS for sub-decimeter precise point positioning. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, pp 2286–2296

14. Zhang Y (2007) StarFire global high-precision wide-area differential GPS system. Equip Geophys Prospect 15(2):321–333
15. Dai L, Hatch RR (2011) Integrated StarFire GPS with GLONASS for real-time precise navigation and positioning. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, pp. 1476–1485

Research and Analysis of S Curve Bias Evaluation Method in GNSS Signals

Hai Sha, Gangqiang Guan, Hongbing Wang and Henglin Chu

Abstract When a MBOC modulated signal is evaluated by the traditional S-curve bias evaluation method, which exists some problems, such as the selection of correlator spacing is not reasonable and the calculation error of the locking point deviation at the platform is relatively large. An improved S-curve bias evaluation method is proposed. In view of the lack of the study on S-curve characteristics in BeiDou, GPS and Galileo on-orbit satellites, the satellite signals in different systems were collected by 13 m antenna signal quality monitoring system, and in accordance with the modulation type, the calculated results on S-curve bias were analyzed. The comparison results show that the S-curve bias of GPS satellites is obviously smaller than that of Beidou test satellite for BPSK signals. For MBOC signals, the difference of implementation methods greatly affects the calculation results. For AltBOC signals, the precision of Beidou satellite is quite the same as that of Galileo satellite.

Keywords Satellite navigation system · Signal quality assessment
Correlation function · S curve bias

1 Introduction

The satellite navigation signal quality assessment plays an important role in GPS SV19 satellite, Galileo GIOVE satellite and other satellite troubleshooting, the signal quality assessment has gradually attracted people's attention. Satellite navigation signal quality assessment includes [1] power spectral density envelopes, constellations, eye diagrams, and some parameters related to the correlation function characteristics such as correlation loss, S-Curve Bias (SCB), pseudo-code coherence and so on. Through signal quality evaluation, the performance of satellite signals and signal distortion can be mastered at different test stages such as ground

H. Sha (✉) · G. Guan · H. Wang · H. Chu
Beijing Satellite Navigation Center, Beijing 100094, China
e-mail: sandhai@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_16

test and on-orbit test. Secondly, when there is a satellite fault, multi-dimension observation of signal quality assessment can help to locate the fault. Finally, by establishing a database of long-term satellite signal quality assessment results, it is easy to identify potential problems and risks of the on-orbit satellites. Therefore, carrying out the satellite navigation signal quality assessment work is of great significance and research value to guarantee the high quality service of Beidou Navigation System and to carry out exchanges and cooperation in the international field.

At present, the signal quality evaluation results on many different satellite navigation systems are published both at home and abroad, including Beidou Experimental Satellite [2–5], Galileo Satellite [4, 6], GLONASS Satellite [7] and GPS Satellite [4, 8, 9]. Among them, the German Aerospace Center [4–8] (Deutsches Zentrum fur Luftund Raumfahrt, DLR) used its 30 m antenna to conduct signal quality assessment of Beidou, Galileo, GPS and GLONASS satellites respectively. MITER [9] used the 45 m antenna of Stanford University to focus on the L5 signal of GPS satellites. In China, however, National University of Defense Technology [2, 10] conducted a preliminary assessment of the signal quality of the Beidou II test satellite using its 7.3 m antenna. The Chinese Academy of Sciences National Time Service Center [3, 11] also used its 40 m antenna to evaluate Beidou satellite. However, most of the analysis results given in these literatures are general indicators in communication systems such as power spectral density, constellation diagrams and eye diagrams. Relative parameters such as correlation loss and S-curve deviation that are of concern to the satellite navigation system are comparatively analyzed less. Among them, the correlation loss analysis may be limited to the acquisition of the authorization code, but the the S curve bias can be obtained based on the acquired satellite signal with high precision processing. At the same time, the traditional SCB evaluation method still exists some drawback for the modulated signal of the multiplexed binary offset carrier (MBOC) jointly proposed by the U.S. and the EU in 2006 and used for international cooperation at the L1/E1 frequency point.

Therefore, based on the discussion of the traditional SCB evaluation method, this paper proposes an improved SCB evaluation method for the lack of evaluating MBOC modulation signal. Then the SCB of GPS, Galileo and Beidou test satellite at each frequency is analyzed by using 13 m antenna signal quality monitoring system, and the SCB of the signal is further compared according to the signal modulation type.

2 Traditional SCB Assessment Method

S-curve usually refers to the navigation receiver delay lock loop phase detection curve. Taking the non-coherent delay locked loop leading/lagging power-type phase detector as an example, the S-curve is obtained by subtracting the square of the output of the correlator before and after the lag, which can be expressed as:

$$\mathrm{SCurve}(\varepsilon,\delta)=\left|CCF\left(\varepsilon-\frac{\delta}{2}\right)\right|^2-\left|CCF\left(\varepsilon+\frac{\delta}{2}\right)\right|^2 \tag{1}$$

where, δ is the correlator lead-lag spacing, and CCF (ε) represents the normalized correlation function of the actual signal with the local reference code.

Ideally, the zero point of the S-curve in the linear region, which is known as the Lock Point, should lie at δ at zero. However, in fact, due to the influence of channel transmission distortion and multipath, the lock point of S curve generally has some deviation. The lock point deviation $\varepsilon_{\mathrm{bias}}$ (δ) is defined as [1]:

$$\mathrm{SCurve}(\varepsilon_{bias}(\delta),\delta)=0 \tag{2}$$

From this, we can get SCB as:

$$SCB=\max_{overall\ \delta}(\varepsilon_{bias}(\delta))-\min_{overall\ \delta}(\varepsilon_{bias}(\delta)) \tag{3}$$

Taking into account the range of δ is [0, $\delta_{\max}$], usually $\delta_{\max}$ values are as follows:

$$\delta_{\max}[chips]=\begin{cases}2 & BPSK(n)\\ \frac{n}{m} & BOC(m,n)\end{cases} \tag{4}$$

Note that δ is usually defined as the correlator spacing in engineering. At this time, δ is exactly half of the early-late spacing of the correlator. Therefore, $\delta_{\max}$ should also be half of (4). In addition, if there are multiple zero crossings (such as BOC signals) on the S curve, the point closest to the peak of the correlation power is selected as the lock point. At the same time, in order to reduce the influence of cross-correlation between different signal components, it is also necessary to calculate the SCB after averaging a plurality of cycle lock-point curves.

3 SCB Evaluation Method on MBOC Signal

The MBOC modulated signal is a signal defined by the power spectrum function. Currently, different modulation schemes are proposed by GPS, Galileo and BeiDou system respectively. Among them, the L1C signal transmitted by the GPS III satellite will use Time-Multiplexed BOC(TMBOC) method. Galileo's E1 OS signals already use Composite BOC (CBOC) mode. For the BeiDou system, in order to achieve compatibility with GPS and Galileo, a Quadrature Multiplexed BOC (QMBOC) mode may be used at the B1 frequency point in the future.

Figure 1 shows the different modulation signal correlation function waveform. Compared with the BOC (1,1) signal, the correlation function of the MBOC (6,1,1/11) signal is not exactly the same in the first zero-crossing point, but there

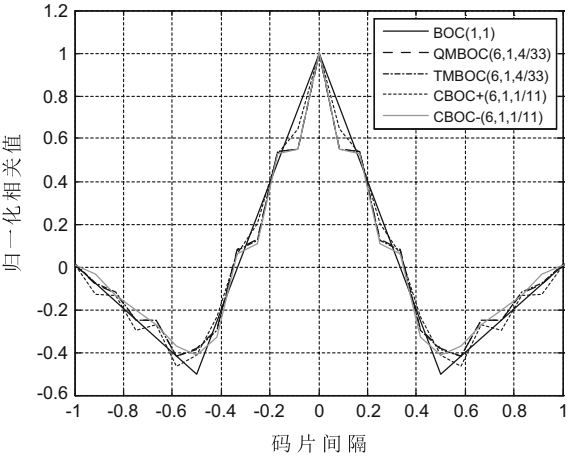


Fig. 1 Different MBOC modulation signal correlation function diagram

are two platform segments with a slight slope, The range of the first platform is 1/12–1/6 chip interval, and the second platform is in the range of 1/4–1/3 chip interval.

According to the traditional SCB evaluation method, the range of the early-late distance δ for the MBOC signal correlator is obviously too small if it is selected according to the BOC (6,1) signal. However, according to the BOC (1,1) signal, the correlation function in the platform segment is easily affected by the channel characteristics. This not only makes the problem of multiple locking points difficult to confirm, but also makes the calculated SCB to be obviously larger. For this reason, the following improvements should be made when evaluating the SCB of MBOC signals:

1. The early-late interval δ range of the correlator should only select the linear part of the second platform, that is

$$\delta[chips] = \begin{cases} [0,2] & BPSK(n) \\ [0,\frac{n}{m}] & BOC(m,n) \\ [0,\frac{1}{6}],[\frac{1}{3},\frac{1}{2}] & MBOC(6,1,1/11) \end{cases} \tag{5}$$

2. In order to avoid sinking the lock point into the platform during tracking, it is necessary to ensure that there is only one zero crossing of the S curve in the range of [0, 1/2].

4 Signal Quality Monitoring System

Signal quality monitoring system consists of 13 m antenna, downlink receiving subsystem, baseband processing subsystem, analysis and evaluation subsystem, time frequency subsystem, monitoring subsystem, as shown in Fig. 2.

The downlink receiving subsystem is used to receive the navigation signal at any frequency in the L-band and output the intermediate frequency to meet the needs of the back-end A/D acquisition. The analysis and evaluation subsystem is used to analyze and evaluate the quality of the navigation signals and achieve the absolute calibration of the receiving link gain and the relative calibration of the navigation signal delay at different frequencies.

5 SCB Test Results and Comparison

In order to analyze and compare the SCB characteristics of GNSS system signals, three satellite signals of BeiDou M2-S, GPS 10 (IIF satellite) and GALILEO 8 were collected respectively on November 23, 2016, with a signal sampling rate of 250 MHz, an effective quantization bit number of 12 Bit. Firstly all the satellite navigation system signal characteristics are analysis, and then according to the different types of signal modulation on the SCB calculation results are discussed.

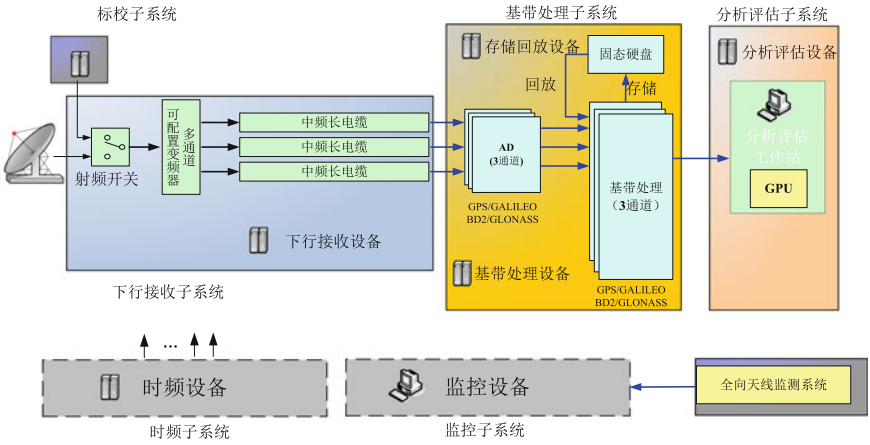


Fig. 2 Signal quality monitoring system structure diagram

5.1 GNSS Signal Analysis

After realizing the GPS modernization plan, GPS satellites will transmit 4 civilian signals at the same time, L1 C/A, L1C, L2C and L5C respectively. Galileo systems provide open-service satellite signals including E1-B, E1-C, E5a-I, E5a-Q, E5b-I and E5b-Q. China BeiDou System is undergoing global system construction. In 2016, the launch of new global test satellites including TMBOC (6,1,1/11), AltBOC (15,10) and other compatible cooperative signals. Civil satellite navigation system of civil signal parameters as shown in Table 1.

As can be seen from Table 1, in order to realize the compatible interoperability between GNSS systems, each satellite navigation system adopts the more complicated BOC modulation technology, and its requirement for signal quality will also be higher.

Table 1 GNSS signal characteristic parameters

	Frequency band	Frequency/MHz	Signal component	Modulation
GPS	L1	1575.42	L1 C/A	BPSK-R(1)
			L1CD	BOC(1,1)
			L1CP	TMBOC(6,1,4/33)
	L2	1227.60	L2CM	BPSK-R(1) TDMA
			L2CL	
	L5	1176.45	L5CD	QPSK-R(10)
			L5CP	
Galileo	E1	1575.42	E1B	CBOC(6,1,1/11,'+')
			E1C	CBOC(6,1,1/11,'−')
	E5	1176.45	E5a-I	AltBOC(15,10)
			E5a-Q	
		1207.14	E5b-I	
			E5b-Q	
BDS Test Satellite [2]	B1	1561.10	B1I	BPSK-R(2)
		1575.42	B1CD	BOC(1,1)
			B1CP	TMBOC(6,1,4/33)
	B2	1176.45	B2aD	TDAltBOC(15,10)
			B2aP	
		1207.14	B2bD	
			B2bP	
	B3	1268.52	B3I	BPSK-R(10)

5.2 SCB Calculation Results

It can be seen from Sect. 5.1 that GNSS signals adopt a variety of modulation schemes and their impact on the SCB curve will also be different. In order to analyze easily, the SCB calculation result of each signal is discussed below according to the modulation type. Among them the locking point deviation of each signal is shown as in Fig. 3.

1. BPSK-R signal

BPSK-R signals include GPS satellites L1C/A, L5CD, L5CP and Beidou test star B1I, B3I signals, SCB calculation results shown in Table 2. It can be seen that the SCB of the GPS10 satellite is significantly smaller than that of the BD M2-S, especially the SCB of the L1 C/A signal is only 0.05 ns.

2. MBOC modulation signal

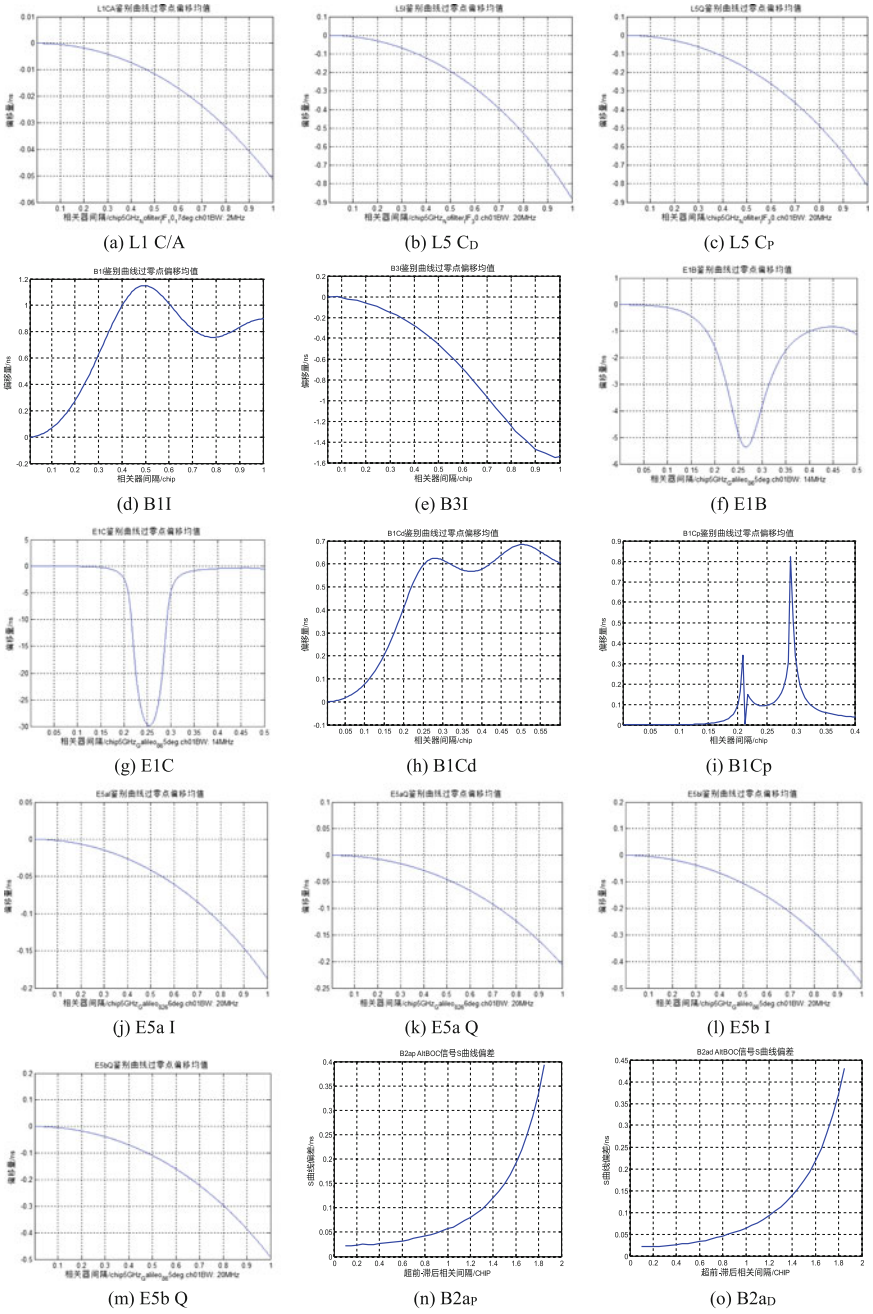
At present, MBOC signals that have been broadcasted include E1B and E1C signals of Galileo satellites and $B1C_D$ and $B1C_P$ signals of Beidou test satellite, and future GPS-III satellites will also transmit L1C signals.

Table 3 shows the SCB calculation results for MBOC signals of Galileo 8 and BeiDou M2-S satellite. It can be seen that the SCB calculation results for E1B and E1C signals are relatively consistent while the SCB differences for $B1C_D$ and $B1C_P$ are quite different. This is in contrast to the two MBOC different realization ways. Among them, E1B and E1C signal power are equal, all need to transmit message information, the two SCB results are relatively close. The $B1C_P$ signal not only has stronger signal power than the $B1C_D$ signal, but also does not need to transmit the message information. The SCB result is obviously better than other signals. Overall, the two modulation methods have their own advantages and disadvantages.

At the same time, the curve of E1B and E1C signal locking point in Fig. 3(f, g) shows that the error of CBOC signal is larger at platform, reaching 5 and 30 ns respectively. It indicates that the calculation of the locking point of the MBOC signal at the platform is very sensitive to the channel characteristics.

3. AltBOC modulation signal

Alternate Binary Offset Carrier (AltBOC) signal is a kind of BOC modulated signal, which not only separates the signal spectrum on both sides of the center zero frequency, but also makes the main lobe on both sides correspond to different pseudo code signals. Galileo satellite E5 frequency is using AltBOC modulation, and Compass experimental satellite is using a new alternative TD-AltBOC modulation.



(a) L1 C/A (b) L5 C_D (c) L5 C_P

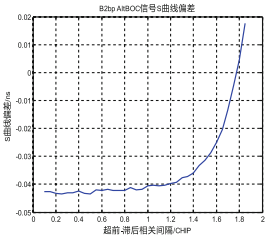
(d) B1I (e) B3I (f) E1B

(g) E1C (h) B1Cd (i) B1Cp

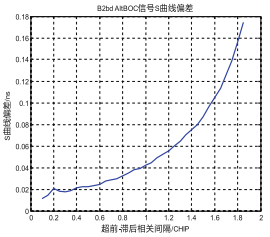
(j) E5a I (k) E5a Q (l) E5b I

(m) E5b Q (n) B2a$_P$ (o) B2a$_D$

Fig. 3 The mean value of the lock point curve in all signals



(p) $B2b_P$



(q) $B2b_D$

Fig. 3 (continued)

Table 2 SCB results of BPSK-R modulation signal

	Signal component	Modulation	SCB(ns)
GPS 10	L1 C/A	BPSK-R(1)	0.05
	L5CD	QPSK-R(10)	0.9
	L5CP		0.8
BD M2-S	B1I	BPSK-R(2)	1.16
	B3I	BPSK-R(10)	1.58

Table 3 SCB results of MBOC modulation signal

	Signal component	Modulation	SCB (ns)
Galileo 8	E1B	CBOC(6,1,1/11,'+')	0.35
	E1C	CBOC(6,1,1/11,'−')	0.4
BD M2-S	B1CD	BOC(1,1)	0.69
	B1CP	TMBOC(6,1)	0.02

Table 4 SCB results of AltBOC modulation signal

	Signal component	Modulation	SCB (ns)
Galileo 8	E5a-I	AltBOC(15,10)	0.18
	E5a-Q		0.2
	E5b-I		0.48
	E5b-Q		0.49
BD M2-S	B2aP	TD-AltBOC (15,10)	0.37
	B2aD		0.42
	B2bP		0.06
	B2bD		0.17

Table 4 shows the SCB calculation results of the Galileo 8 satellite E5 signal and the BeiDou M2-S satellite B2 signal. It can be observed that SCB of the one sideband outperforms the other sidebands in two modulated signals. The Galileo 8 satellite is superior to the lower sideband for the upper sideband and the lower sideband for the Beidou M2-S satellite is superior to the upper sideband. Overall, the accuracy of the two SCBs is comparable.

6 Conclusion

In order to overcome the shortcomings of the traditional SCB evaluation method in processing MBOC signals, an improved method is proposed from two aspects: the range of correlator spacing and the evaluation method at the platform. At the same time, the SCB results of Beidou, GPS and Galileo orbiting satellites are compared and calculated. Through analysis we can draw:

(1) For the BPSK-R signal, the L1C/A, $L5C_D$ and $L5C_P$ signals SCB of the GPS 10 satellite are better than the B1I and B3I signals of the BeiDou M2-S satellite.
(2) For the MBOC modulated signal, the $B1C_P$ signal SCB of the Beidou M2-S satellite is obviously superior to the $B1C_D$ and the E1B and E1C signals of Galileo 8 satellites, which explains for MBOC modulated signal, if there is no message data greater impact on SCB.
(3) For the AltBOC modulated signal, the SCB results of the Galileo 8 satellite E5 signal and the BeiDou M2-S satellite B2 signal all appear inconsistent accuracy of the upper and lower sidebands. However, the accuracy of the two is quite good.

Acknowledgements Thanks to Liu Yuqi, China Electronics Technology Group 29, Associate Professor, Yao Zheng and Dr. Zhang Jiayi from Tsinghua University for their assistance in SCB calculation and MBOC modulation signal evaluation.

References

1. Soellner M, Kurzhals C, Hechenblaikner G et al (2008) GNSS offline signal quality assessment. In: The 21th international technical meeting of the satellite division, Savannah, GA
2. Xiao W, Liu W, Sun G (2016) Modernization milestone: BeiDou M2-S initial signal analysis. GPS Solution. https://doi.org/10.1007/s10291-015-0496-7
3. He C, Guo J, Lu X et al (2013) Offline analysis of Beidou MEO-3 signal quality. In: The 26th international technical meeting of the satellite division, Nashville, TN
4. Thoelert S, Furthner J, Meurer M (2012) New birds in the sky—signal in space (SIS) analysis of new GNSS satellites. In: The 25th international technical meeting of the satellite division, Nashville, TN

5. Thoelert S, Erker S, Meurer M et al (2010) New signals in the sky—a high gain antenna analysis of GPS IIF and COMPASS. In: The 23th international technical meeting of the satellite division, Portland, OR
6. Montenbruck O, Gunther C, Graf S et al (2006) GIOVE-a initial signal analysis. GPS Solution. https://doi.org/10.1007/s10291-006-0027-7
7. Thoelert S, Erker S, Furthner J et al (2011) First signal in space analysis of GLONASS K-1 [C]. In: The 24th international technical meeting of the satellite division, Portland, QR
8. Erker S, Tholert S, Furthner J et al (2009) GPS L5 "Light's On"—a first comprehensive signal verification and performance analysis. In: The 22th international technical meeting of the satellite division, Savannah, GA
9. Hegarty CJ, Ross JT (2010) Initial results on nominal GPS L5 signal quality. In: The 23th international technical meeting of the satellite division, Portland, OR
10. Ouyang X (2013) Evalution technology and analysis of BDS signal quality. National University of Defense Technology, Changsha
11. He C (2013) Research on evaluation methods of GNSS signal quality and the influence og GNSS signal on ranging performance. Chinese Academy of Sciences, Xian
12. USA & EU (2006) Joint statement on Galileo and GPS signal optimization. By the European Commission and the United States, Brussels

Analysis Methods of Linear Distortion Characteristics for GNSS Signals

Tao Yan, Ying Wang, Bo Qu, Xiao Liu and Guoyong Wang

Abstract The non-ideality of payload channel can cause the navigation signal distortions, which would lead to the performance degradation. The distortions due to non-ideal channel fall into the linear and nonlinear signal distortions. The linear distortion is usually introduced by filter and triplexer, and non-linear distortion mainly depends on high power amplifier. Analyzing the non-ideal characteristics of channel is helpful to optimize the navigation payload design and improve the navigation signal quality. This paper only focuses on the analysis of linear distortion characteristics. The linear distortion characteristics include the amplitude-frequency characteristic and group delay characteristic. The non-ideality of amplitude-frequency characteristic can result in the distortion of signal frequency spectrum and power loss. The non-ideality of group delay characteristic can lead to the distortion of correlation function, and then introduce ranging bias error. This paper firstly presents the analysis methods from the view of correlation-domain and time-domain. The correlation-domain method exploits the real correlation function and ideal correlation function to obtain directly linear distortion characteristics. The time-domain method needs to recover the baseband waveform with the periodicity of short code signal component, and generate approximately the linear distortion characteristics using a FIR filter. Then, taking BPSK signals as an example, the performance of above two methods is studied by simulation.

Keywords Linear distortion · Correlation function · Time-domain waveform GNSS

T. Yan (✉) · Y. Wang · B. Qu · X. Liu · G. Wang
China Academy of Space Technology (Xi'an), 504 Chang'an Avenue, Xi'an 710100, China
e-mail: iamyantao@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_17

1 Introduction

The channel of satellite navigation signal generation is composed of up-converter, high power amplifier, filter and triplexer [1]. The characteristics of channel determine the navigation signal quality. The overall performance of the system is closely related to the navigation signal quality. To ensure the service performance of the global navigation satellite system (GNSS), the signal quality of GPS and Galileo have been fully analyzed and evaluated [2–4].

In the practical environment, each component of the channel may be non-ideal. The non-ideality would lead to the signal distortion, including linear distortion and nonlinear distortion [5, 6]. Nonlinear distortion is mainly introduced by high power amplifier. In general, AM/AM and AM/PM can be used to describe the nonlinear distortion [7]. The linear distortion characteristics include amplitude-frequency characteristic and group delay characteristic. The non-ideal amplitude-frequency characteristic can result in signal spectrum distortion and power loss, and the undesirable group delay characteristic will lead to the distortion of correlation function and introduce the ranging bias error [8].

The linear distortion characteristics, especially the group delay characteristic, have a direct influence on the navigation signal, which is the focus of this paper. Analyzing the distortion characteristics of the channel is helpful for improving the navigation signal quality. For example, introducing the predistortion scheme [9]. The group delay characteristic of the channel can be measured by instrument on the ground [10]. However, it cannot represent the characterize on orbit satellites in real time. It is an effective method to extract the channel characteristics from the received signal.

The methods of analyzing linear distortion mainly include the time-domain method and the correlation-domain method. The time domain method uses a FIR filter to approach the fact linear distortion characteristic. The ideal signal is distorted by the FIR filter. When the time-domain waveforms of the distorted signal and the actual received signal have the minimum mean square error, the coefficients of the FIR filter are obtained [11]. For the constant envelope multiplexing signals with multiple signal components, the Ref. [12] proposes to solve the FIR coefficients by satisfying the maximum likelihood probability under the premise of considering the distortion of each signal component. The correlation-domain analysis method firstly obtains the correlation functions of the real signal and the ideal signal, and then transforms the correlation functions to the frequency domain. Finally, the linear distortion transfer function is directly obtained [13].

This paper first presents the signal model. On this basis, the analysis methods of time-domain and correlation-domain to estimate the linear distortion characteristics are shown. Finally, the performance of the above two methods are analyzed by taking BPSK as example. The results can provide an reference for optimizing the signal quality and enhancing the accuracy of channel characteristics analysis.

2 Signal Model

Different from the traditional GPS signal, the modernized GNSS signal needs to broadcast multiple signal components at the same frequency, and the constant envelope multiplexing method is applied. For the N binary signal components, the baseband expression of the constant envelope signal is [14]:

$$s(t) = \sum_{n=1}^{N} A_n e^{j\theta_n} s_n(t) + IM(t) \tag{1}$$

where $IM(t)$ is the inter-modulation terms to achieve the constant envelope. The power ratio and phase relationship between signal components are determined. The signal component in GNSS is usually binary signal, such as BPSK, BOC and TMBOC. They can be expressed as

$$s_n(t) = d(t)c(t)sc(t) \tag{2}$$

where $d(t)$ denotes the navigation message or secondary code. $c(t)$ is the PRN code. $sc(t)$ denotes the binary subcarrier. The PRN codes are usually ideal. Thus, all the signal terms can be treated as uncorrelated components.

The linear distortion characteristics of channel can be characterized by the unit impulse response $h(t)$. The distorted signal is expressed as:

$$\hat{s}(t) = s(t) \otimes h(t) \tag{3}$$

where $\otimes$ denotes the convolution operation. Accordingly, the single signal component n after distortion is expressed as:

$$\hat{s}_n(t) = s_n(t) \otimes h(t) \tag{4}$$

At the receiving end, the signal component n is generated locally, and the cross-correlation function is expressed as:

$$\begin{aligned} \hat{R}_n(\tau) &= \frac{1}{T_p} \int_{T_p} \hat{s}(t) s_n^*(t-\tau) dt \approx \frac{1}{T_p} \int_{T_p} \hat{s}_n(t) s_n^*(t-\tau) dt \\ &= s_n(\tau) \otimes s_n^*(\tau) \otimes h(\tau) = R_n(\tau) \otimes h(\tau) \end{aligned} \tag{5}$$

where T_p is the coherent integrated time, which is usually a prime code period. $R_n(\tau)$ is the auto-correlation function of the signal component n. Therefore, the receiving performance is determined by the cross-correlation function [12]. By the Fourier transform, the cross-power spectrum of the signal component n can be obtained, i.e.

$$\tilde{G}_n(f) = \mathcal{F}\left[\hat{R}_n(\tau)\right] = G_n(f)H(f) \tag{6}$$

where $\mathcal{F}[\cdot]$ denotes the Fourier transform, $G_n(f) = \mathcal{F}[G(f)]$ is the power spectrum density of signal component n. $H(f) = \mathcal{F}[h(t)]$ is the corresponding transfer function of $h(t)$. According to $H(f)$, the required amplitude-frequency characteristic and group delay characteristic can be calculated. $H(f)$ can be expressed as $H(f) = |H(f)|e^{j\varphi(f)}$. The amplitude-frequency characteristic and group delay characteristic are respectively:

$$\begin{aligned} \mathrm{Mag}(f) &= 20\lg(|H(f)|) \\ \mathrm{GD}(f) &= -\frac{1}{2\pi f}\frac{d\varphi(f)}{df} \end{aligned} \tag{7}$$

3 Analysis Methods

3.1 Time-Domain Method

The time-domain method is based on (3) and (4). The group delay characteristic is estimated by using the time domain waveform of the ideal signal and the time domain waveform after distortion. In the time-domain method, the FIR filter is usually used to approximate the $h(t)$.

When only using the time-domain waveform of single signal component n to estimate the linear distortion characteristics, it is necessary to separate the baseband time domain waveform from the multiplexed signal. The foundation of baseband waveform separation is the periodicity of civilian signal, and the randomness of message secondary code of different signal components. The baseband waveform can be obtained by accumulating multiple periods [15].

The FIR filter is used to approximate the channel characteristics. The separated time-domain waveform and the ideal time-domain waveform are discretized, which are denoted as $\tilde{s}_n(p)$ and $s_n(p)$, $p = 0,1,2, \ldots, P$. pT_s is the sampling time, and $f_s = 1/T_s$ is the sampling rate. The order of FIR filter is M order, i.e. there are a total of $M + 1$ coefficients. After FIR filtering, we obtain

$$h(p) \otimes s_n(p) = \sum_{m=0}^{M} h(m)s_n(p-m) \tag{8}$$

In order to solve the coefficients of FIR filter, we need to minimize the next expression.

$$E=\sum_{p=1}^{P}\left|\tilde{s}_n(p)-h(p)\otimes s_n(p)\right|^2 \tag{9}$$

The order of the FIR filter is finite, and the solution of formula (9) is as follows:

$$h=\left(\mathbf{A}^{\mathrm{H}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{C} \tag{10}$$

where $(\cdot)^{\mathrm{H}}$ denotes the conjugate transposed. $\mathbf{A}$ is the convolution matrix of $s_n(p)$, and $\mathbf{C}$ is a column vector of $\tilde{s}_n(p)$. It is obvious that the estimation accuracy of the single component time domain waveform is related to the separated precision of the baseband waveform.

The single signal component method uses the recovered baseband waveform to estimate the linear distortion characteristics. In Ref. [12], a method of estimating the linear distortion characteristics using the multiplexing signal directly is presented, which is similar to the single signal component method.

After receiving the navigation signal using a high gain antenna and removing the Doppler, the data is expressed as:

$$\tilde{y}(t)=y(t)\otimes h(t)+n(t) \tag{11}$$

The digital distortion of each signal component and the linear distortion of the channel are considered at the same. The digital distortion of the nth signal component is characterized by parameters η_n, which indicates the delay of the falling edge of the waveform is η_n, i.e.

$$y(t)=\sum_{n=1}^{N}A_n e^{j\theta_n}s_n^{\eta_n}(t)+IM^{\boldsymbol{\eta}}(t) \tag{12}$$

The sampling data of $y(t)$ and$\tilde{y}(t)$ are denoted as

$$\begin{aligned}\mathbf{y}&=[y_1,y_2,\ldots,\mathrm{y}_P]^T\\ \tilde{\mathbf{y}}&=[\tilde{y}_1,\tilde{y}_2,\ldots,\tilde{\mathrm{y}}_P]^T\end{aligned} \tag{13}$$

respectively. Let the unit impulse response be expressed as

$$\mathbf{h}=[h_1,h_2,\ldots,\mathrm{h}_M]^T \tag{14}$$

The digital distortion of the N signal components is expressed as

$$\boldsymbol{\eta}=[\eta_1,\eta_2,\ldots,\eta_n,\ldots,\eta_N]^T,\quad \eta_n\in[-1,1] \tag{15}$$

The multiplexing signal with digital distortion but no linear distortion is expressed as

$$\mathbf{x}(\boldsymbol{\eta})=[x_1, x_2, \ldots, \mathrm{x}_P]^T \tag{16}$$

Let $\mathbf{X}(\boldsymbol{\eta})$ denote the convolution matrix of $\mathbf{x}(\boldsymbol{\eta})$. The estimation problem of digital distortion and linear distortion can be written as:

$$\left(\hat{\eta}, \hat{\mathbf{h}}\right) = \arg\min_{\boldsymbol{\eta},\mathbf{h}} \left\{ \|\tilde{\mathbf{y}} - \mathbf{X}(\boldsymbol{\eta})\mathbf{h}\|_2^2 \right\} \tag{17}$$

The solution of (17) is

$$\begin{aligned} \hat{\eta} &= \arg\min_{\boldsymbol{\eta}} \left\{ \tilde{\mathbf{y}}^H \mathbf{X}(\boldsymbol{\eta}) \left(\mathbf{X}^H(\boldsymbol{\eta})\mathbf{X}(\boldsymbol{\eta})\right)^{-1} \mathbf{X}(\boldsymbol{\eta})\tilde{\mathbf{y}} \right\} \\ \hat{\mathbf{h}} &= \left(\mathbf{X}^H(\hat{\eta})\mathbf{X}(\hat{\eta})\right)^{-1} \mathbf{X}(\hat{\eta})\tilde{\mathbf{y}} \end{aligned} \tag{18}$$

The advantage of this method is that it can simultaneously estimate the digital distortion parameters and linear distortion parameters. However, the estimation accuracy of multiplexing signal is limited with a relative higher computation complexity, and the high SNR condition is required.

3.2 Correlation-Domain Method

The correlation-domain method is based on (5) and (6). The correlation function of the distorted signal $R_n^{\mathrm{Real}}(\tau)$ and the correlation function of the ideal signal $R_n^{ideal}(\tau)$ can be obtained through the correlation operation. Theoretically, a correlation function can be generated during each code period, and then a transfer function can be estimated. However, due to the existence of cross-correlation and thermal noise, the error would be introduced. In order to reduce the influence of random errors, the correlation functions of multiple periods are averaged.

According to (6), the estimated transfer function is expressed as:

$$H(f) = \frac{\tilde{G}_n(f)}{G_n(f)} \tag{19}$$

Figure 1 shows the ideal power spectral density of BPSK (1) and BOC (1,1) modulated signals. It can be seen that the ideal power spectrum includes multiple zero values. When they become denominator, it would lead to large deviation in the estimation at zero value.

In order to reduce the error at the zero value of spectrum, the following formula can be used.

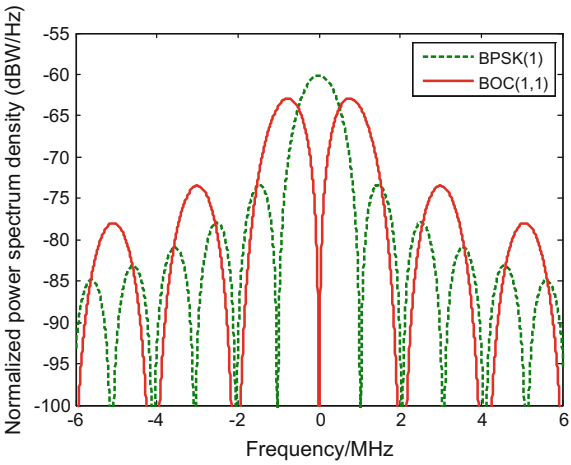


Fig. 1 Power spectrum density of BPSK and BOC

$$H(f) \cong \frac{\int H_{Average}(f') \cdot \tilde{G}_n(f - f')df'}{\int H_{Average}(f') \cdot G_n(f - f')df'} \tag{20}$$

where $H_{Average}(f)$ is a moving average filter, whose bandwidth is much less than the minimum chip rate. The $H_{Average}(f)$ can be a rectangular window function, and the bandwidth is determined according to the required accuracy.

4 Simulation Analysis

In this paper, the accuracy of group delay estimation is considered in the case of single signal component. The transfer function $H(f)$ of linear distortion characteristics come from the Ref. [9]. The amplitude-frequency characteristic and group delay characteristic are shown in Fig. 2. The center frequency of the channel characteristics is 1575.42 MHz.

The BPSK(10) signal is analyzed. The signal bandwidth is 20.46 MHz. The carrier-to-noise-density ratio (C/N_0) is 70 dB-Hz. The estimation results are obtained through 100 code periods. Each code period is 1 ms. The sampling rate is 300 MHz.

Firstly, the time-domain method is applied to estimate the linear distortion characteristics. The time domain waveform of the undistorted signal and the distortion signal, are shown in Fig. 3a and b under 20.46 MHz ideal band-limitation. Linear distortion leads to a large component in both the real and the imaginary parts of the time domain waveform. The coefficients of FIR filter are calculated according

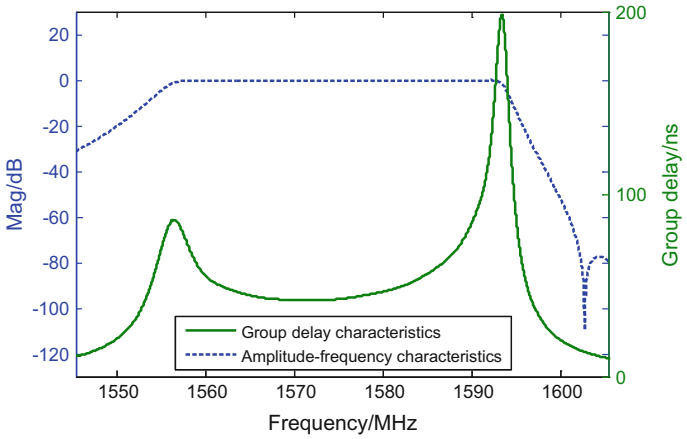


Fig. 2 The linear distortion characteristics

to (10). The estimation results of linear distortion characteristics are shown in Fig. 4. The order of FIR filter is 10, 25, 50 respectively. It can be seen that the time-domain method can have a better estimation accuracy of the linear distortion characteristics during the main-lobe bandwidth ±10.23 MHz. With the increase of the order, the estimated distortion characteristics would be more close to the theoretical results. It should be noted that the amplitude of the estimated results and the group delay values are moved to the theoretical results at the center frequency for the aim of comparison.

Then the correlation-domain method is used. Figure 5a and b show the cross-correlation functions of the undistorted signal and the distortion signal respectively. According to (20), the estimation results of the linear distortion characteristics are shown in Fig. 6. The smooth window function is a rectangular window with 25 kHz bandwidth. Compared with the results of the time-domain method, it can be found that the correlation-domain method has a larger estimation error near the zero value of the spectrum. However, in other regions, the estimated result of the correlation domain method is closer to the theoretical distortion.

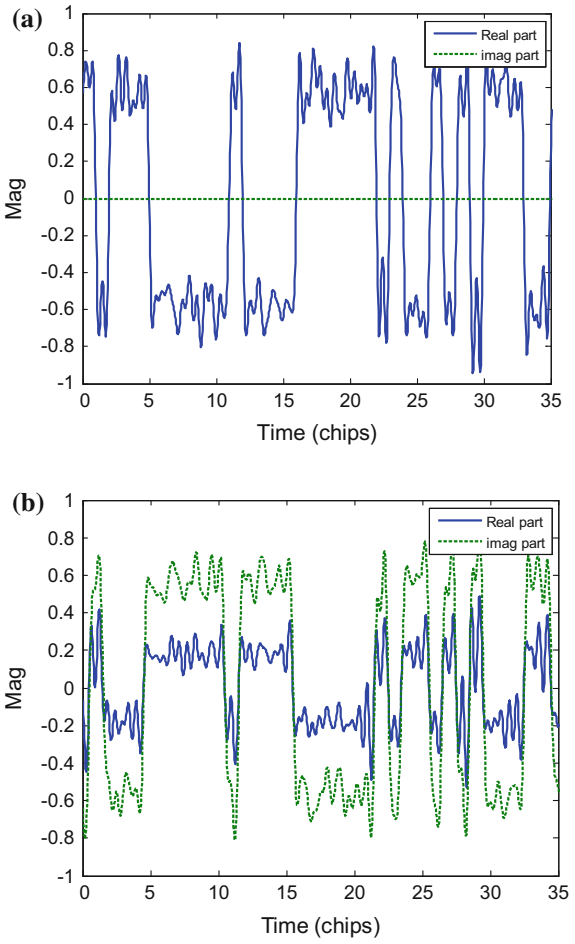


Fig. 3 Time-domain waveform. **a** Undistorted, **b** linear distortion

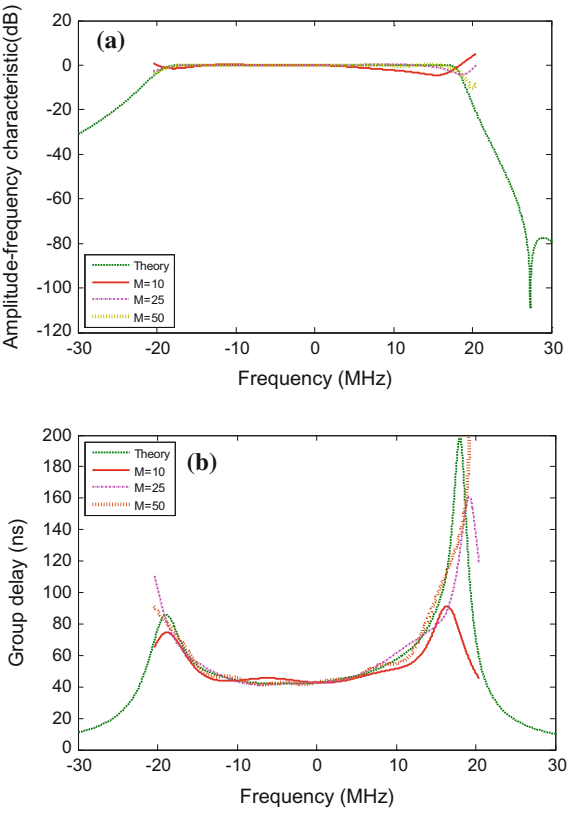


Fig. 4 Time-domain results. **a** Amplitude frequency characteristics, **b** group delay characteristics

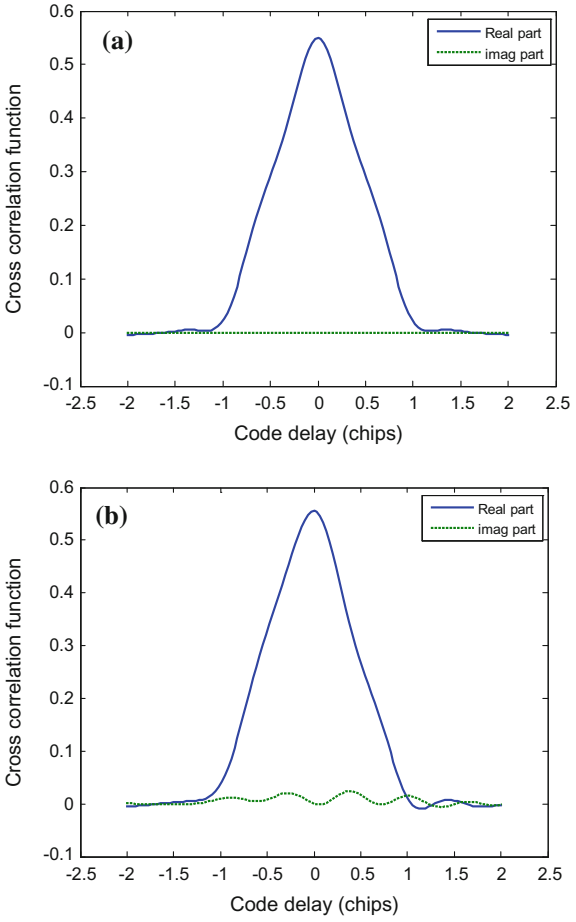


Fig. 5 Cross-correlation functions. **a** Undistorted, **b** linear distortion

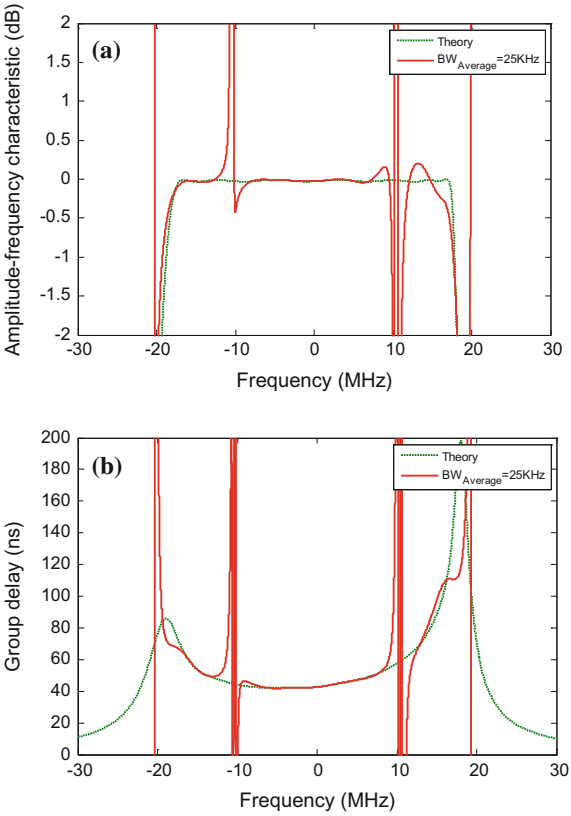


Fig. 6 Correlation-domain results. **a** Amplitude frequency characteristics, **b** group delay characteristics

5 Conclusions

The linear distortion of channel can lead to the degradation of the navigation signal quality. Thus, it is of great significance to improve the signal quality by analyzing the distortion characteristics of the channel. From the two aspects of time-domain and correlation-domain, this paper proposes the analysis method of linear distortion characteristics. The BPSK signal is used to evaluate the performance of above two kinds of methods. The results show that these two kinds of methods have their own advantages and disadvantages. The estimation error of the correlation-domain method is larger at the zero value of the spectrum, but the estimation accuracy during other regions is higher. The correlation-domain method can be used when the details are concerned. These is no existing zero-value problem in the time-domain method, but the overall estimation accuracy is worse than the correlation-domain method, which is more suitable for evaluating the overall distortion characteristics. The results of the paper can provide references for evaluating the linear distortion characteristics of the channel.

Acknowledgements This work is supported by National Natural Science Foundation of China (Grant 61627817; Grant 91438107).

References

1. Chen Y, Kou Y, Zhang Z (2012) Analog distortion of wideband signal in satellite navigation payload. In: CSNC 2012 proceedings, lecture notes in electrical engineering, vol 160, pp 89–100
2. Christie JRI, Bentley PB (2004) GPS signal quality monitoring system. In: Proceedings of the 17th international technical meeting of the satellite division (ION GNSS 2004), vol 9. SunDiego, USA, pp 21–24
3. Spelat M, Hollreiser M, Crisici M, Falcone M (2006) GIOVE-A signal in space test activity at ESTEC. In: Proceedings of the 19th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS 2006), pp 981–993. Fort Worth, TX, Sept 2006
4. Thoelert S, Furthner J, Meurer M (2013) GNSS survey–signal quality assessment of the latest GNSS satellites. In: ION GNSS+2013, 2013, pp 608–615
5. Soellner M, Kohl R, Luetke W, Erhard P (2002) The impact of linear and non-linear signal distortions on Galileo code tracking accuracy. In: ION GPS 2002, pp 1270–1285
6. Rapisarda M, Angeletti P, Casini E (2007) A simulation framework for the assessment of navigation payload non-idealities. In: 2nd workshop on GNSS signals & signal processing-GNSS signals
7. Li X, Geng S, Ou G, Zhuang Z (2008) A study on PN code ranging error caused by nonlinear power amplifier. J Nat Univ Defense Technol 30(5):49–55
8. Zeng W, Li B, Liu W, Ou G (2015) Impact of non-ideal channel characteristic on correlation peak of navigation signal. GNSS World of China 40(5):71–85
9. Yang Z, Xu Q, Han H, Zhang L (2015) Research on pre-distortion technology for navigation transmitting channel. In: CSNC 2015
10. Patidar P, Upadhyay D, Sarkar S (2016) Measuring navigation payload absolute delay in radiation mode, pp 49–53. Inside GNSS, Jan/Feb 2016
11. Graf S, Gunther C (2016) Analysis of GIOVE-A L1-signals. In: ION GNSS: 19th international technical meeting of the satellite division. Fort Worth, TX, pp 1560–1566
12. Thoelert S, Vergara M, Sgammini M, Enneking C, Antreich F (2014) Characterization of nominal signal distortions and impact on receiver performance for GPS (IIF) L5 and Galileo (IOV) E1/E5a Signals. In: ION GNSS+2014, pp 3113–3128
13. Soellner M, Kurzhals C, Rapisarda M, Burger T, Erker S, Furthner J, Grunert U, Meurer M, Tholert S (2008) GNSS offline signal quality assessment. In: ION GNSS 2008, pp 909–920
14. Yan T, Tang Z, Wei J, Qu B, Zhou Z (2015) A quasi-constant envelope multiplexing technique for GNSS signals. J Navig 68(4):791–808
15. Lu H, Lian B, Yan H (2016) A single channel separation algorithm for new multiplexing satellite navigation signals. J Xi'an Jiaotong Univ 50(12):32–37

The Discussion on Local Optimization of Navigation Constellation Based on STK/MATLAB

Jiangyan Sun, Caibo Hu, Yueyue Li and Lianqing Lin

Abstract Combined with the development background of "The Belt and Road Initiative", the constellation configuration model of the BeiDou-2 is simulated in this paper. Without increasing the number of satellite, the track of subsatellite point is shifted through changing the orbit parameters to expand the system coverage area. The STK and MATLAB are interlinked, and the orbit parameters are optimized by MATLAB with powerful programming and calculation function. The optimization result shows that service ability in extended area can be realized meanwhile high accuracy positioning in key area to some extent can be improved. The system will offer service at the lowest cost and the shortest time to achieve the greatest benefits. In addition, it also provides reference for future location of satellite drift to enhance effectively local combat security system.

Keywords The belt and road initiative · STK/MATLAB · Satellite drift Constellation optimization

1 Introduction

Since 2004, BeiDou-2 Satellite Navigation System project has started, and research construction has been carried out. In 2012, 14 working satellites networking were successfully established and the system was officially put into operation which provided regional services to China and its surrounding areas. In order to enhance the system construction, it has entered the peak period of satellite networking from 2009. By the end of 2018, 18 satellites will be launched serving the countries along "The Belt and Road Initiative" and it is expected to provide service to users around

J. Sun (✉) · C. Hu · Y. Li · L. Lin
Beijing Satellite Navigation Center, Beijing, China
e-mail: sjy7297@126.com

C. Hu
GNSS Research Center of Wuhan University, Wuhan, China
e-mail: hucaibo@whu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_18

the world by 2020 [1, 2]. In 2015, China put forward the great strategic concept of building the "Silk Road Economic Belt" and the "21st Century Maritime Silk Road." which is an important strategic platform for the cooperation and development of emerging economies [3]. The development of countries and regions along the "The Belt and Road Initiative" is bound to be escorted by Beidou Satellite Navigation System so as to achieve military and civilian integration and international cooperation. Given the timeframe (around 2020) of Beidou Global System networking and the provision of services, combining with the complex relationship between the navigation geometric precision factor and the visible number of satellite in which there is a positive correlation in a certain range, and a negative correlation in a certain range, it is necessary to apply a reasonable constellation geometry to balance the two in accordance with the number of existing satellites [4].

In this paper, given the current operation situation of the three kinds of satellite hybrid networking in BeiDou-2, the discussion on local optimization of system constellation configuration was made. Without increasing the number of satellites, and a constraint that service performance (only from a precision point of view, without considering the integrity, continuity and availability) in the key area not declining, and taking the part of the satellite orbit parameters as the control variables and the GDOP value weighted in the service area grid as the objective function of optimization, and combining MATLAB with STK, the Levenberg-Marquardt algorithm based on MATLAB is iteratively calculated to determine the optimal orbital parameters of the satellite and STK is used to verify the analysis results. The optimization result shows that service ability in extended area can be realized meanwhile high accuracy positioning in key area to some extent can be improved.

2 Navigation Constellation Model

2.1 STK Software

STK full name is Satellite Tool Kit (Satellite Toolbox), which is a commercial analysis software developed by the United States Analytical Graphics company and mainly used in the aerospace field simulation. The software is an important tool for space technology and application researchers, which supports diverse tasks in aerospace, missile defense, intelligence reconnaissance and surveillance [5]. For orbital design and control, simulations and theoretical calculations are helpful in identifying the best constellation to achieve system service performance before the actual satellite launch. Digital simulation technology is the best choice of simulation methods, which has the advantages of high accuracy, wide adaptability and good universality [6]. Using the interface module provided by STK and MATLAB

for joint simulation, the satellite orbit simulation system is constructed and the satellite orbit and attitude control is realized [7].

2.2 BeiDou-2 Constellation

As of the end of 2012, the working satellites of Beidou-2 on-orbit is a hybrid constellation consisting of 5 GEO+5 IGSO+4 MEO in 3 different orbital altitude. Among them: the 5 GEO satellites have a height of 35786 km, fixed at 58.75°E, 80°E, 110.5°E, 140°E and 160°E respectively; IGSO satellite with the orbital altitude 35786 km and the orbit inclination 55°, run on 3 orbital planes and longitude of ascending node of difference of 120° respectively. IGSO-4 and IGSO-5 satellites are in-plane backup with IGSO-1 and IGSO-2 respectively, whose longitude of intersection is 95°E; 4 MEO satellites are located on the 7,8 phases of the first orbital plane and the 3,4 phase of the second orbital plane in the Walker 24/3/1 constellation respectively, whose orbital height is 21528 km and the orbital inclination is 55°, and longitude of the ascending node in the first orbital plane is 0° [8]. According to the six parameters of satellite orbit, STK is applied to simulation modeling, as shown in Figs. 1 and 2.

3 Extended Area Settings

Due to the limitation of orbital resources, the visible time of different satellite and the different track of subsatellite point, in addition, the current networking situation and the mid-orbit satellite used for paving the way for the global satellite navigation

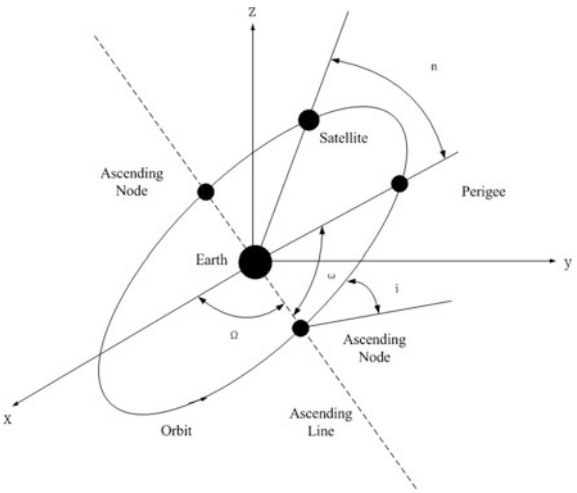


Fig. 1 Six parameters of orbit

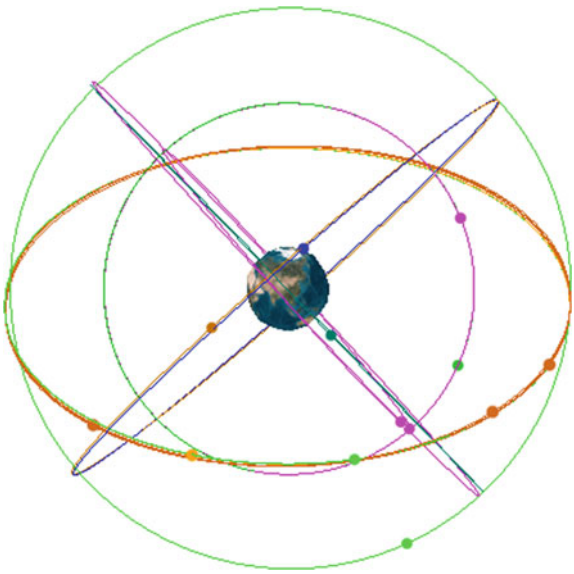


Fig. 2 Three-dimensional constellation model of BeiDou-2

system, so IGSO satellites with "8" shaped track of subsatellite point become the preferred satellite type for constellation optimization. This paper considers starting from the orbit parameters of 5 IGSO satellites whose true anomaly (0°–360°) being designed as control variables (only changing the phase of the satellite on-orbit and not changing the orbit of the satellite), and the track of subsatellite point of a preferable IGSO satellites will be shifted overall, as shown in Fig. 3.

BeiDou-2 focuses on services covering the vast region of Asia Pacific. Without reducing its service performance, according to the geographical distribution of countries and regions along the "The Belt and Road Initiative", the extended range of analysis is set to be 35°E–140°E and 0°N–55°N in this paper, as shown in Fig. 4.

4 Constellation Optimization Analysis

4.1 Interface Settings and Simulation Optimization Process

The limitation of STK is that it requires the user to constantly set the parameters of the satellite orbit to change the value and output the analysis result (such as the DOP value), which is too inefficient to directly give the relationship between the DOP value and the orbit parameters so as to optimize too slowly. So the third-party applications are used to interconnect. As shown in Fig. 5, the interface between MATLAB and STK is set in this paper that the constellation model is constructed by MATLAB and the Least Squares problem is solved with Levenberg-Marquardt

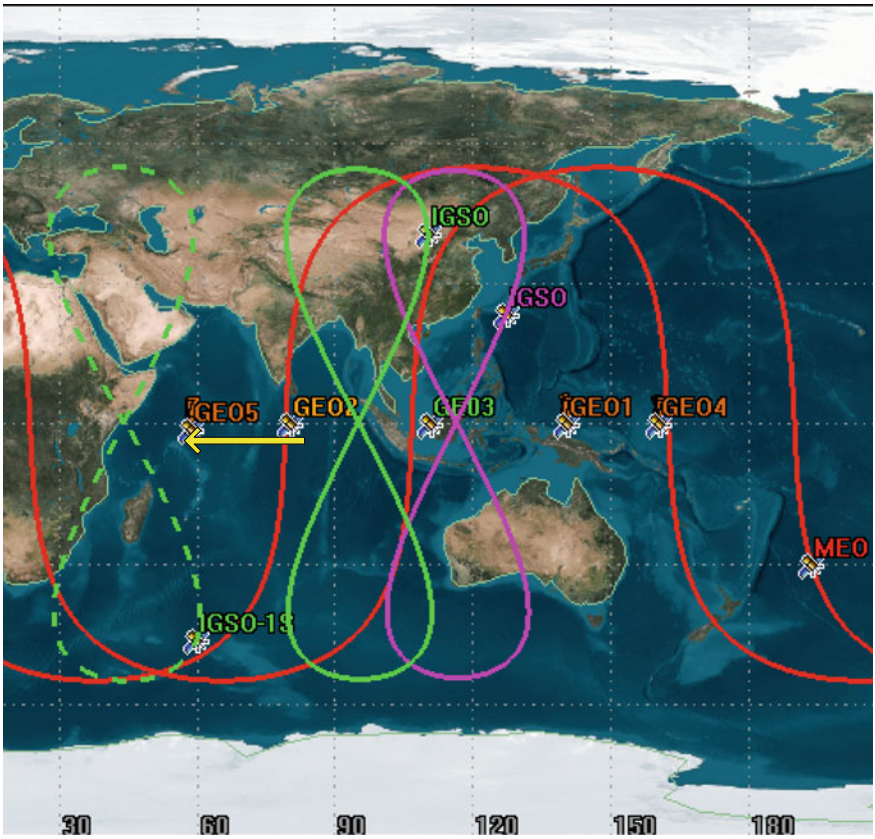


Fig. 3 The offset of IGSO subsatellite point track diagram

algorithm, and the optimal orbit parameters is obtained through iterative looping, as shown in Fig. 6.

4.2 Single Point GDOP Value Calculation

To determine the user's three-dimensional position (x_u, y_u, z_u) and the time offset t_u, pseudo-range measurement of j satellites were measured and equations were generated.

$$\rho_j = \left\| \boldsymbol{s}_j - \boldsymbol{u} \right\| + ct_u \tag{1}$$

A single pseudorange is expressed as

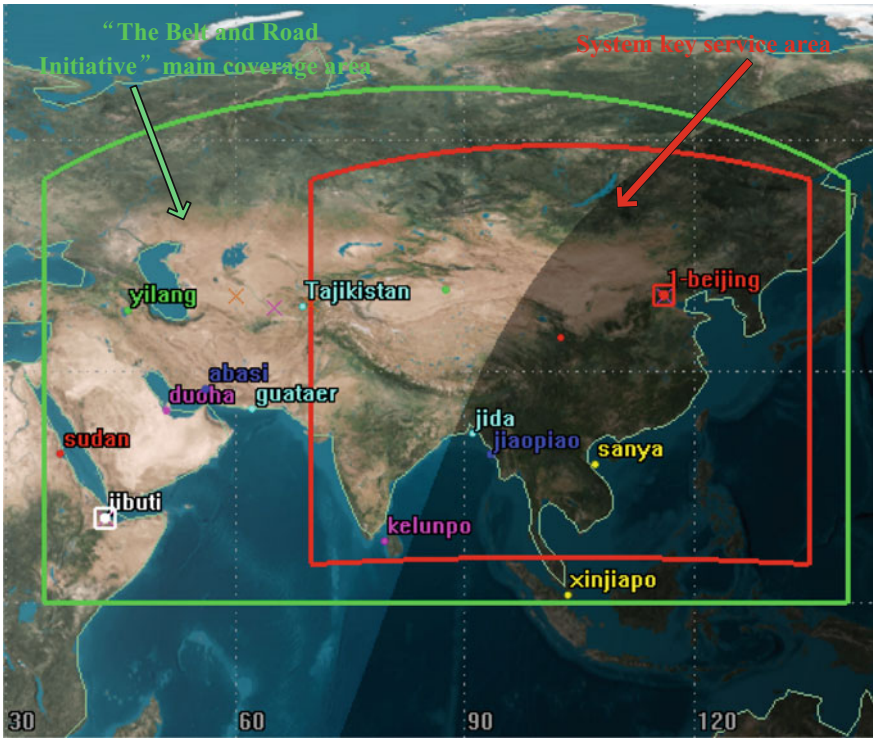


Fig. 4 Key service area and main coverage of "The Belt and Road Initiative"

$$\begin{aligned}\rho_j &= \sqrt{(x_j - x_u)^2 + \left(y_j - y_u\right)^2 + (z_j - z_u)^2} + ct_u \\ &= f(x_u, y_u, z_u, t_u)\end{aligned} \tag{2}$$

where,

j	refers to different satellites;
x_j, y_j and z_j	represent the three-dimensional position of jth satellite;
t_u	represents the offset between the receiver clock and the system time; c represents the speed of light.

If we know the receiver's position approximately, then the offset between true position (x_u, y_u, z_u) and approximate position $(\widehat{X_u}, \widehat{Y_u}, \widehat{Z_u})$ can be marked by displacement $(\Delta x_u, \Delta y_u, \Delta z_u)$. The formula (1) can be expanded according to the Taylor series in a approximate position. The position offset $(\Delta x_u, \Delta y_u, \Delta z_u)$ can be expressed as a linear function of the known coordinates and pseudorange measurements [9]. Writing in matrix form:

```
function Untitled1
remMachine = stkDefaultHost;
delete(get(0,'children'));
conid=stkOpen(remMachine);

dtr = pi/180;
rtd = 180/pi;
scen_open = stkValidScen;

if scen_open == 1
   rtn = questdlg('Close the current scenario?')
   if ~strcmp(rtn,'Yes')
      stkClose(conid)
      return
   else
      stkUnload('/*')
   end
end
stkNewObj('/','Scenario','yidaiyilu');

......
```

Fig. 5 Interface settings between STK and MATLAB

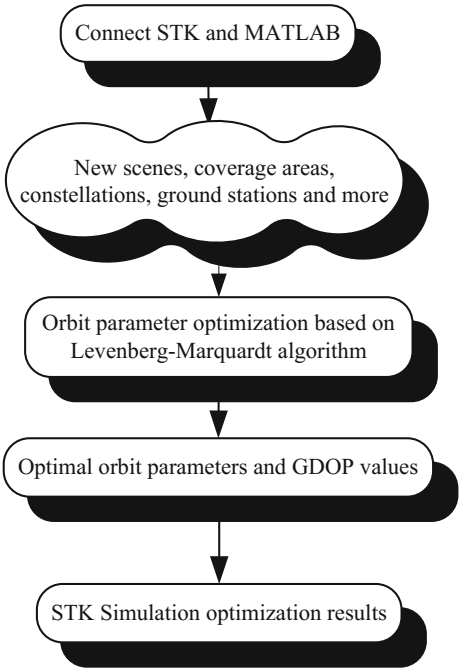


Fig. 6 Simulation analysis and optimization process

$$\Delta\rho = \begin{bmatrix} \Delta\rho_1 \\ \Delta\rho_2 \\ \vdots \\ \Delta\rho_j \end{bmatrix} \quad H = \begin{bmatrix} a_{x1} & a_{y1} & a_{z1} & 1 \\ a_{x2} & a_{y1} & a_{z1} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{xj} & a_{y1} & a_{z1} & 1 \end{bmatrix} \quad \Delta x = \begin{bmatrix} \Delta x_u \\ \Delta y_u \\ \Delta z_u \\ -c\Delta t_u \end{bmatrix} \tag{3}$$

where $a_{xj} = \frac{x_j - \widehat{x_u}}{\widehat{r_j}}, a_{yj} = \frac{y_j - \widehat{y_u}}{\widehat{r_j}}, a_{zj} = \frac{z_j - \widehat{z_u}}{\widehat{r_j}}$;

$\widehat{r_j} = \sqrt{\left(x_j - \widehat{x_u}\right)^2 + \left(y_j - y_u\right)^2 + \left(z_j - \widehat{Z_u}\right)^2} (j \geq 4)$ is the distance from satellite to the user's.

$$\Delta\boldsymbol{\rho} = \mathbf{H}\Delta\mathbf{x} \tag{4}$$

Its solution is

$$\Delta\mathbf{x} = \mathbf{H}^{-1}\Delta\boldsymbol{\rho} \tag{5}$$

The relation of GDOP is given in the form of $\left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}$ component

$$\left(H^{T}H\right)^{-1} = \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & D_{24} \\ D_{31} & D_{32} & D_{33} & D_{34} \\ D_{41} & D_{42} & D_{43} & D_{44} \end{bmatrix} \tag{6}$$

GDOP can be calculated as the square root of $\left(H^{T}H\right)^{-1}$ matrix trace

$$\mathrm{GDOP} = \sqrt{D_{11} + D_{22} + D_{33} + D_{44}} \tag{7}$$

4.3 Optimization Algorithm Design

On the one hand, the navigation constellation optimization problem can be simply described as an optimization problem that the orbit parameters of satellite are control variables and the weighted geometric precision factor WGDOP in the region is the objective function. This paper designs a constraint that service performance in the key area not declining and the objective function WGDOP being the weighted average of the GDOP values at each point in the extended area within 7 days and the true anomaly of IGSO satellite being the control variables (range 0°–360°). The horizontal grid of this extended area is set up and the grid range is set to be 2° × 2° in which the center point of latitude and longitude of each grid is

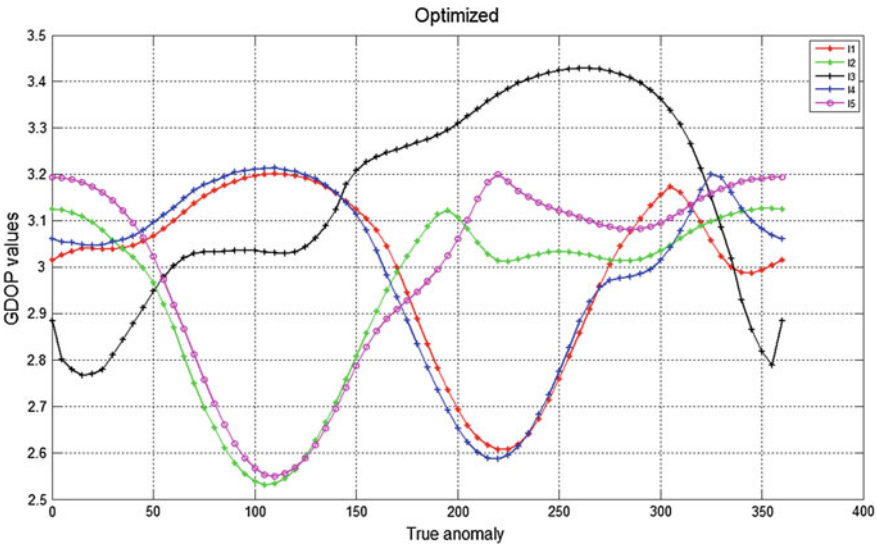


Fig. 7 The variation of GDOP values in extended area in case of 5 IGSO satellites with different true anomaly

considered as the sampling point. The weight of each grid are all set to be 1 and the stack of GDOP values calculated for all grids is obtained.

$$G = \sum_{i=1,j=1}^{m,n} g(i,j)GDOP_{ij} \tag{8}$$

Levenberg-Marquardt algorithm is used to solve the least squares problem and the optimal true anomaly of five IGSO satellites are respectively obtained through loop iteration. The results are shown in Fig. 7, the track of subsatellite point of IGSO-2 is shifted which will have a relatively optimal effect on the GDOP values in extended area.

4.4 Optimization Results Analysis

Through the above MATLAB calculation and STK simulation validation, the changes of average GDOP value in key position and area before and after the optimization are obtained in Table 1. The average GDOP distribution in the extended area before and after optimization is shown in Figs. 8 and 9.

Table 1 GDOP value comparison before and after optimization

	Before optimization	Optimized	Reduction (%)
Key area	2.532	2.355	6.99
Expanded area	4.880	2.534	48.07
Beijing	2.623	2.529	3.72
Sanya	2.268	2.260	0.35
Kashi	3.450	2.725	26.61
Gwadar	3.621	2.278	58.96

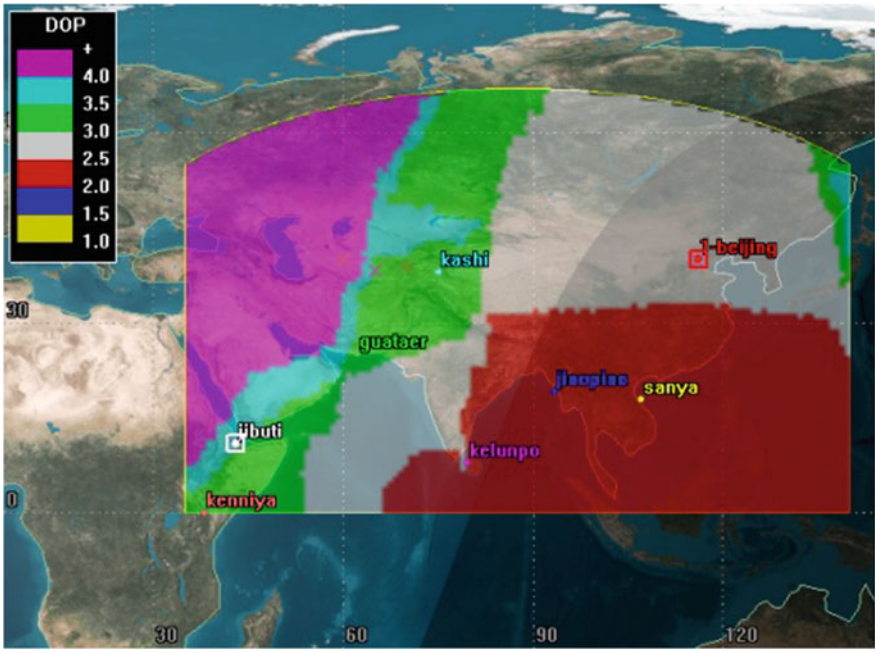


Fig. 8 GDOP value distribution in extended area before optimization

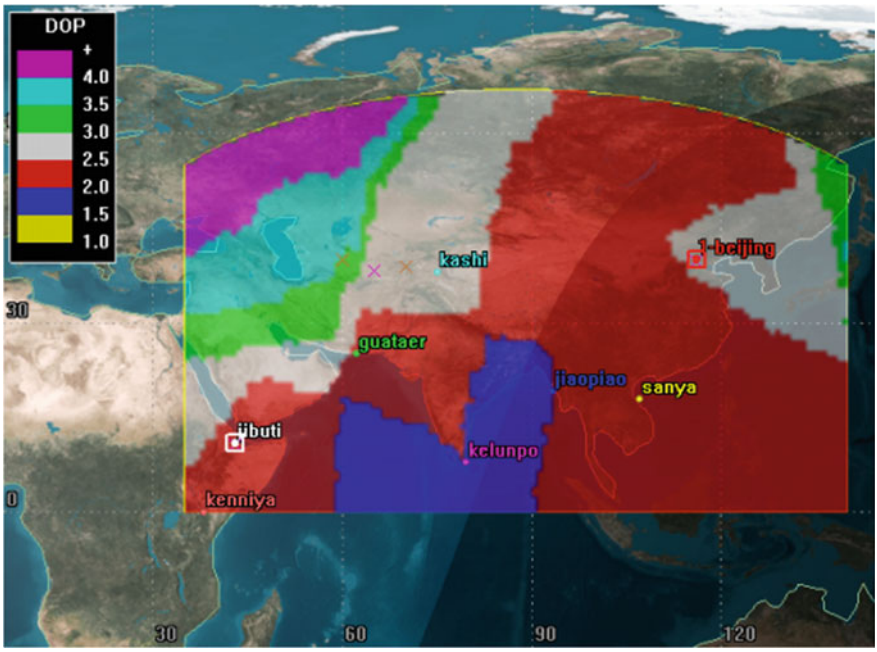


Fig. 9 GDOP value distribution in extended area after optimization

5 Conclusion

The implementation of the "The Belt and Road Initiative" development strategy can not be achieved without the Beidou Satellite Navigation System. Due to the limited number of satellites and orbit resources and the track of subsatellite point of satellites on different orbit, IGSO satellites become the preferred satellite to expand system coverage. The track of subsatellite point is shifted through optimizing satellite orbit parameters that is also an important means of improving service performance in the short term at low cost. The following discussions are mainly made in this paper:

1. STK and MATLAB are effectively interconnected through the interface setting, and the model of system constellation is constructed by MATLAB;
2. With a constraint that service performance in the key area not declining, the true anomaly of IGSO satellite being the control variable and the GDOP value weighted in the extended area being the objective function, the optimal value of orbit parameters is obtained by using MATLAB with powerful programming calculation and design optimization algorithm.
3. The analysis result shows that service ability in the extended area along "The Belt and Road Initiative" can be realized meanwhile high accuracy positioning in key area to some extent can be improved.

4. During the period of satellite failure or overterm service, it laid the theoretical foundation for changing the orbit parameter to adjust the constellation at the lowest cost and at the shortest time. In addition, it also provides reference for future location of satellite drift to enhance effectively local combat security system.

References

1. Ran C (2014) BeiDou satellite navigation system operation and development. J Satell Appl, 7–10
2. Ye L (2018) China Beidou shines in space. In: Bright Daily. Available via Dialog. http://www.beidou.gov.cn/. Accessed 4 Jan 2018
3. Cai C (2014) Tactic study on emerging economies participating in the construction of new silk road. Int Trade, 25–29
4. Liu X et al (2012) Optimization of navigation constellation by a multi-objective evolutionary algorithm based on decomposition. Microcomput Inf, 214–216
5. Zhang Y, Fan P, Liu H (2011) Master and proficient in the satellite toolbox STK, pp 1–2
6. Ding S (2016) STK use techniques and manned space engineering applications. National Defense Industry Press, Beijing, pp 1–2
7. Liu M, Du H, Xu C (2009) TT&C station tracking prediction based on STK/Matlab. In: System simulation technology and application conference proceedings
8. BDS-SIS-ICD-B1I-1.0 (2012) BeiDou navigation satellite system signal. In: Space interface control document open service signal B1I (Version 1.0). China Satellite Navigation Office
9. Kaplan ED (2007) GPS principle and application. Electronics Industry Press, Beijing, pp 39–41

Orbits Quality and PPP Application for Galileo

Guofeng Ji, Hang Chang, Zhiqiang Yang, Xiaolin Jia and Liang Zhang

Abstract The orbit quality and PPP application of the Galileo precise products from five MGEX analysis centers (AC) are analyzed in this article, the influence of Solar Radial Pressure (SRP) model on Precise Orbit Determination (POD) is studied. The results show that: the quality of Galileo orbit from GFZ and TUM is significantly improved if the Cuboid Box-Wing (CBW) priori model is used. The SLR residual amplitudes reduced from 20 cm down to 10 cm, the STD has decreased about 50%, and the mean SLR bias is reduced from −2.0 to −6.0 cm to millimeters level; and PHMs clock-offsets daily linear fitting results also show that the periodic signature could be significantly reduced, which is consistent with the results of SLR residuals. CODE's ECOM2 model shows slightly worse performance than CBW + ECOM1 model, but much better than classic ECOM1 model. In addition, the mean SLR bias of CODE's and WHU's orbits is about −2.0 to 4.0 cm, too larger than the result of other Analysis Center, so that the impact of albedo model and antenna thrust require appropriate consideration. Precise point positioning (PPP) error of Galileo-only is about 1.0 cm in horizontal, which is very close to the results of GPS-only, but larger than 3.0 cm in vertical component due to the limited number of active satellites. The convergence speed of GPS/Galileo combined solution is faster than any single system, the positioning accuracy in horizontal is improved respect to GPS-only, but a small degradation can be observed in vertical component.

Keywords Galileo system · Precise orbit determination · SLR residuals Solar radiation model · PPP

G. Ji · H. Chang · Z. Yang (✉) · L. Zhang
College of Geology Engineering and Geomantics, Chang'an University, Xi'an, China
e-mail: yang_gps@chd.edu.cn

X. Jia
National Key Laboratory of Geographic Information Engineering, Xi'an, China

X. Jia
Xi'an Research Institute of Surveying and Mapping, Xi'an, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_19

1 Introduction

The European global navigation satellite system Galileo announced its operation services on December 15, 2016. All Galileo satellites are equipped with two passive hydrogen masers (PHMs) and satellite laser ranging (SLR) retro reflector array, routinely transmit signals at 5 frequencies [1]. Galileo has gone through three phases from planning to operation. The first is the so-called Galileo In-Orbit Validation Element (GIOVE) phase, were GIOVE-A and GIOVE-B, launched in December 2005 and April 2008, respectively. The primary task of GIOVE was to validate critical technology components and transmitting Galileo signals from space to ensure the frequency allocation. GIOVE-A and GIOVE-B were decommissioned by the European Space Agency (ESA) in June and July 2012. The second is In-Orbit Validation (IOV) phase, four IOV satellites were placed in two different orbital planes by two dual launched in 2011 and 2011, and they will be part of the operational Galileo constellation. In May 2014, the GAL-104 satellite experienced a sudden power loss and a permanent failure of the E5 and E6 signals transmission. The third is Full Operational Capability (FOC) phase. In August, 2014, the first pair of FOC satellites launched with a Soyuz ST rocket, resulting into wrong, highly eccentric orbits. Fortunately, the perigee have be corrected by a series of orbit-raising maneuvers. After that, another twelve FOC satellites were launched in 2015 and 2016, all of which reached their target orbits. As a result, today there are 17 active satellites have providing initial services for navigation and positioning. On December 12, 2017, another four new FOC satellites were launched with Ariane V and currently under commissioning. Until 2020, Galileo will be evolving towards its global navigation capability with a full constellation consisting of 24 active satellites and 6 spares.

In 2012, the International GNSS Service (IGS) established the Multi-GNSS Experiment (MGEX) [2] in order to prepare the service for new and upcoming GNSS, such as Galileo and BeiDou, and for Regional Navigation Satellite Systems, such as QZSS and IRNSS. The IGS MGEX pilot project provides freely available Galileo observation data as well as orbit and clock products. Preliminary analysis of orbits and clocks provided by different MGEX ACs was carried out by Montenbruck et al. [3], Hackel et al. [4], Steigenberger et al. [5], and systematic errors related to the solar radiation pressure (SRP) mismodeling have been found in all orbit products. And then, CODE [6], GFZ [7], WHU [8] gradually released MGEX solution strategies and evaluated the performance of their own Galileo products. Since SRP modeling is a key issue for precise orbit determination, two different approaches have been developed to cope with the systematic errors. The first one is the so-called Cuboid Box-Wing model established by Montenbruck et al. [9]. This new priori model can reduce the radial orbit errors from 20 cm down to 5 cm outside eclipse phases. The second one is change the orbit model from the classical Empirical CODE Orbit Model (ECOM1) [10] to the extended ECOM2 model [11]. This new SRP model was used by CODE MGEX solutions since January 2015, significantly orbit and clock improvement can be found when Galileo

and QZSS satellites at times when the satellite's attitude is maintained by yaw-steering [12]. Guo et al. [13] also evaluated the Galileo, BDS and QZSS products provided by MGEX, but no Galileo-only PPP analysis was performed due to the limited number of active satellites.

Although several studies on Galileo MGEX products are already available as listed above. However, most of these are focus on Galileo-IOV satellites or only for individual AC. Therefore, this article evaluates the quality of Galileo orbits using SLR data, PHMs clock-offsets daily linear fitting and PPP applications. The orbit and clock characteristics of five individual ACs are provided. Different SRP models and their impact on orbit quality of different types of Galileo spacecraft are analyzed. Finally, the results are summarized and discussed.

2 Solar Radiation Pressure Models

The SRP is the largest non-conservative perturbation, which directly restricts the POD and orbit prediction. The SRP modeling arises with the development of the GPS system, which can mainly divided into empirical type, analytical type and semi-empirical type. The empirical model is widely used because it does not require the detailed structural information of the new satellite. Among them, the reduced ECOM1 model [10] with five empirical acceleration parameters has the best performance. These parameters can be estimated in a Sun-fixed orthogonal frame, as shown in Eqs. (1) and (2):

$$\vec{a}_{SRP} = \vec{a}_0 + D(u) \cdot \vec{e}_D + Y(u) \cdot \vec{e}_Y + B(u) \cdot \vec{e}_B \tag{1}$$

$$\begin{aligned} D(u) &= D_0 \\ Y(u) &= Y_0 \\ B(u) &= B_0 + B_C \cdot \cos\mu + B_S \cdot \sin\mu \end{aligned} \tag{2}$$

where $\vec{a}_0$ is a priori SRP model; $\vec{e}_D$ points along the satellite-Sun direction; $\vec{e}_Y$ points along the solar panel axis; $\vec{e}_B$ completes the orthogonal frame; the main argument of the model is μ, the argument of latitude of the satellite; D_0, Y_0, B_0, B_C, B_S are parameters to be estimated. At the beginning, due to limited knowledge about the attitude control, dimensions and the optical properties of the satellite surfaces, and the lack of processing experience, this SRP model is widely used within MGEX ACs.

Steigenberger et al. firstly evaluated and analyzed the Galileo orbit and clock products provided by the MGEX ACs, and found that there was systematic errors, whose amplitudes vary as a function of the elevation angle of the Sun above the satellite orbital planes [5]. Montenbruck et al. [9] point out that this systematic error is due to the ECOM1 modeling deficiencies, and established a CBW prior model based on the ECOM1 parameters and clock residuals of the Galileo IOV satellites.

The CBW is parameterized by a cubic-like parameter a_c and a stretching parameter a_s and depends only on the sun-satellite-earth angle γ:

$$\begin{aligned}\vec{a}_D &= -a_C \cdot \left(|\cos\gamma| + \sin\gamma + \frac{2}{3}\right) - a_S \cdot \left(|\cos\gamma| - \sin\gamma - \frac{4}{3}\sin^2\gamma + \frac{2}{3}\right) \\ \vec{a}_B &= -\frac{4}{3} a_S \cdot (\cos\gamma \cdot \sin\gamma)\end{aligned} \tag{3}$$

where coefficients $a_c = 14.5$ nm/s^2 and $a_s = 5.0$ nm/s^2, are given by Montenbruck et al. [9]. When Galileo POD processing, $\vec{a}_0$ in formula (1) could be replaced by formula (3), and those remaining deficiencies are eliminated by the 5 parameter ECOM1 model.

Since classical ECOM1 model introduces spurious signals into orbit and clock products of the elongated satellite bodies, such as GLONASS-M or Galileo. Arnold et al. [9] developed the ECOM2 model for a better absorption of the SRP. The ECOM2 model adds second- and fourth-order harmonic terms to the D-component of the ECOM1 model:

$$\begin{aligned}D(\Delta\mu) &= D_0 + D_{C2} \cdot \cos 2\Delta\mu + D_{S2} \cdot \cos 2\Delta\mu + D_{C4} \cdot \cos 4\Delta\mu + D_{S4} \cdot \cos 4\Delta\mu \\ Y(\Delta\mu) &= Y_0 \\ B(\Delta\mu) &= B_0 + B_C \cdot \cos\Delta\mu + B_S \cdot \sin\Delta\mu\end{aligned} \tag{4}$$

where D_{C2}, D_{S2}, D_{C4}, D_{S4} are four added parameters on D-component, other parameters are same as ECOM1 model. Instead of the argument of latitude μ, the angular argument $\Delta\mu = \mu - \mu_S$ is used, as shown in Fig. 1, where μ_S stands for the argument of latitude of the sun projected on the satellite's orbital plane. Prange et al. [6] applied this new model to CODE's contribution to MGEX since January 2015.

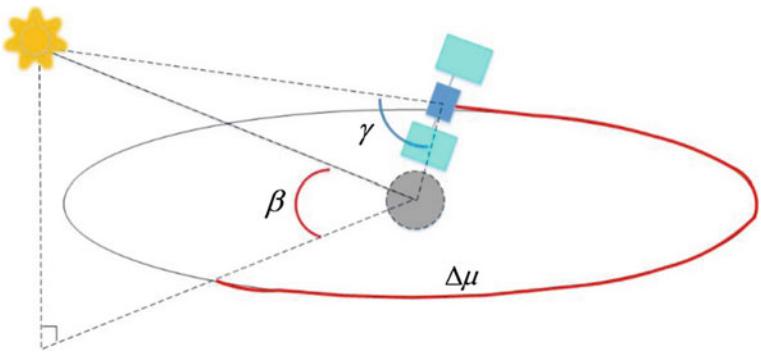


Fig. 1 Geometrical relationship of the satellite, earth and sun

Currently, five MGEX ACs provide precise Galileo orbit and clock products. The CODE, GFZ, WHU provide full multi-GNSS solutions. CNES/CLS provide GPS/GLONASS/Galileo, and TUM only provide Galileo/QZSS products. The most important characteristics of the Galileo processing strategies of above five ACs are summarized in Table 1.

The MGEX ID stands for a three-character abbreviation that is used as prefix for the filenames of the MGEX products. It will be also used as identifier for the individual AC products within this article. All ACs process Galileo data together with GPS or GPS + GLONASS, the ionosphere-free linear combination of E1/E5a are used. CODE use double-differenced data for computation of the satellite orbits and kept then fixed when solving for satellite clocks based on un-differenced data. All other ACs use un-differenced data. GFZ and CNES/CLS process GPS, GLONASS (or together with BeiDou) and Galileo data in one step, TUM and WHU employ a two-step approach: station coordinates, troposphere parameters, and receiver clock offsets are estimated from GPS (or together with GLONASS) date; these above parameters are kept fixed in the second step solving for Galileo orbit and clock products. TUM used CODE's rapid products, and WHU used their own final GPS/GLONASS core products of the IGS. CODE use 3 days orbit solution, and TUM use 5 days orbit solution. CNES/CLS estimating orbits and clocks in a data interval of 30 h, GFZ and WHU solving orbits and clocks in daily solution. CODE used the ECOM2 model for the SRP, and GFZ and TUM used the reduced ECOM1 model, WHU used the reduced ECOM1 model as well as an additional empirical constant acceleration bias in along-tracking direction with

Table 1 Comparison of the processing strategy among MGEX

	CODE	GFZ	CNES/CLS	TUM	WHU
MGEX ID	COM	GBM	GRM	TUM	WUM
Software	Bernese 5.3	EPOS. P8	CNES POD GINS	Bernese 5.0	PANDA
Frequency	E1/E5a	E1/E5a	E1/E5a	E1/E5a	E1/E5a
Differencing	Double diff. (orbit) Zero diff. (clock)	Zero diff.	Zero diff.	Zero diff.	Zero diff.
Elevation cutoff	3° (orbit) 5° (clock)	7°	12°	5°	7°
Data interval	3 day (orbit) 1 day (clock)	1 day	3 + 24 + 3 h	5 day (orbit) 1 day (clock)	1 day
SRP parameters	D_0, Y_0, B_0, D_{C2}, D_{C4}, D_{S2}, D_{S4}, B_C, B_S	D_0, Y_0, B_0, B_C, B_S	D_0, Y_0, B_C, B_S, D_C, D_S	D_0, Y_0, B_0, B_C, B_S	D_0, Y_0, B_0, B_C, $B_S Along_{emp}$
A priori SRP model	None	CBW	Box-wing	CBW	Unknown
Ambiguities	float	float	fixed	float	fixed

1.0^{-10} m/s^2 constrain. CNES/CLS applying a different set of SRP parameters. No a priori SRP model, no albedo model, and no antenna thrust are used for Galileo orbits generated at CODE. Only CNES/CLS applying a boxing-wing, GFZ and TUM using the CBW since November 2016, but the WHU's priori SRP model is still unknown.

3 Results and Analysis

3.1 SLR Residuals

The quality of Galileo orbit products from above five MGEX ACs are evaluated by using SLR data in this article. We analyze the SLR residuals for the period between January 2016 and June 2017. During this period, GFZ, TUM and WHU have adjusted their POD strategies, respectively. Among them, GFZ and TUM used the ECOM1 model together with the CBW priori model since November 2016, but the specific strategy of WHU is still unknown. Different types of Galileo spacecraft from different ACs are analyzed, and Galileo-FOC2 stands for E14 and E18 in this article. The relationships between the SLR residuals and the γ angle are shown in Fig. 2, the red dots indicate the SLR residuals after POD strategy adjustment. The statistics of the SLR residuals from each AC are summarized in Table 2, and the

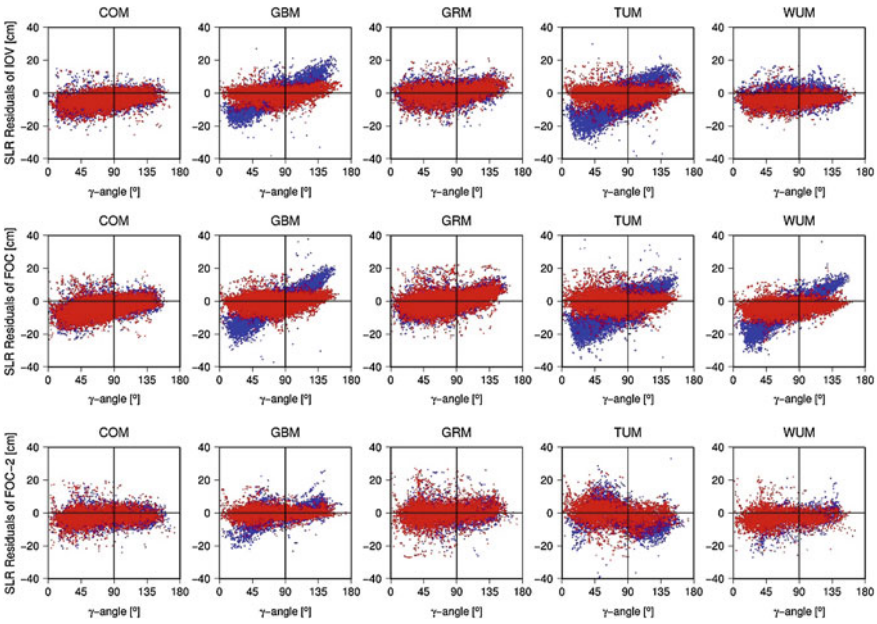


Fig. 2 SLR residuals of Galileo MGEX orbits versus the γ angle

Table 2 Statistics of SLR validation of Galileo MGEX orbits (units: cm)

AC	Galileo-IOV		Galileo-FOC		Galileo-FOC2	
	Offsets	STD	Offsets	STD	Offsets	STD
COM	−4.2	+3.9	−3.1	+3.9	−2.3	+4.3
GBM	−1.1(−0.7)	+7.4(+3.3)	−3.2(−0.4)	+7.9(+3.2)	−1.3(−0.2)	+6.4(+3.1)
GRM	+0.0	+4.1	−0.1	+4.2	+0.0	+5.5
TUM	−5.3(−1.1)	+8.7(+4.0)	−6.4(−0.7)	+9.7(+3.6)	−2.2(−0.4)	+7.4(+5.7)
WUM	−3.2	+3.9	−5.9(−3.4)	+9.1(+3.1)	−3.1	+3.1

values in brackets indicate the results after POD strategy adjustment. From Fig. 2 and Table 2, one can see that: (1) the quality of Galileo orbit from GFZ and TUM is significantly improved if the CBW priori model is used. The SLR residual amplitudes reduced from 20 to 10 cm, the STD has dropped about 50%. (2) When adding the CBW priori model, the offsets of GBM and TUM are reduced from −2 to −6 cm to mm level, which indicates that using of CBW priori model may also eliminate the antenna thrust and earth radiation pressure. (3) The orbit quality of WHU's Galileo-IOV and Galileo-FOC2 satellites do not change when the POD strategies adjustment, while the quality of Galileo-FOC satellite is significantly improved. (4) The systematic errors is mitigated but still existing in the CODE's orbit products. The standard deviations of SLR residuals of ECOM2 is smaller than that of the ECOM, but larger than that of using CBW prior model. (5) There is a mean bias about −2 to −4 cm can be found in COM and WUM orbit products, in contrast, the offsets in GRM is almost zero, which might be attributed to the antenna thrust and earth radiation pressure.

3.2 Fitting Analysis of PHMs

There is a strong correlation between the orbital radial error and the clock parameter, and orbit error will be reflected by Galileo's PHMs. In this article, E30, E26, E19 and E14 satellites from different orbital plane were selected for clock-offsets daily linear fitting, and only the period of on-board PHMs were counted. Figure 3(a–d) are the time series of clock bias from CODE, GFZ, CNES/CLS and WHU, respectively, the dotted line in the figure indicate the absolute value of the elevation angle of the Sun above the satellite orbital planes, and the vertical lines in Fig. 3b and d indicate the day of the POD strategy adjustment. Since the fitting accuracy of TUM's clock-offsets is similar as the results of GBM, it will not be given in this article for the sake of brevity.

From Fig. 3, one can see that: (1) the quality of PHMs clock-offsets from COM and CNES/CLS is always good in the whole arc, varying within 0.2 ns generally. The accuracy of clock products on different orbital planes is not changed significantly. (2) A periodic variation could be observed in GBM's RMS when without

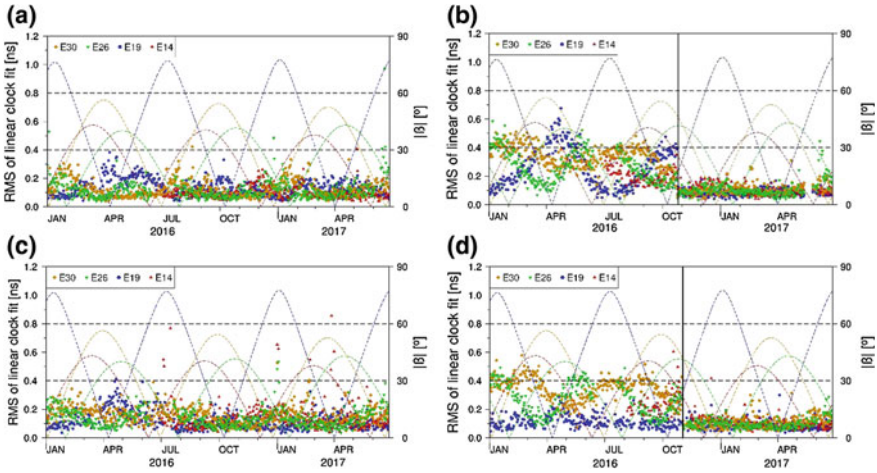


Fig. 3 RMS of daily linear fit for Galileo PHMs clock-offsets

using the CBW model. But the fitting accuracy of the PHMs clock-offsets is about 0.1 ns if the prior CBW model is used. (3) The fitting accuracy of WHU's Galileo-IOV clock-offsets did not change significantly when the POD strategy is adjusted. However, the fitting accuracy of the Galileo-FOC clock-offsets is improved, and the RMS of daily linear fitting is about 0.1 ns. The performance of the PHMs evaluation is consistent with that of the SLR validation.

3.3 PPP Test

We selected 15 stations to analyze the PPP applications for Galileo-only and GPS/Galileo combined solution. The basic information of the stations are listed in Table 3. The analysis period is DOY 145–175 in 2017. The parameters were estimated by sequential least squares adjustment. The sampling interval is 30 s and the elevation cut-off is 7°. The estimation parameters are daily station coordinates, epoch-wise receiver clock parameters, 2-h tropospheric zenith wet delay, float ambiguity parameters, and inter-system biases (ISB) parameters are estimated. The ambiguity parameter is used as a constant in the continuous arc and is re-estimated when the cycle slip occurs. PCOs and PCVs values in igs_absolute.atx are used for GPS. Galileo satellite's PCO values provided by Steigenberger [14] are used in this article, but receiver's PCO and PCV have not been provided, so only satellite's PCO corrections were corrected when processing the Galileo data. Polar tide, solid tide and ocean tide are also considered. The static coordinate solutions are compared with the IGS SNX weekly solutions. The average RMS of all the stations in the analysis period are shown in Table 4. As TUM do not provide GPS products, so only the Galileo-only PPP are analyzed. Taking JFNG station in Wuhan as an

Table 3 Basic information of selected stations

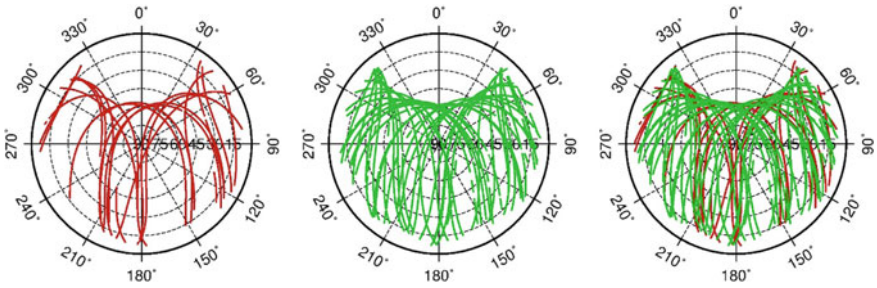
Station	Receiver	Antenna	Radome	Firmware
ALIC	LEICA GR25	LEIAR25.R3	NONE	4.11.606/6
CAS1	TRIMBLE NETR9	LEIAR25.R3	LEIT	5.22
CHOF	JAVAD TRE_G3TH DELTA	TRM57971.00	NONE	3.6.9
CKIS	TRIMBLE NETR9	TRM59800.00	NONE	5.22
GRAC	LEICA GR25	TRM57971.00	NONE	4.02
JFNG	TRIMBLE NETR9	TRM59800.00	NONE	5.15
KIRU	SEPT POLARX4	SEPCHOKE_MC	SPKE	2.9.5-extre
KZN2	TRIMBLE NETR9	TRM59800.00	SCIS	5.22
MAS1	SEPT POLARX4	LEIAR25.R4	NONE	2.9.5-extre
MGUE	SEPT POLARX4	LEIAR25.R4	NONE	2.9.5-extre
NICO	LEICA GR25	LEIAR25.R4	LEIT	4.11/6.523
REYK	LEICA GR25	LEIAR25.R4	LEIT	3.11.1639/6
SUTM	JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T	NONE	3.6.7
UCAL	TRIMBLE NETR9	TRM57971.00	NONE	5.22
ULAB	JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T	NONE	3.6.7

Table 4 PPP results of different analysis center (units: cm)

	GPS				Galileo				GPS/Galileo			
	E	N	U	3D	E	N	U	3D	E	N	U	3D
COM	0.82	0.56	1.12	1.49	0.96	0.77	3.44	3.65	0.72	0.55	1.55	1.80
GBM	0.54	0.44	0.99	1.21	0.62	0.40	3.51	3.59	0.48	0.33	1.31	1.43
GRM	0.53	0.39	0.83	1.06	0.76	0.61	3.38	3.52	0.55	0.37	1.22	1.38
TUM	–	–	–	–	1.51	0.89	3.28	3.71	–	–	–	–
WUM	0.45	0.42	0.82	1.03	0.55	0.70	3.48	3.59	0.44	0.43	1.21	1.36

example, the sky plots of Galileo-only, GPS-only and GPS/Galileo combined solution is shown in Figs. 4 and 5 shows a positioning residuals series calculated by GFZ's orbit and clock products.

From Table 3 and Figs. 4 and 5, one can see that: (1) the positioning error of GPS-only is within 1 cm in North and East component, but slightly larger in vertical component; 3D positioning is accuracy about 1 cm, and COM's positioning accuracy is the lowest. (2) Positioning error of Galileo-only is about 1.0 cm in horizontal, which is very close to the results of GPS-only, but larger than 3.0 cm in vertical component due to the limited number of active satellites. (3) The convergence time of GPS/Galileo combination is significantly shortened. Positioning accuracy in North and East component is better than that of GPS-only, but the accuracy in vertical component is reduced due to the error term related to Galileo.



(a) Galileo (b) GPS (c) GPS/GPS

Fig. 4 Sky plots (azimuth vs. elevation) for the various satellite systems at JFNG

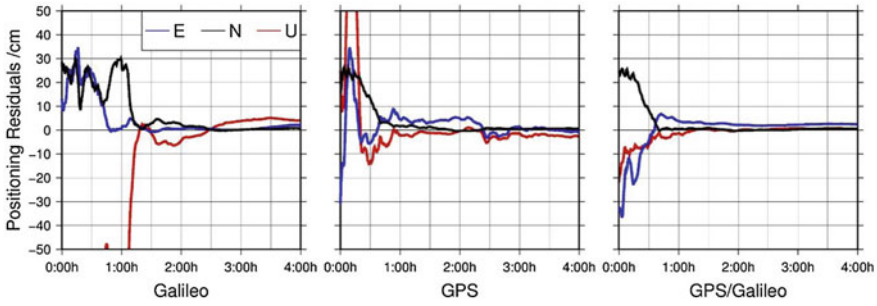


Fig. 5 Positioning residuals of static PPP at JFNG

4 Conclusions

The multi-GNSS not only helps users to obtain more stable and reliable navigation, positioning, timing services, but also contributes to the application in high-precision geodetic fields. But the premise is to provide high precision satellite orbit and clock products under a unified space-time reference system. The orbit quality and PPP application of the Galileo precise products from five MGEX ACs are analyzed in this article, the influence of the SRP model on POD is studied.

The quality of Galileo orbit from GFZ and TUM is significantly improved if the CBW priori model is used. The SLR residual amplitudes reduced from 20 to 10 cm, the STD has dropped about 50%. When adding the CBW priori model, the offsets of GBM and TUM are reduced from −2 to −6 cm to mm level, which indicates that using of CBW priori model may also eliminate the antenna thrust and earth radiation pressure. The orbit quality of WHU's Galileo-IOV and Galileo-FOC2 satellites do not change when the POD strategies adjustment, while the quality of Galileo-FOC satellite is significantly improved. The systematic errors is mitigated but still existing in the CODE's orbit products. The standard deviations of SLR residuals of ECOM2 is smaller than that of the ECOM, but larger than that of using

CBW prior model. There is a mean bias about −2 to −4 cm can be found in COM and WUM orbit products, in contrast, the offsets in GRM is almost zero, which might be attributed to the antenna thrust and earth radiation pressure. The RMS of clock daily linear fitting indicates that the accuracy of the Galileo clocks is significantly improved if the CBW priori model is used. The accuracy of clock products on different orbital planes is not changed significantly. The result of the PHMs evaluation is consistent with that of the SLR validation. Positioning error of Galileo-only is about 1.0 cm in horizontal, which is very close to the results of GPS-only, but larger than 3.0 cm in vertical component due to the limited number of active satellites. The convergence speed of GPS/Galileo combined solution is faster than any single system, the positioning accuracy in horizontal is improved compared to GPS-only, but a small degradation can be observed in vertical component.

Acknowledgements The IGS MGEX and ILRS are greatly acknowledged for providing the multi-GNSS and SLR tracking data. This work was supported by National Key Research Program of China "Collaborative Precision Positioning Project" (No. 2016YFB0501900).

References

1. https://www.gsc-europa.eu/system-status/Constellation-Information. Accessed 20 Oct 2017
2. Rizos C, Montenbruck O, Weber R et al (2013) The IGS MGEX experiment as a milestone for a comprehensive multi-GNSS service. In: ION Pacific PNT conference. DLR, pp 289–295
3. Montenbruck O, Steigenberger P, Khachikyan R et al (2014) IGS-MGEX: preparing the ground for multi-constellation GNSS science. Inside GNSS 9(1):42–49
4. Hackel S, Steigenberger P, Hugentobler U et al (2015) Galileo orbit determination using combined GNSS and SLR observations. GPS Solutions 19(1):15–25
5. Steigenberger P, Hugentobler U, Loyer S et al (2015) Galileo orbit and clock quality of the IGS multi-GNSS experiment. Adv Space Res 55(1):269–281
6. Prange L, Dach R, Lutz S et al (2015) The CODE MGEX orbit and clock solution. In: IAG Scientific Assembly, pp 1–7
7. Uhlemann M, Gendt G, Ramatschi M et al (2015) GFZ global multi-GNSS network and data processing results. Springer, Berlin
8. Guo J, Xu X, Zhao Q et al (2016) Precise orbit determination for quad-constellation satellites at Wuhan University: strategy, result validation, and comparison. J Geodesy 90(2):1–17
9. Montenbruck O, Steigenberger P, Hugentobler U (2015) Enhanced solar radiation pressure modeling for Galileo satellites. J Geodesy 89(3):283–297
10. Springer TA, Beutler G, Rothacher M (1999) A new solar radiation pressure model for GPS. GPS Solutions 2(3):50–62
11. Arnold D, Meindl M, Beutler G et al (2015) CODE's new solar radiation pressure model for GNSS orbit determination. J Geodesy 89(8):775–791
12. Prange L, Orliac E, Dach R et al (2016) CODE's five-system orbit and clock solution—the challenges of multi-GNSS data analysis. J Geodesy 91(4):1–16
13. Guo F, Li X, Zhang X et al (2016) Assessment of precise orbit and clock products for Galileo, BeiDou, and QZSS from IGS Multi-GNSS experiment (MGEX). GPS Solutions, 1–12
14. Steigenberger P, Montenbruck O (2016) Galileo status: orbits, clocks, and positioning. GPS Solutions, 1–13

Research on the Detection Technology Based on INSGNSS for the Dynamic Positioning Performance of Satellite Navigation Terminals

Pengfei Zhao, Shuguo Pan, Fei Ye, Xueyong Lin, Jie Zhang and Xiang Cao

Abstract The positioning performance of satellite navigation equipment is the foundation and core of each performance. Among them, dynamic positioning performance is one of the most important indexes to evaluate the overall performance of GNSS receivers. The detection of GNSS receiver positioning performance under high dynamic conditions is also the difficult point of international research. In this paper, a high precision and reliability vehicle INS/GNSS combined system is applied as a calibration system to realize the detection of GNSS receiver's dynamic positioning performance and further verify the feasibility and reliability of the technology through the dynamic car test under the real road inspection environment.

Keywords Dynamic positioning · INS/GNSS · Calibration system Detection

1 Introduction

As a basic product of the national economy and the people's livelihood, the production of satellite navigation equipment is increasing by scientific research institutions and navigation enterprises both at home and abroad. The main products are: GNSS receivers, real-time dynamic positioning terminals and car navigation systems etc. They are widely used in engineering surveying, waterway surveying and precise Agriculture, hydrography, mapping, precision timing and geodynamics and other fields. In recent years, the rapid development of satellite navigation technology makes the satellite navigation equipments extend from the conventional

P. Zhao · S. Pan (✉) · F. Ye
School of Instrument Science and Engineering, Southeast University, Nanjing, China
e-mail: psg@seu.edu.cn

X. Lin · J. Zhang · X. Cao
Nanjing Institute of Measurement and Testing Technology, Nanjing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_20

support for single GPS constellation mode to multi-constellation or even full constellation and from static reference positioning applications to dynamic application such as dynamic vehicle-mounted, airborne and shipborne applications, from low-precision navigation applications to high-precision measurement applications. The types, performances, and application modes of the product have been continuously optimized and promoted [1].

The detection methods of the static positioning performance of satellite navigation equipments has been proven. Conventional static detection is usually done at the baseline detection field. The static detection method is simple and easy to operate and can significantly reflect the static positioning performance of satellite navigation products. The dynamic positioning mode of satellite navigation equipments has the characteristics of high positioning accuracy, high data update rate and high dynamic performance and so on. These notable features pose a huge challenge in establishing the calibration for the dynamic positioning performance of satellite navigation equipments [2]. Domestic and foreign scholars have done a lot of related researchs on GNSS dynamic detection, mainly aimed at how to establish a suitable dynamic detection calibration, and put forward different positioning methods and evaluation indicators and designed different systems. But there still exist the deficiencies of the single track, calibration errors and the complexity of data processing etc [3, 4]. Based on the above research, a set of high precision and high reliability INS/GNSS detection calibration is established in this paper to realize the real road inspection of the dynamic positioning performance of satellite navigation equipments.

2 INS/GNSS Calibration System

2.1 Principle Introduction

The advantage of GNSS positioning is that unconstrained divergence will not occur during long time observation. The disadvantage is that the environmental dependence is large. The reflection and occlusion of GNSS signals in the Observational environment greatly reduce the accuracy of GNSS measurement and even lead to unreliable positioning. There are two distinct characteristics of INS compared with GNSS:

1. Divergence with time. The navigation error of INS will appear unconstrained divergence over time;
2. Navigation autonomy. The navigation error of INS mainly comes from the error of measurement device itself in the system, which has nothing to do with the external environment.

This shows that the error characteristics of GNSS and INS are complementary. In addition, the measurement update frequency of GNSS is low, while that of INS up to 1 kHz or higher [5, 6]. Therefore, the combination of GNSS and INS can

build a detection calibration system that is environmentally unaffected and provides a continuous, high-precision, and highly reliable positioning.

2.2 Combination Solution

As the calibration of a vehicle dynamic detection system, the vehicle calibration system is based on deep coupling solution technology. It consists of a highly accurate and reliable SINS measurement unit and a high-precision, multi-mode and multi-frequency, direction-finding GNSS receiver compatible with BDS/GPS/GLONASS. The complement of GNSS and INS can guarantee that when the number of satellites is less than 4, high precision position, speed and attitude information can be obtained by using INS data. It is also ensured that the INS system error will be corrected by using GNSS data when the INS system error becomes larger. It can consistently output high-precision and reliable position information and attitude information in a variety of complex environments such as complete openness (without high buildings and trees), sporadic shading (lush trees), severe obstruction (viaduct bottom) and complete occlusion (tunnel) (Fig. 1).

3 Vehicle Dynamic Detection System

3.1 System Composition

Vehicle dynamic detection system is composed of vehicle calibration system, monitoring display system, GNSS signal receiving and forwarding system, testing

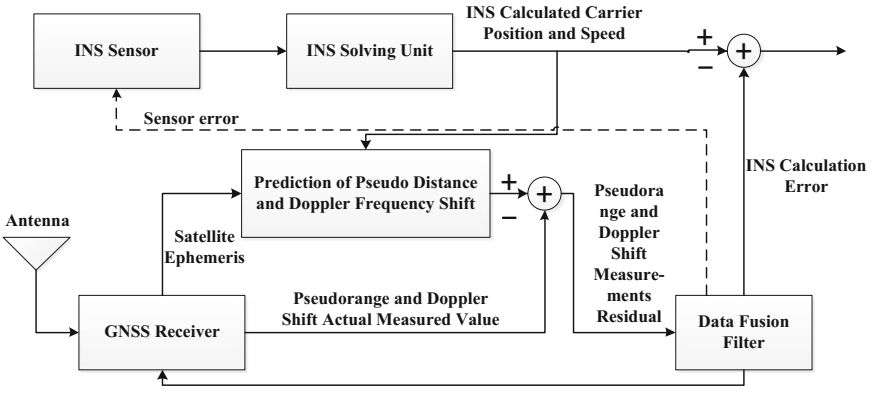


Fig. 1 Deep coupled navigation system

unit for equipments to be checked, power supply system, communication system etc and carries out data exchange with the CORS service center (Fig. 2).

3.2 Data Exchange

There are mainly three types of data flow in vehicle dynamic detection system: GNSS data flow, INS data flow and data flow of equipments uploaded to CORS.

The specific process of data exchange is as follows:

1. Satellite signal homologous transmission: in order to ensure that the received signal of the satellite navigation and positioning terminal to be detected is homologous with that of the vehicle calibration system, a quasi-zero baseline solution is adopted for receiving the same antenna forwarded signal for them. The receiving antenna receives the satellite signal and transmits the satellite signal to the reference system and the satellite navigation positioning terminal to be tested in real time through the GNSS signal forwarding system.
2. Data exchange in the calibration system: while the base GNSS device receives the satellite signal, INS transmits its data to the reference GNSS device via the physical interface for real-time combinatorial computation of the deep coupled solution module of the base GNSS device.

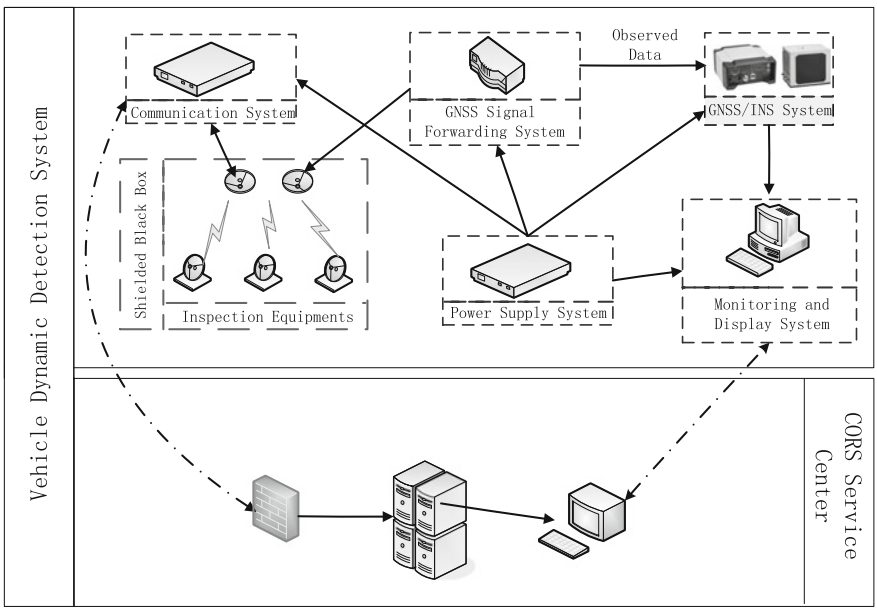


Fig. 2 Vehicle dynamic detection system

3. Access to CORS for the calibration system: the vehicle calibration system login to the CORS service center to upload its own positioning of the GGA information through the 4G network, and real-time receive the corresponding differential correction information for high-precision RTK positioning.
4. Access to CORS for the device to be detected: After receiving the satellite signal, the inspection equipment will log on to the CORS service center through the WIFI network or self-provided other network to upload its own GGA location information, and get the corresponding differential correction information in real time for achieving high-accuracy RTK positioning (Fig. 3).

4 Dynamic Car Test

The positioning accuracy test of the vehicle calibration system refers to that in the process of driving, on the one hand, high-precision positioning solution of calibration system is real-time stored to achieve dynamic detection for RTD; on the other hand, real-time GNSS original observations, INS original data and CORS base station is collected and applied to high-precision GNSS+INS commercial

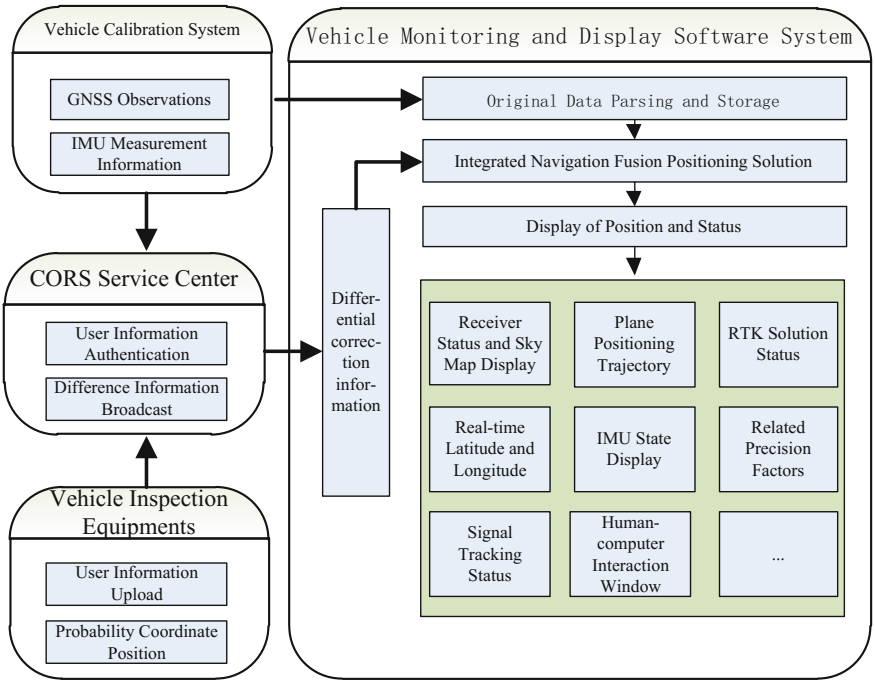


Fig. 3 Data interaction flow chart

post-processing software business to obtain high-precision positioning solution for RTK dynamic detection. Finally, through the real road inspection on the dynamic positioning performance of the high-precision GNSS positioning terminal, the feasibility and reliability of the technology are further verified.

4.1 Centimeter-Level Accuracy for RTD Detection

In the actual measurement process, the calibration system can output centimeter-level high-precision positioning information in real time and support real-time storage. By comparing the real-time positioning solution between the calibration system and the satellite navigation and positioning terminal to be detected, dynamic detection can be realized for RTD positioning terminals.

As can be seen from Fig. 4, in the aspect of the real-time position accuracy of the reference system, the N, E, U errors in the three directions is within 2 cm in most cases, which meets dynamic detection requirements for RTD positioning terminals.

4.2 Millimeter-Level Accuracy for RTK Detection

In the post-processing process, the calibration system can achieve millimeter-level high-precision positioning solution. By comparing the post-positioning solution of

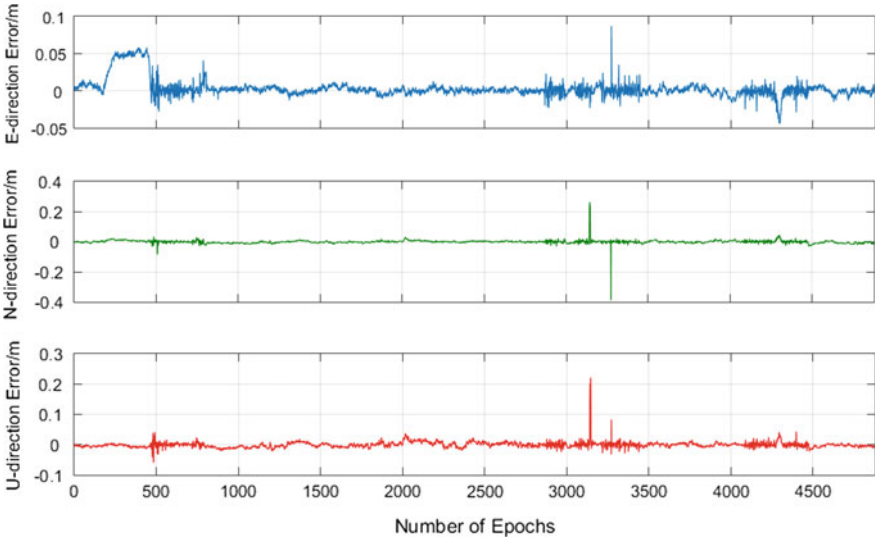


Fig. 4 Real-time position error of calibration system

the calibration system with the real-time positioning solution of the satellite navigation and positioning terminal to be detected, dynamic detection can be realized for RTK positioning terminals.

As can be seen from Fig. 5, in the aspect of the post-position accuracy of the reference device, errors of the N direction and E direction error are within 1 cm, and the error of the U direction within 1.5 cm, in most cases less than 1 cm. It meets the dynamic detection claim for RTK positioning terminals.

4.3 Dynamic Positioning Performance Detection of GNSS Terminals

A more extensive test environment is selected for dynamic car detection. The coordinates of the points of the satellite navigation and positioning terminals to be detected is stored in the hand book at the frequency of 1 Hz. After the test is completed, the dynamic positioning result file of the device under test is exported from the handbook. This test adopts the statistical method of external coincidence accuracy. The measured value of the vehicle calibration system is taken as a true value to calculate the error of each point of the satellite navigation and positioning terminal to be tested. And the external conforming accuracy is counted. The specific statistical results are shown in Fig. 6.

As can be seen from Fig. 6, in most cases, the positioning errors in the N direction and the E direction are within 0.02 m, and the positioning error in the U direction is within 0.06 m. According to the statistics, the positioning accuracy of the satellite navigation and positioning terminals to be detected is 0.012 m in plane and 0.012 m in height, which is in line with the actual situation and further verifies that the vehicle dynamic detection technology based on INS/GNSS is feasible and reliable.

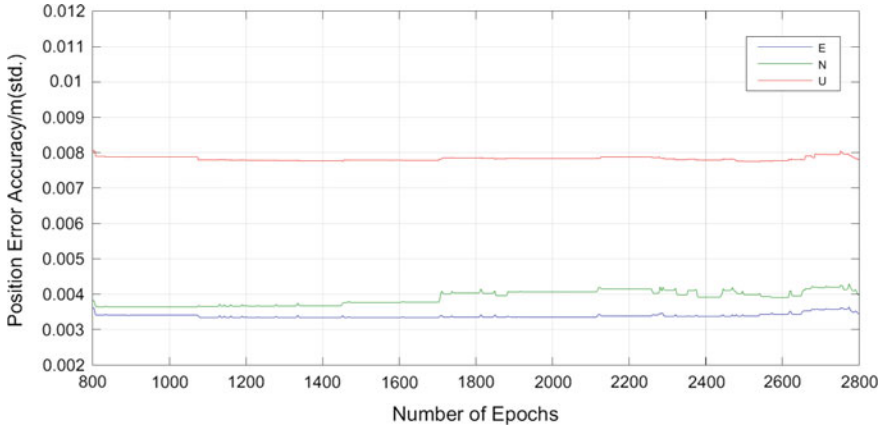


Fig. 5 Post position error of calibration system

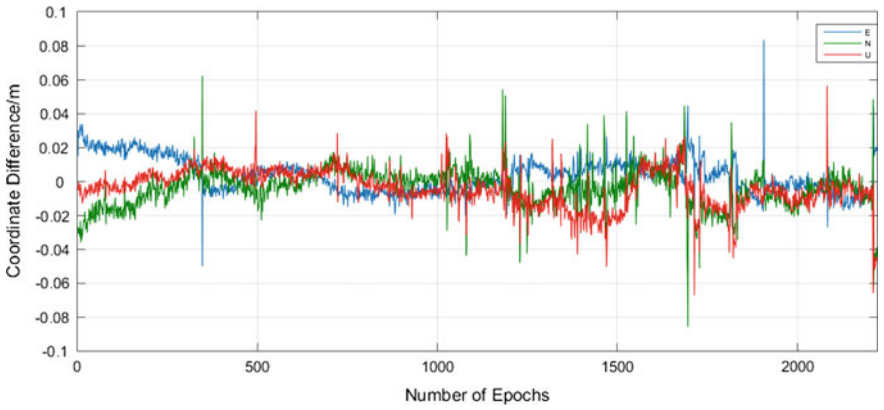


Fig. 6 Dynamic positioning performance detection of terminals

5 Conclusion and Prospect

This paper presents a new real-time dynamic detection technology for RTK terminals, and design a set of vehicle dynamic detection system. A quasi-zero baseline program in which the vehicle calibration system and the satellite navigation and positioning terminals to be detected receive the same antenna forwarding signal is adopted to assure signal homology and the credibility of the detection. In addition, the vehicle equipments seamlessly access CORS for real-time upload and retrieval of information to realize real-time dynamic differential positioning and high-precision positioning results. Most importantly, the sophisticated INS/GNSS deep coupling technology can ensrue that the calibration system realizes the high-precision, highly reliable continuous positioning from environmental influences. All of these provide a guarantee for the dynamic positioning performance detection of satellite navigation and positioning terminals.

Although the vehicle dynamic detection system has been set up. But the space-time synchronization errors still exist. They mainly includes the factors such as antenna phase center deviation, multipath effect, signal forwarding delay and terminal clock error (including clock jitter) etc. In this paper, these factors are not taken into account. Therefore, the suppression and elimination of the errors of the entire vehicle detection system still need to be further studied and improved.

Acknowledgements This work is partially supported by the National Natural Science Foundation of China (Grant No. 41574026, 41774027), the Primary Research & Development Plan of Jiangsu Province (Grant No. BE2016176), the National Key Technologies R&D Program (Grant No. 2016YFB0502101) and the Six Talent Peaks Project in Jiangsu Province (Grant No. 2015-WLW-002) and AQSIQ Science and Technology Plan (Grant No. 2014QK089).

References

1. Cong D, Xu Q (2015) The necessity and preliminary assumption of establishing GNSS dynamic positioning calibration. Satell Internet 04:64–67
2. Cong D, Xu Q (2016) Research status and advances of GNSS accuracy evaluation methods for kinematic positioning. J Navig Positioning. https://doi.org/10.16547/j.cnki.10-1096.20160201
3. Mao Q (2016) Research on real-time dynamic technologies of vehicle GNSS positioning. Southeast University, Dhaka. http://kns.cnki.net/kns/detail/detail.aspx?FileName=1016329062.nh&DbName=CMFD2017
4. Mao Q, Pan S, Liu H, Gao W (2015) Research on real-time dynamic detecting technologies of RTK terminal equipment. Eng Surveying Mapp 24(03):47–52
5. Xu Z (2014) Research on key technologies of integration of GPS/INS for land vehicle applications. Southeast University, Dhaka. http://kns.cnki.net/kns/detail/detail.aspx?FileName=1016752934.nh&DbName=CDFD2016
6. Han S (2010) Novel GPS/SINS integration architecture and systematic error compensation methods. National University of Defense Technology, China. http://kns.cnki.net/kns/detail/detail.aspx?FileName=1011073920.nh&DbName=CDFD2011

Accuracy Analysis of NavIC Pseudo Range Point Positioning

Liang Zhang, Pengfei Sun, Haichun Wang, Miaomiao Liu and Ding Li

Abstract Navigation with Indian Constellation (NavIC) is a regional navigation system built by India that is mainly used for positioning, navigation and timing services throughout India and the surrounding areas. Using the data of TID1 and KAT1 stations from 2017 160 days to 166 days, the accuracy of single-point location accuracy of NavIC is analyzed. Experimental results show that the accuracy of NavIC is greater than 20 m and discontinuity, the PDOP is greater than 10, the number of visible satellites is less than 5 for the most time in Australia.

Keywords NavIC · PDOP · Positioning accuracy · Pseudo-range single point positioning

1 Introduction

Navigation with Indian Constellation (NavIC) is an autonomous regional navigation satellite system implemented by the Indian Space Research Organization (ISRO) [1]. The system provides real-time positioning, speed and time services throughout the day for many kinds of platform user. It also can providing positioning, navigation and timing services throughout India and surrounding areas [2].

In this paper, the data come from IGS data network. Analysis for precision of pseudo range point positioning.

L. Zhang (✉) · P. Sun · H. Wang · M. Liu · D. Li
Chang'an University, Xian 710064, China
e-mail: 2017126054@chd.edu.cn

P. Sun
e-mail: 2017226025@chd.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_21

2 NavIC System Overview

NavIC launch the first satellite in July 2013 and the seventh satellite in April 2016, have successfully completed the seven-satellite network so that can operate independently. The NavIC system includes three GEO satellites and four IGSO satellites, as shown in Fig. 1. The orbital distribution to ensure that satellites can be tracked by at least one ground station at any given point in any time. It can be guaranteed to see most satellites at any point in India [3].

The NavIC system, like the GPS system, also offers two types of services: "Standard Services SPS" and "Limited/Authorized Services RS". The former is provided to all users for free with an accuracy of 20 m. The latter is only available to authorized users, the accuracy better than 10 m [3]. Fig. 2 shows the NavIC constellation PDOP situation. It can be seen that NavIC has a PDOP about 3 in India and surrounding areas. In most parts of Australia, PDOP are greater than 10.

NavIC system broadcast S-band signals and L5-band signals. S signal and L5 signal respectively contain standard service SPS signal, limited/authorized service RS data signal and limited/authorized service RS pilot signal and others signals. In addition, the NavIC has developed a CDMA precision ranging technology to achieve rapid and precise determination of the distance and satellite orbit [4].

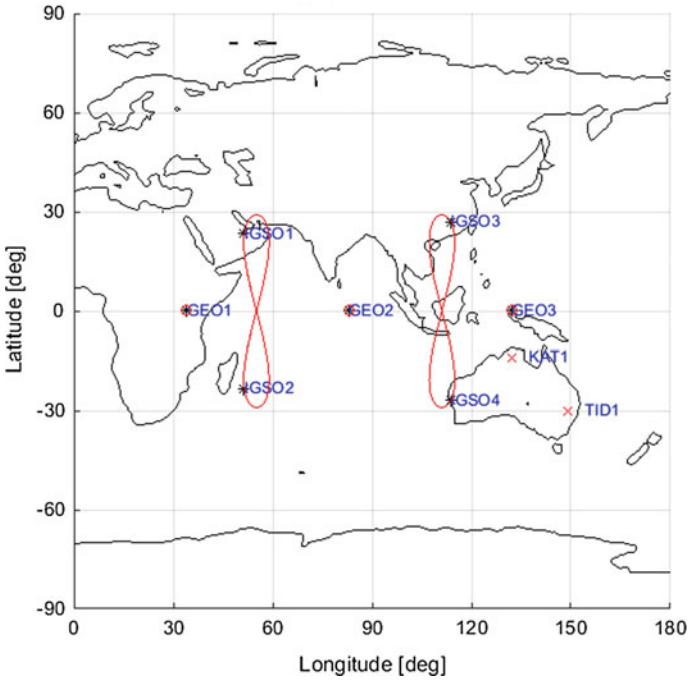


Fig. 1 Ground track of NavIC

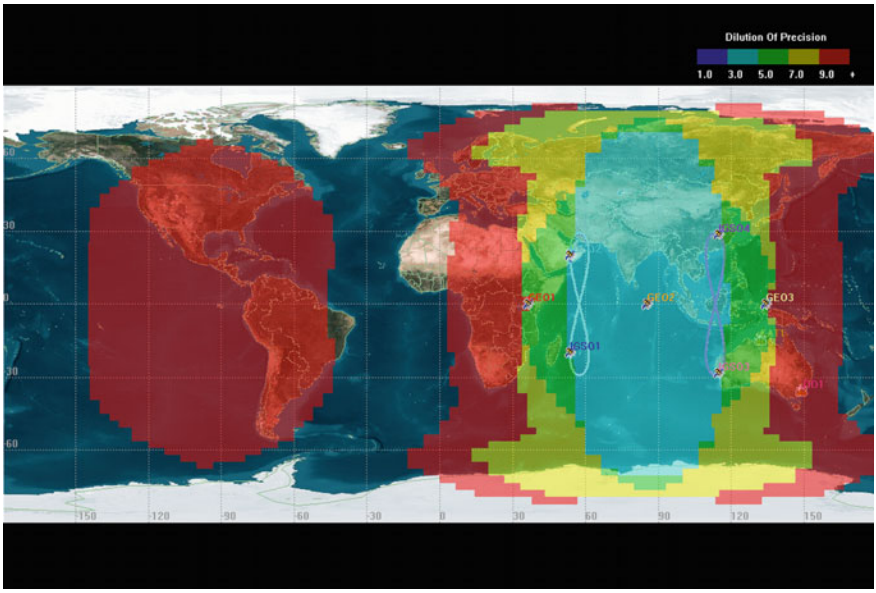


Fig. 2 PDOP of NavIC system

In coordinate and time bases, the NavIC system maintain a high level of compatibility and interoperability with GPS. The WGS-84 coordinate system was used by NavIC; The NavIC system time starts at August 22, 1999 (00:00 UT, August 21–22, midnight) corresponds to 23:59:47 UTC on August 21, 1999.

3 NavIC Pseudorange Point Positioning Accuracy

3.1 Data Acquisition

At present, the collected NavIC data are single-frequency L5 data in Australia. In this paper, TID1, KAT1 two stations (Fig. 1) were selected, time span about 7 days (2017 160 days to 166 days), sampling interval of 30 s.

3.2 Data Processing Strategy

The corrections include: satellite relativistic clock, tropospheric delay, ionospheric delay, relativistic effects, earth rotation correction. The estimated parameters are single epoch station coordinates (Table 1).

Table 1 Software parameters setting

Parameters	Setting
Observations	L5
Interval	30 s
Cut-off elevation angle	10°
Earth rotation parameters	IERS
Troposphere model	Saastamoinen
Ionospheric model	Klobuchar

The exact coordinates of these two stations are provided by the IGS Weekly Solution documents, within mm. Compare the accuracy of NavIC pseudorange point positioning with it.

In service performance evaluation, the accuracy describes the difference between the calculated position of the point and the actual position of the point during the positioning process. It is the most directly performance evaluation of the user's requirement. Through the comparison with the known coordinates, the positioning performance of the NavIC user receiver can be verified. Through the statistical analysis of the positioning accuracy of multiple receivers located in the coverage area, the service performance of the NavIC can be judged. That positioning accuracy can be described using statistical methods [5]. The accuracy of measurement results can be expressed in terms of RMS error [6]. The RMS error is calculated as:

$$\begin{bmatrix} \Delta x_{rms} \\ \Delta y_{rms} \\ \Delta z_{rms} \end{bmatrix} = \begin{bmatrix} \sqrt{\sum_{k=1}^{n} (x_k - \tilde{x})^2/(n-1)} \\ \sqrt{\sum_{k=1}^{n} (y_k - \tilde{y})^2/(n-1)} \\ \sqrt{\sum_{k=1}^{n} (z_k - \tilde{z})^2/(n-1)} \end{bmatrix}$$

3.3 Positioning Precision Analysis

3.3.1 Visible Satellite Number and DOP Value

To analyze NavIC positioning accuracy, first calculate the number of visible satellites (Figs. 3 and 8) and PDOP (Figs. 4 and 9) at TID1 and KAT1 stations. It can be seen that the number of visible satellites in TID1 station is 2–4 and the number of KAT1 stations is 4–5. The PDOP values of both stations are large and discontinuous. When the number of visible satellites reaches 5 in KAT1 station, the PDOP value better than 10, the other times are greater than 10.

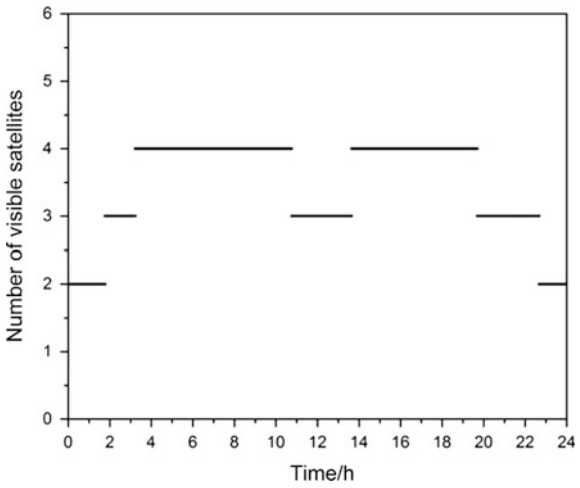


Fig. 3 Satellites of TID1 station in the 165 day

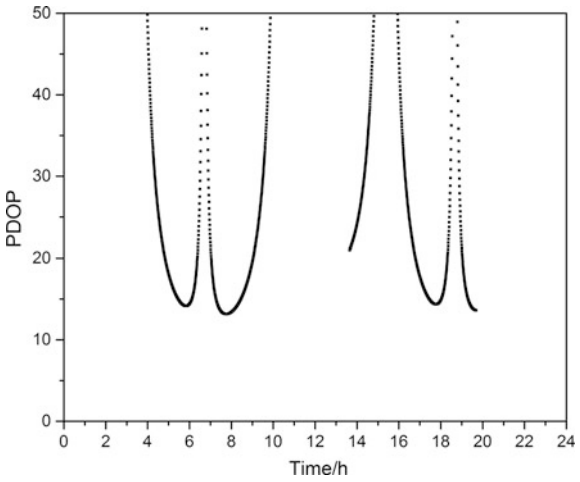


Fig. 4 PDOP of TID1 station in 165 day

3.3.2 Positioning Results

In view of the large and discontinuous PDOP values of the TID1 station and the KAT1 station, we use the PDOP value of less than 50 to conduct the positioning result analysis. Figures 5, 6 and 7 show the positioning errors of the TID1 station in three directions of the E, N and U (Figs. 8 and 9). Figures 10, 11 and 12 show the positioning errors of the KAT1 station in the three directions of the E, N and U. The scattered points of different colors indicate different days.

Table 2 shows the one-week RMS averages of the TID1 and KAT1 stations in the E, N and U directions. It can be seen TID1 station and KAT1 station positioning accuracy more than 20 m and not continuous.

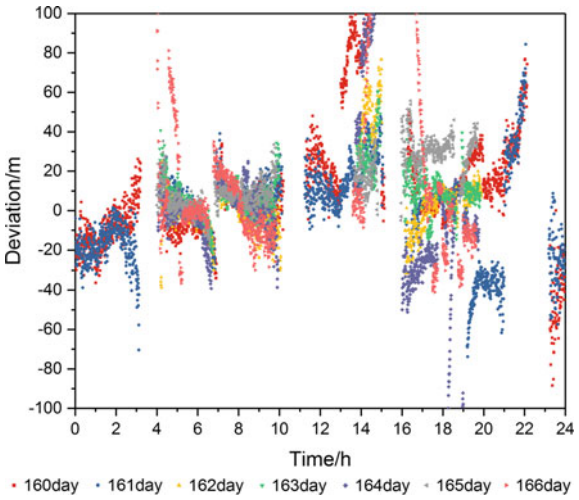


Fig. 5 Deviation sequence of TID1 station in the E direction

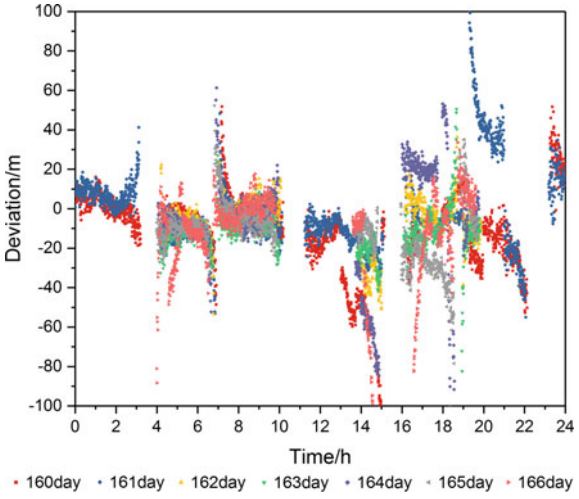


Fig. 6 Deviation sequence of TID1 station in the N direction

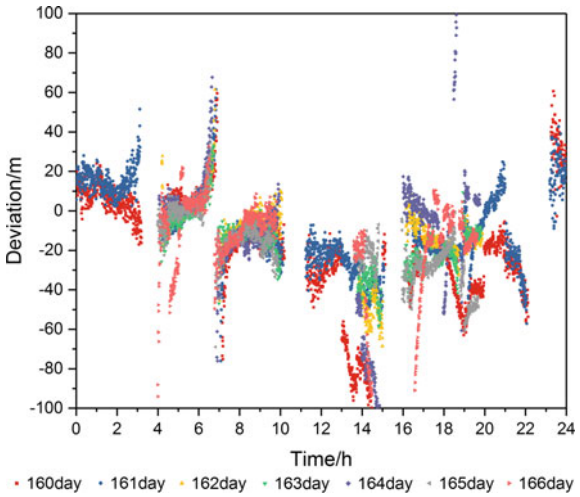


Fig. 7 Deviation sequence of TID1 station in the U direction

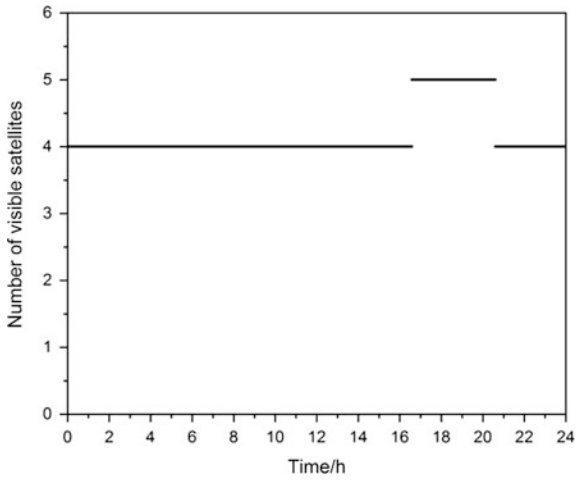


Fig. 8 Satellites of KAT1 station in 165 day

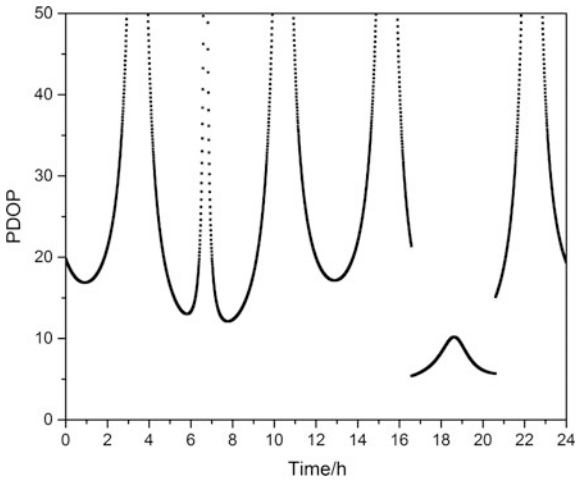


Fig. 9 PDOP of KAT1 station in 165 day

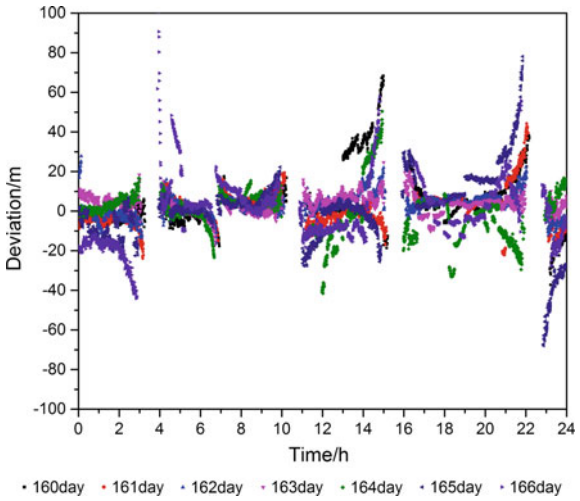


Fig. 10 Deviation sequence of KAT1 station in the E direction

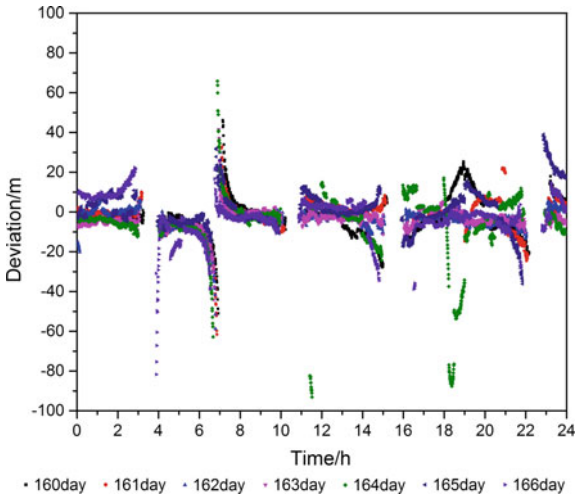


Fig. 11 Deviation sequence of KAT1 station in the N direction

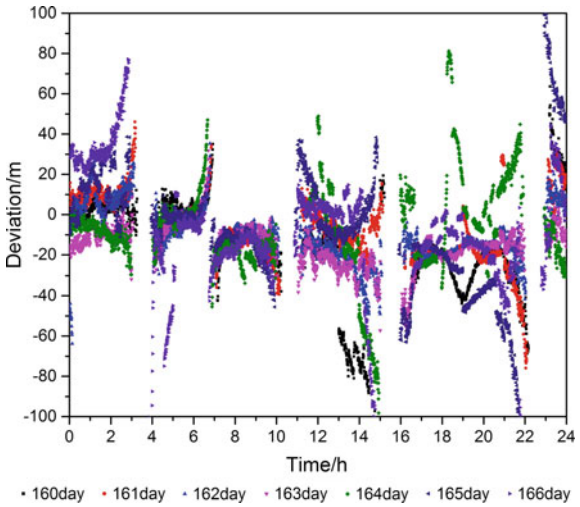


Fig. 12 Deviation sequence of KAT1 station in the U direction

Table 2 RMS mean statistics for two stations in the ENU direction over a week

Station	E (m)	N (m)	U (m)	P (m)
TID1	23.43	20.09	24.81	39.60
KAT1	11.45	8.96	24.13	28.17

4 Conclusions

The coverage of the NavIC system in Australia is not very good. The number of visible satellites is less than 5 for most of the time. The PDOP value is greater than 10 for most of the time, and the positioning accuracy is more than 20 m and not continuous.

At present, NavIC system data that can be collected is the single-frequency data in Australia. Due to the coverage reason, the NavIC system performance cannot be effectively evaluated. In the next step is to gather more data, especially data of India and surrounding areas, get the location accuracy of NavIC system in different regions.

Acknowledgements This work was supported by National key Research Program of China "Collaborative Precision Positioning Project" (No. 2016YFB0501900).

References

1. Zhang C, Zhao X, Li H (2014) Summary of IRNSS. Telecommun. Eng. 54(02):231–235
2. Gowrisankar D, Kibe SV (2008) India's satellite navigation programme. Hanoi Vietnam space for sustainable development, 10 Dec 2008
3. Wei Y (2016) Rename IRNSS to NavIC itself. Mod Navig 7(04):282
4. Li X (2013) Characteristics Analysis of IRNSS. Satell Netw (08):52–58
5. Li Z (2013) Research on performance monitoring and evaluation of satellite navigation system. Information Engineering University
6. Wan S, Xing C, Li H et al (2010) Research on measurement accuracy and measurement uncertainty. J Geodesy Geodyn 30(b11):78–80

Availability Modeling and Simulation of Satellite Navigation System Based on Integration of PDOP and Reliability Maintainability Supportability

Zhuopeng Yang, Yujun Gou, Yanhua Zhu and Heng Zheng

Abstract Due to the satellite-launcher-station interactive operation, coupling of PDOP, reliability and maintainability and supportability (RMS), an availability modeling and simulation method of a navigation system is proposed by GSPN, Bayesian network and outage analysis. The availability of a per-slot is conformed based on GSPN, considering short scheduled or unscheduled outage of bus and payload,long outages, on-orbit or ground backup, testing time in launching site as well as launching reliability. Taking PDOP under different outages as the relationship of slots and the service availability is conformed based on Bayesian network. The vulnerable slots are identified and different optimization designs are compared. The simulation results have demonstrated that this method can be applied for navigation system design optimization, and satellite-launcher-station quantitative decision.

Keywords Navigation system · PDOP · Designed life · RMS
GSPN · Bayesian network

1 Introduction

A service availability, which is important evidence to construction capability and service quality, is the core indicator of a satellite navigation system. As a composite indicator, the service availability is related to the constellation configuration, back strategy, per-slot availability, outage design, and reliability of launcher, fault detection and reconfiguration. Those factors can be decomposed into PDOP and reliability maintainability supportability (RMS) of satellite- launcher-control segment. Due to the interactive operation of satellites, launchers and control segments, the relationship of PDOP and RMS is extremely complicated. The availability

Z. Yang (✉) · Y. Gou · Y. Zhu · H. Zheng
China Academy of Aerospace Standardization and Product Assurance,
Beijing 100071, China
e-mail: yangzhuopeng88@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_22

modeling and simulation based on integration of PDOP and RMS is significant for the navigation system design, state evaluation, and maneuver strategy.

There exist some interesting papers about availability of a satellite navigation system. A per-slot availability and constellation availability is modeling based on Markov Chain, considering long-term failures, short-term failures and maneuvers [1]. The definition and algorithm of PDOP availability, accuracy availability, availability of integrity and availability of continuity are given in [2]. A per-slot availability, constellation availability and PDOP availability [3–5] are analyzed by considering short-term scheduled (STS), short-term unscheduled (STU), long-term scheduled (LTS) and long-term unscheduled (LTU), or orbital maneuvers, single event upset, and end-of-life disposal. The concept of availability, the relationship of a per-slot availability, constellation availability and PDOP availability are discussed successfully in above papers. On the other hand, the bus and payload faults under different outages, processes of fault detection and reconfiguration, and design ability of satellite- launcher-control segment are researched scarcely.

The concept of satellite navigation system availability is given firstly. Then considering fault detection and reconfiguration and parameters design, the availability based on integration of PDOP and RMS is modeling by generalized stochastic Petri net (SCPN), Bayesian network (BN) and outage analysis, and the vulnerable slots are identified and improved schemes are proposed in this paper.

2 The Navigation System Availability Concept and Process

2.1 Availability Concept

Availability is defined as a measure of the degree to which an item is in an operable and committable state at the start of a mission when the mission is called for at an unknown (random) time. Availability is initially applied to weapon equipment system, then to satellite navigation system. The satellite navigation system availability (A_{NAV}) is divided into SIS availability (A_{SIS}) and service availability ($A_{SERVICE}$) in GPS SPS PS [6].

(1) The SIS availability can be divided into per-slot availability (A_{slot}) and constellation availability ($A_{constellation}$). The per-slot availability is the fraction of time that a slot in the GPS constellation will be occupied by a satellite that is transmitting a trackable and healthy SPS SIS. The constellation availability is the fraction of time that a specified number of slots in the GPS constellation are occupied by satellites that are transmitting a trackable and healthy SPS SIS, and it can be measured as state probability of at least 21/20 Slots out of the 24 Slots.

(2) The service availability is the fraction of time that a service is availability for users. It can be divided into PDOP availability (A_{PDOP}), availability of position ($A_{position}$), availability of timing (A_{timing}), availability of integrity ($A_{integrity}$), and availability of continuity ($A_{continuity}$).

The SIS availability is a manufacturer-facing indicator, and service availability is a user-facing indicator. The classification of availability is given in Fig. 1.

The per-slot availability is a basic unit of SIS availability and service availability, and the PDOP availability is a basic unit of service availability. The per-slot availability and PDOP availability are researched in this paper.

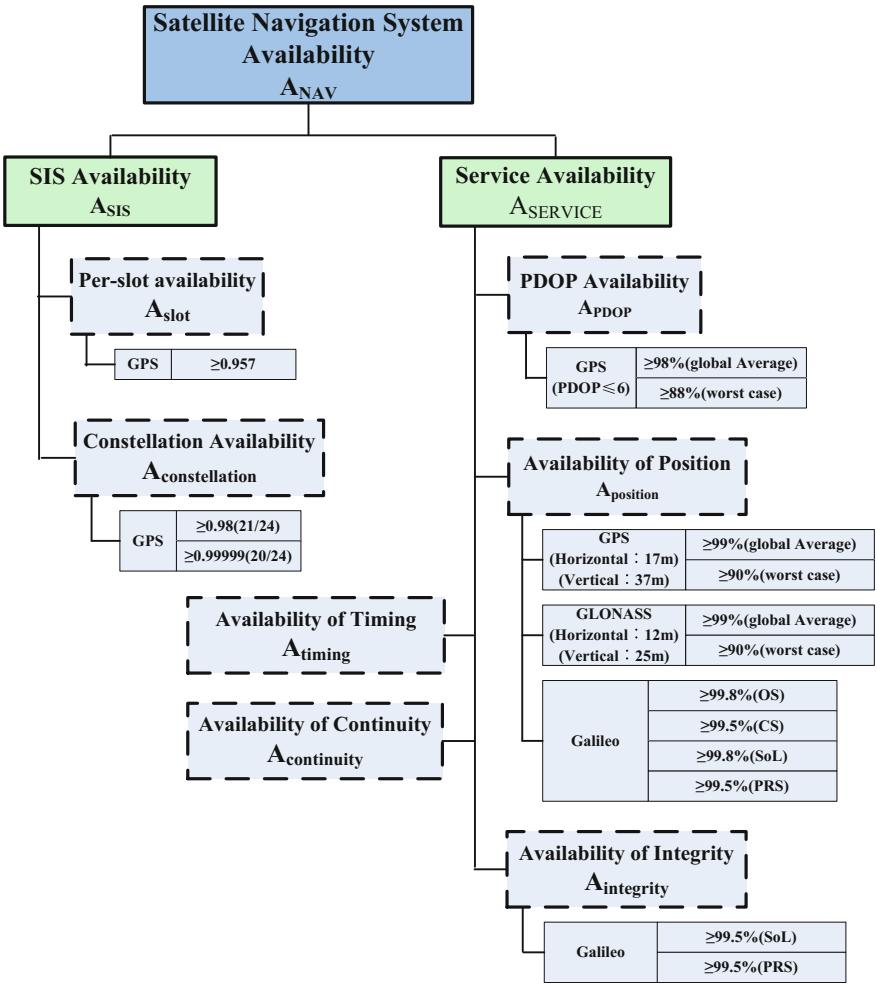


Fig. 1 Availability of satellite navigation system

2.2 Processes of Per-Slot Availability

The per-slot outages are divided into STS, STU, LTS and LTU. Because the LTS is a rare event, LTS and LTU are combined into the LT. Furthermore, bus faults and payload faults are considered respectively due to the different faults ratio and reconfiguration, as Fig. 2.

(1) STS (bus): taking orbital maneuver (OM) and business recovery by control segment.
(2) STS (payload): taking business recovery by control segment.
(3) STU (bus): taking fault detection, working condition recovery and business recovery by control segment.
(4) STU (payload): taking fault detection, strategies research and business recovery based on spares by control segment.
(5) LT: judging on-orbit state, if on-orbit, then taking constellation reconfigure and on-orbit testing, if not, then developing a new satellite or launching a backup satellite, then taking on-orbit testing.

The parameters illustrated in Table 1.

3 Availability Modeling and Analysis

3.1 A Per-Slot Availability Modeling and Analysis

Because of the flow characteristic of fault reconfigure, a per-slot availability ($GSPN_{slot}$) is modeled by GSPN, thus as $GSPN_{STSB}$ (STS(bus)), $GSPN_{STUB}$ (STU(bus)), $GSPN_{STSP}$ (STS(payload)), $GSPN_{STUP}$ (STU(payload)), and $GSPN_{LT}$ (LT):

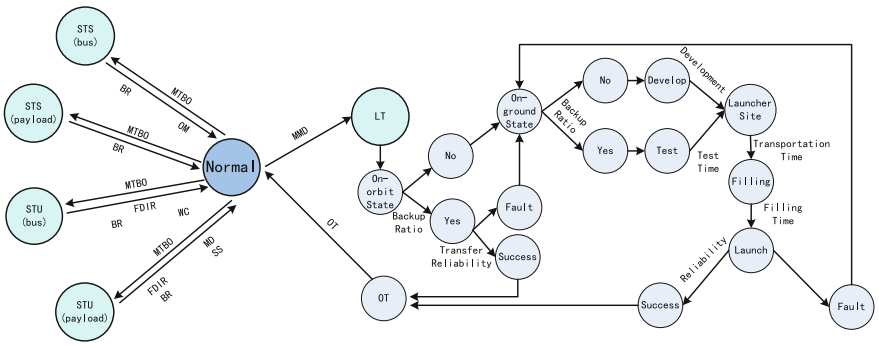


Fig. 2 Operational process of a per-slot availability

Table 1 Parameters of a per-slot availability

Outage	Parameters
STS (bus)	Mean time between outage ($T_{STSB1} = 1/\lambda_{STSB1}$[a])
	Mean orbit moving time ($T_{STSB2} = 1/\lambda_{STSB2}$)
	Mean business recovery ($T_{STSB3} = 1/\lambda_{STSB3}$)
STS (payload)	Mean time between outage ($T_{STSP1} = 1/\lambda_{STSP1}$)
	Mean business recovery ($T_{STSP2} = 1/\lambda_{STSP2}$)
STU (bus)	Mean time between outage ($T_{STUB1} = 1/\lambda_{STUB1}$)
	Mean fault detection time ($T_{STUB2} = 1/\lambda_{STUB2}$)
	Mean working condition recovery time ($T_{STUB3} = 1/\lambda_{STUB3}$)
	Mean business recovery ($T_{STUB4} = 1/\lambda_{STUB4}$)
STU (payload)	Mean time between outage ($T_{STUP1} = 1/\lambda_{STUP1}$)
	Mean fault detection time ($T_{STUP2} = 1/\lambda_{STUP2}$)
	Mean administrative delay time ($T_{STUP3} = 1/\lambda_{STUP3}$)
	Mean short-supply response time ($T_{STUP4} = 1/\lambda_{STUP4}$)
	Mean long-supply response time ($T_{STUP5} = 1/\lambda_{STUP5}$)
	Mean business recovery ($T_{STUP6} = T_{STUP7} = 1/\lambda_{STUP6} = 1/\lambda_{STUP7}$)
	Probability of spares sufficiency ($f_{STUP1} = 1 - f_{STUP2}$[b])
LT	Mean mission during ($T_{LT1} = 1/\lambda_{LT1}$)
	Satellite development time ($T_{LT2} = 1/\lambda_{LT4}$)
	Backup satellite testing time ($T_{LT3} = 1/\lambda_{LT3}$)
	Transportant time ($T_{LT4} = 1/\lambda_{LT5}$)
	Testing and filling time ($T_{LT5} = 1/\lambda_{LT6}$)
	On-orbit testing time ($T_{LT6} = 1/\lambda_{LT7}$)
	Satellite design life (T_{LT7})
	On-orbit ratio ($f_{LT1} = 1 - f_{LT2}$)
	On-ground ratio ($f_{LT3} = 1 - f_{LT4}$)
	Reliability of constellation reconfigure ($f_{LT5} = 1 - f_{LT6}$)
	Reliability of launch ($f_{LT7} = 1 - f_{LT8}$)

[a]Fault rate or Repair Rate of the Parameter, [b]Opposite event Parameter

$$\begin{aligned} GSPN_{slot} &= GSPN_{STS} + GSPN_{STU} + GSPN_{LT} \\ &= GSPN_{STSB} + GSPN_{STSP} + GSPN_{STUB} + GSPN_{STUP} + GSPN_{LT} \end{aligned} \tag{1}$$

(1) $\mathbf{GSPN_{STS}}$

Considering bus faults and payload fault respectively, the STS is modeled, as Fig. 3.

In Fig. 3, P_{STS} is state of un-STS, P_{STSB1} and P_{STSP1} is STS of bus and payload, P_{STSB2} is state after orbit manuver. Corresponding to the continuous time Markov chain:

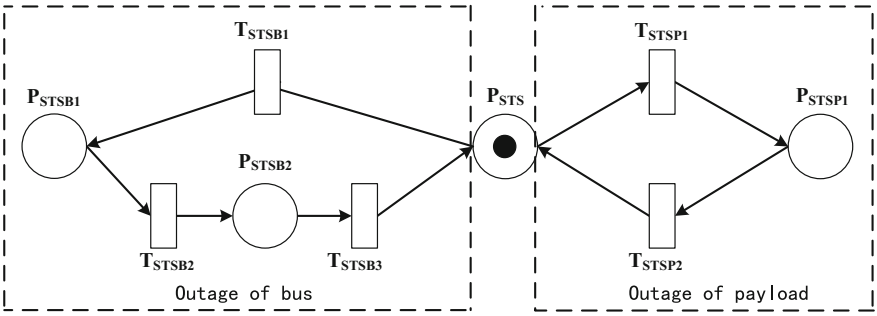


Fig. 3 The model of STS based on GSPN

$$\begin{bmatrix} \dot{P}_{STS}(t) \\ \dot{P}_{STSB1}(t) \\ \dot{P}_{STSB2}(t) \\ \dot{P}_{STSP1}(t) \end{bmatrix} = \begin{bmatrix} -\lambda_{STSB1} - \lambda_{STSP1} & 0 & \lambda_{STSB3} & \lambda_{STSP2} \\ \lambda_{STSB1} & -\lambda_{STSB2} & 0 & 0 \\ 0 & \lambda_{STSB2} & -\lambda_{STSB3} & 0 \\ \lambda_{STSP1} & 0 & 0 & -\lambda_{STSP2} \end{bmatrix} \cdot \begin{bmatrix} P_{STS}(t) \\ P_{STSB1}(t) \\ P_{STSB2}(t) \\ P_{STSP1}(t) \end{bmatrix} \tag{2}$$

$$P_{STS}(t) + P_{STSB1}(t) + P_{STSB2}(t) + P_{STSP1}(t) = 1 \tag{3}$$

and the steady of un-STS:

$$P_{STS} = A_{STS}/(A_{STS} + B_{STS} + C_{STS} + D_{STS}) \tag{4}$$

where: $A_{STS} = \lambda_{STSB2} \cdot \lambda_{STSB3} \cdot \lambda_{STSBP2}$, $B_{STS} = \lambda_{STSB1} \cdot \lambda_{STSB3} \cdot \lambda_{STSBP2}$, $C_{STS} = \lambda_{STSB1} \cdot \lambda_{STSB2} \cdot \lambda_{STSBP2}$, $D_{STS} = \lambda_{STSB2} \cdot \lambda_{STSB3} \cdot \lambda_{STSBP1}$.

(2) **$\mathbf{GSPN_{STU}}$**

Considering fault detection time, administrative delay time, supply response time, business recovery time and probability of spares sufficiency, STU is modeled, as Fig. 4.

In Fig. 4, P_{STU} is state of un-STU, $P_{STUB1}$and $P_{STUP1}$is STU of bus and payload, P_{STUB2} and P_{STUB3} are state after fault detection and working condition recovery, P_{STUP2}, P_{STUP3}, P_{STUP4}, P_{STUP5}, P_{STUP6} and P_{STUP7} are state after fault detection, strategy research, short-supply waiting, long-supply waiting, short-supply response, long-supply response. Corresponding to the continuous time Markov chain:

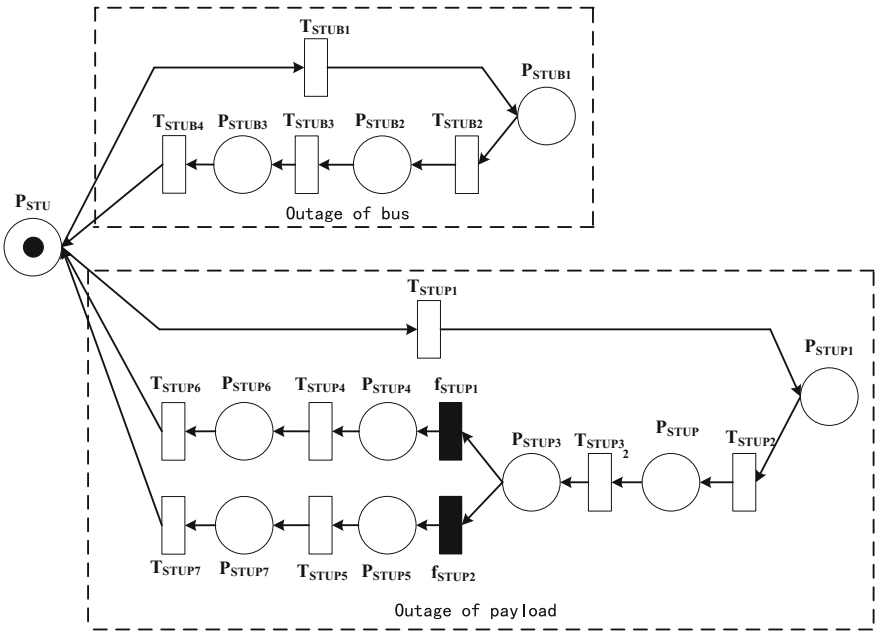


Fig. 4 The model of STU based on GSPN

$$
\begin{aligned}
&\left[\dot{P}_{STU}(t)\quad \dot{P}_{STUB1}(t)\quad \dot{P}_{STUB2}(t)\quad \dot{P}_{STUB3}(t)\quad \dot{P}_{STUP1}(t)\quad \dot{P}_{STUP2}(t)\quad \dot{P}_{STUP4}(t)\quad \dot{P}_{STUP5}(t)\quad \dot{P}_{STUP6}(t)\quad \dot{P}_{STUP7}(t)\right]^{\mathrm{T}}\\
&=\begin{bmatrix}
-\lambda_{STUB1}-\lambda_{STUP1} & 0 & 0 & \lambda_{STUB4} & 0 & 0 & 0 & 0 & \lambda_{STUP6} & \lambda_{STUP7}\\
\lambda_{STUB1} & -\lambda_{STUB2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & \lambda_{STUB2} & -\lambda_{STUB3} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \lambda_{STUB3} & -\lambda_{STUB4} & 0 & 0 & 0 & 0 & 0 & 0\\
\lambda_{STUP1} & 0 & 0 & 0 & -\lambda_{STUP2} & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & \lambda_{STUP2} & -\lambda_{STUP3} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & \lambda_{STUP3}\cdot f_{STUP1} & -\lambda_{STUP4} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & \lambda_{STUP3}\cdot f_{STUP2} & 0 & -\lambda_{STUP5} & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & \lambda_{STUP4} & 0 & -\lambda_{STUP6} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{STUP5} & 0 & -\lambda_{STUP7}
\end{bmatrix}\\
&\cdot\begin{bmatrix}
P_{STU}(t)\\ P_{STUB1}(t)\\ P_{STUB2}(t)\\ P_{STUB3}(t)\\ P_{STUP1}(t)\\ P_{STUP2}(t)\\ P_{STUP4}(t)\\ P_{STUP5}(t)\\ P_{STUP6}(t)\\ P_{STUP7}(t)
\end{bmatrix}
\end{aligned}
\tag{5}
$$

$$P_{STU}(t) + P_{STUB1}(t) + P_{STUB2}(t) + P_{STUB3}(t) + P_{STUP1}(t) + P_{STUP2}(t) + P_{STUP4}(t) + P_{STUP5}(t) + P_{STUP6}(t) + P_{STUP7}(t) = 1 \tag{6}$$

and the steady of un-STU:

$$P_{STU} = A_{STU}/(A_{STU} + B_{STU} + C_{STU} + D_{STU} + E_{STU} + F_{STU} + G_{STU} + H_{STU} + I_{STU} + J_{STU}) \tag{7}$$

where:

$A_{STU} = \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7}$,
$B_{STU} = \lambda_{STUB1} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7}$,
$C_{STU} = \lambda_{STUB1} \cdot \lambda_{STUB2} \cdot \lambda_{STUB4} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7}$,
$D_{STU} = \lambda_{STUB1} \cdot \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7}$,
$E_{STU} = \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP1} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7}$,
$F_{STU} = \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP1} \cdot \lambda_{STUP2} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7}$,
$G_{STU} = \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP1} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7} \cdot f_{STUP1}$, $H_{STU} = \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP1} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP6} \cdot \lambda_{STUP7} \cdot f_{STUP2}$, $I_{STU} = \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP1} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP7} \cdot f_{STUP1}$, $J_{STU} = \lambda_{STUB2} \cdot \lambda_{STUB3} \cdot \lambda_{STUB4} \cdot \lambda_{STUP1} \cdot \lambda_{STUP2} \cdot \lambda_{STUP3} \cdot \lambda_{STUP4} \cdot \lambda_{STUP5} \cdot \lambda_{STUP6} \cdot f_{STUP2}$.

(3) **GSPN$_{LT}$**

Considering mean mission during, satellite development time, transportant time, testing and filling time, on-orbit ratio, reliability of constellation reconfigure, reliability of launch, LT is modelled, as Fig. 5.

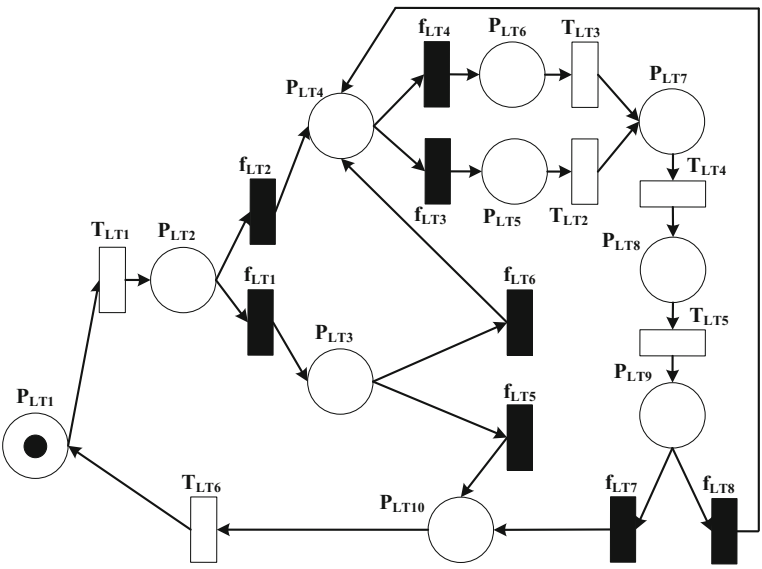


Fig. 5 The model of LT based on GSPN

In Fig. 5, P_{LT1} is state of un-LU, P_{LT2}, P_{LT3}, P_{LT4}, P_{LT5}, P_{LT6}, P_{LT7}, P_{LT8}, P_{LT9} and P_{LT10} are state of LT, on-orbit backup, on-ground backup, testing waiting, development waiting, transportant, testing and filling. Corresponding to the continuous time Markov chain:

$$\begin{bmatrix} \dot{P}_{LT1}(t) & \dot{P}_{LT5}(t) & \dot{P}_{LT6}(t) & \dot{P}_{LT7}(t) & \dot{P}_{LT8}(t) & \dot{P}_{LT10}(t) \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} -\lambda_{LT1} & 0 & 0 & 0 & 0 & \lambda_{LT6} \\ \lambda_{LT1}\cdot f_{LT2}\cdot f_{LT3}+\lambda_{LT1}\cdot f_{LT1}\cdot f_{LT3}\cdot f_{LT6} & -\lambda_{LT2} & 0 & 0 & \lambda_{LT5}\cdot f_{LT3}\cdot f_{LT8} & 0 \\ \lambda_{LT1}\cdot f_{LT2}\cdot f_{LT4}+\lambda_{LT1}\cdot f_{LT1}\cdot f_{LT4}\cdot f_{LT6} & 0 & -\lambda_{LT3} & 0 & \lambda_{LT5}\cdot f_{LT4}\cdot f_{LT8} & 0 \\ 0 & \lambda_{LT2} & \lambda_{LT3} & -\lambda_{LT4} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{LT4} & -\lambda_{LT5} & 0 \\ \lambda_{LT1}\cdot f_{LT1}\cdot f_{LT5} & 0 & 0 & 0 & \lambda_{LT5}\cdot f_{LT7} & -\lambda_{LT6} \end{bmatrix} \cdot \begin{bmatrix} P_{LT1}(t) \\ P_{LT5}(t) \\ P_{LT6}(t) \\ P_{LT7}(t) \\ P_{LT8}(t) \\ P_{LT10}(t) \end{bmatrix} \tag{8}$$

$$P_{LT1}(t)+P_{LT5}(t)+P_{LT6}(t)+P_{LT7}(t)+P_{LT8}(t)+P_{LT10}(t)=1 \tag{9}$$

and the steady of un-LT:

$$P_{LT1}=A_{LT}/(A_{LT}+B_{LT}+C_{LT}+D_{LT}+E_{LT}+F_{LT}+G_{LT}+H_{LT}+I_{LT}+J_{LT}) \tag{10}$$

where: $A_{LT}=\lambda_{LT2}\cdot\lambda_{LT3}\cdot\lambda_{LT4}\cdot\lambda_{LT5}\cdot\lambda_{LT6}\cdot f_{LT7}$, $B_{LT}=\lambda_{LT1}\cdot\lambda_{LT3}\cdot\lambda_{LT4}\cdot\lambda_{LT5}\cdot\lambda_{LT6}\cdot f_{LT2}\cdot f_{LT3}$, $C_{LT}=\lambda_{LT1}\cdot\lambda_{LT3}\cdot\lambda_{LT4}\cdot\lambda_{LT5}\cdot\lambda_{LT6}\cdot f_{LT1}\cdot f_{LT3}\cdot f_{LT6}$, $D_{LT}=\lambda_{LT1}\cdot\lambda_{LT2}\cdot\lambda_{LT4}\cdot\lambda_{LT5}\cdot\lambda_{LT6}\cdot f_{LT2}\cdot f_{LT4}$, $E_{LT}=\lambda_{LT1}\cdot\lambda_{LT2}\cdot\lambda_{LT4}\cdot\lambda_{LT5}\cdot\lambda_{LT6}\cdot f_{LT1}\cdot f_{LT4}\cdot f_{LT6}$, $F_{LT}=\lambda_{LT1}\cdot\lambda_{LT2}\cdot\lambda_{LT3}\cdot\lambda_{LT5}\cdot\lambda_{LT6}\cdot f_{LT2}$, $G_{LT}=\lambda_{LT1}\cdot\lambda_{LT2}\cdot\lambda_{LT3}\cdot\lambda_{LT5}\cdot\lambda_{LT6}\cdot f_{LT1}\cdot f_{LT6}$, $H_{LT}=\lambda_{LT1}\cdot\lambda_{LT2}\cdot\lambda_{LT3}\cdot\lambda_{LT4}\cdot\lambda_{LT6}\cdot f_{LT2}$, $I_{LT}=\lambda_{LT1}\cdot\lambda_{LT2}\cdot\lambda_{LT3}\cdot\lambda_{LT4}\cdot\lambda_{LT1}\cdot f_{LT6}$, $J_{LT}=\lambda_{LT1}\cdot\lambda_{LT2}\cdot\lambda_{LT3}\cdot\lambda_{LT4}\cdot\lambda_{LT5}\cdot f_{LT7}$.

Mean Mission During (T_{LT1}) is a parameter related to satellite reliability and design life. Satellite reliability is a function of random failure and wearout failure:

$$R_{satellite}(t)=e^{-(t/\alpha)^{\beta}}\int_{t}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \tag{11}$$

where: α is scale parameter, β is shape parameters, μ is expectation of satellite design life, σ is standard deviation.

Thus, MMD is:

$$T_{LT1} = \int_0^T R_{satellite}(t)dt = \int_0^T e^{-(t/\alpha)^\beta} \int_t^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x dt \tag{12}$$

where: T is a satellite design life.

A per-slot availability is modeled by STS, STU and LT. Due to the complexity of model, Monte-Carlo is applied.

3.2 Service Availability Model and Analysis

(1) Relationship of satellites

A service availability is a function of per-slot availability and relationship of satellites. PDOP is selected for the relationship of satellites.

$$\alpha_{k,n} = \frac{\sum_{t=t_0}^{t_0+k\cdot\Delta T} \sum_{j=1}^{L} bool(PDOP_{t,j} \le PDOP_{req}) \cdot area_j}{k \cdot \sum_{i=1}^{L} area_i} \tag{13}$$

where: t_0 is start time, ΔT is simulation step time interval, k is the total steps of simulation, $PDOP_{t,j}$ is PDOP of j at t, $PDOP_{req}$ is requirement value, $area_i$ is the i area, L is the total areas. $bool(x)$ is boolean function, when x is true, $bool(x) = 1$, thus $bool(x) = 0$.

Determine PDOP of each per-slot by outage analysis, then the relationship is given.

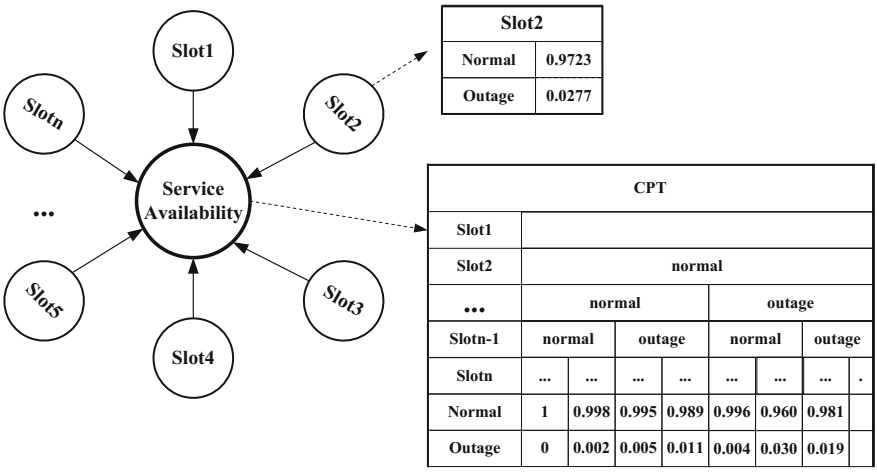
(2) Service Availability model and prediction

Due to the "0 ~ 1" relationship characteristic, Bayesian network is applied, as Fig. 6.

In Fig. 6, the values of root nodes "slot1…n" are availability and outage rate, the input value of central node "service availability" are PDOP under different outages, and the output value of central node is service availability $A_{SERVICE}$:

$$\begin{aligned} A_{SERVICE} &= P(A_{SERVICE} = 1) \\ &= \sum_{A_{sloti}} P(A_{slot1}, A_{slot2}, \ldots, A_{slotn}, A_{SERVICE}) \\ &= \sum_{A_{sloti}} P(A_{SERVICE} = 1 | A_{slot1}, A_{slot2}, \ldots, A_{slotn}) P(A_{slot1}) \ldots P(A_{slotn}) \end{aligned} \tag{14}$$

where: A_{sloti} is the availability of the i slot.



Slot2	
Normal	0.9723
Outage	0.0277

CPT								
Slot1								
Slot2	normal							
...	normal				outage			
Slotn-1	normal		outage		normal		outage	
Slotn	...	...	...	...	...	...	...	.
Normal	1	0.998	0.995	0.989	0.996	0.960	0.981	
Outage	0	0.002	0.005	0.011	0.004	0.030	0.019	

Fig. 6 The service availability model based on BN

3.3 *Fault Diagnosis and Design Improvement*

BN is applied for fault detection. Supposing there is a outage of service availability, then the fault probability (the importance) of each per-slot is determined. Set "1" is availability, and "0" is outage. Thus the importance of the i slot is:

$$P(A_{sloti} = 0|A_{SERVICE} = 0) = \frac{P(A_{sloti} = 0)P(A_{SERVICE} = 0|A_{sloti} = 0)}{P(A_{SERVICE} = 0)} \quad (15)$$

where: $i = 1, 2, \ldots, n$.

After the important slots are determined, the service availability can be improved by optimizing MTBO, satellite design life, BR, satellite development, reliability of launch, probability of spares sufficiency, and on-orbit backup rate.

4 Simulation and Validation

The constellation consists of 14 satellites of 5GEO + 5IGSO + 4MEO. The service area is 55°–180° east longitude, 55° south latitude ~55° north latitude. Assuming that the technical conditions of same type satellites are consistent, the MTBO, fault recovery time, satellite design life, satellite development time of different types of satellites are different. Therefore, the satellite types are respectively considered as shown in Table 2.

Table 2 A per-slot parameters of GEO/IGSO/MEO

Outage	Parameters	GEO	IGSO	MEO
STS (bus)	T_{STSB1}	3000 h	5000 h	8000 h
	T_{STSB2}	0.5 h	0.5 h	0.5 h
	T_{STSB3}	20 h	20 h	20 h
STS (payload)	T_{STSP1}	10000 h	12000 h	16000 h
	T_{STSP2}	20 h	20 h	20 h
STU (bus)	T_{STUB1}	30000 h	25000 h	20000 h
	T_{STUB2}	1 h	1 h	1 h
	T_{STUB3}	0.5 h	0.5 h	0.5 h
	T_{STUB4}	20 h	20 h	20 h
STU (payload)	T_{STUP1}	3000 h	5000 h	10000 h
	T_{STUP2}	0.5 h	0.5 h	0.5 h
	T_{STUP3}	0.5 h	0.5 h	0.5 h
	T_{STUP4}	24 h	24 h	24 h
	T_{STUP5}	72 h	72 h	72 h
	T_{STUP6} T_{STUP7}	20 h	20 h	20 h
	f_{STUP1}	0.9	0.9	0.9
LT	T_{LT1}	84096 h	84096 h	72270 h
	T_{LT2}	24 h	24 h	24 h
	T_{LT3}	4380 h	4380 h	2100 h
	T_{LT4}	360 h	360 h	360 h
	T_{LT5}	72 h	72 h	72 h
	T_{LT6}	720 h	720 h	720 h
	T_{LT7}	10a	12a	12a
	f_{LT1}	0.3	0.3	0.3
	f_{LT3}	0.8	0.8	0.8
	f_{LT5}	0.99	0.99	0.99
	f_{LT7}	0.98	0.98	0.98

(1) A per-slot availability

According to Table 2, the steady solutions and availability of per-slot, as Table 3.

In Table 3, the rank of per-slot availability is $A_{SLOTMEO} > A_{SLOTIGSO} > A_{SLOTGEO}$, LT has more significant impact on per-slot availability than STS and STU. In STS, bus faults is more frequent than payload faults, which is the main factor affecting STS. In STU, payload faults is more frequent than bus faults, which is the main factor affecting STU. The satellite design life, development time and on-orbit testing time are main factor affecting LT.

Table 3 Steady-state solutions and a per-slot availability of GEO/IGSO/MEO

Parameters	Slot (GEO)	Slot (IGSO)	Slot (MEO)
P_{STS}	0.9912	0.9943	0.9962
P_{STU}	0.9814	0.9883	0.9935
P_{LT}	0.9805	0.9805	0.9816
A_{SLOT}	0.9556	0.9648	0.9723

(2) Service Availability

Simulating the PDOP values of different outage combinations. According to the actual operation, three or more slots on outaged at the same time are rare, so the paper consider the following three cases: all the slots are available, any one slot is outage, any two slots are outage, a total of three categories 106 kind of state, as shown in Table 4. Figure 7 shows the constellation values for the first and second cases.

In Fig. 7, Slot (GEO1) and Slot (IGSO5) are main factors in PDOP than other solts.

Similarly, simulating the third case, and the PDOP are determined.

Table 4 Assumption conditions of PDOP Simulation

No.	Description	状态数
1	All 14 slots are available	1
2	Any one slot is outage	14
3	Any two slots are outage	91

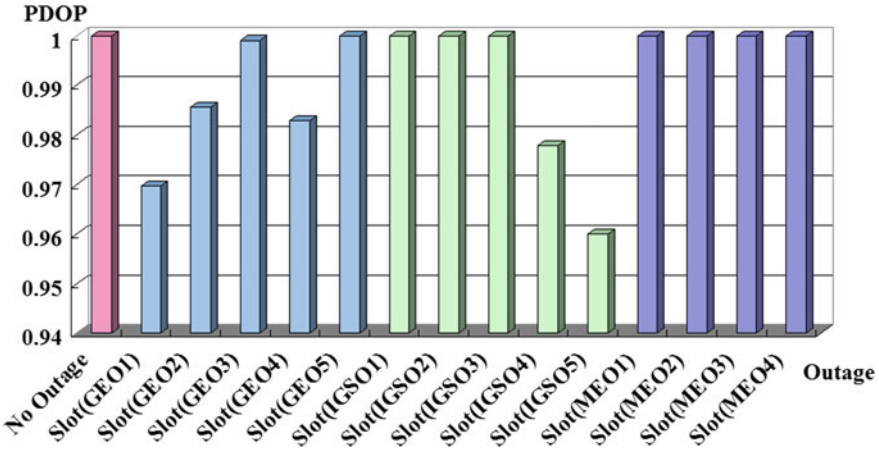


Fig. 7 The PDOP under different outage combinations

The service availability is modeled by BN:

$$A_{SERVICE} = P(A_{SERVICE} = 1) = \sum_{A_{sloti}} P(A_{slot_1}, A_{slot_2}, \ldots, A_{slot_{14}}, A_{SERVICE}) = 0.9914$$

(3) Fault Diagnosis and Design improvement

Taking fault diagnosis of service availability by BN, as Fig. 8.

In Fig. 8, the slot (GEO) has the most significant impact on service availability, followed by the slot (IGSO) and slot (MEO). The reason is that the choice of service area has a significant impact on service availability and the impact of the slot (MEO) will increase significantly as the service area expands to a global area.

As STU and LT are major factors, the STU and LT of Slot (GEO) and Slot (IGSO) are optimized.

Strategy 1, optimizing the outage design. Extend the MTBO of slot (GEO) and slot (IGSO) by 10, 20, 30, 40 and 50% respectively. The service availability after optimization is respectively 0.9918, 0.9920 and 0.9921, 0.9922, 0.9925.

Strategy 2: optimize the design life. Extend MMD for GEO and IGSO by 10, 20, 30, 40 and 50% respectively. The service availability after optimization is respectively 0.9917, 0.9920, 0.9924 and 0.9926, 0.9928.

Strategy 3: optimize on-orbit testing time. Extend on-orbit test time of GEO and IGSO by 10, 20, 30, 40 and 50% respectively. The service availability after optimization is respectively 0.9915, 0.9917, 0.9919 and 0.9921, 0.9923.

Figure 9 shows the comparison between different optimization strategies and the original plan.

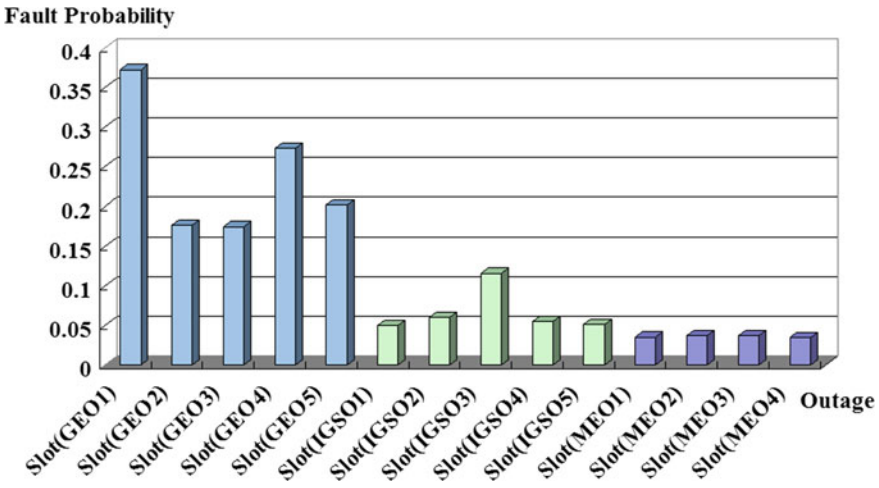


Fig. 8 The fault probability of different slots outage

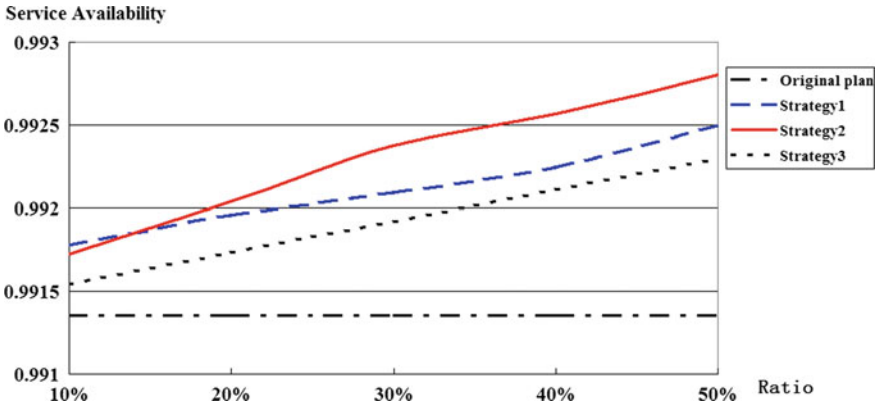


Fig. 9 The service availability of different schemes

In Fig. 9, under the same optimization ratio, service availability improvement of strategy 2 is more significant than that of strategy 1 and strategy 3. The service availability can be improved by selecting long-life components, carrying out reliability accelerated life test, and extending satellite design life. The strategy 1 is slightly better than strategy 3, but it is constrained by design alternatives, product features, and process capabilities. As the optimization ratio increases, its implementation becomes more difficult. According to the engineering practice, the on-orbit testing time is shortened and there is still much room for improvement. It is of great significance to improve the service availability of the system by optimizing on-orbit testing procedures and improving on-orbit testing efficiency.

5 Conclusions

The service availability is modeled and analyzed by GSPN, BN and outage analysis:

(1) The PDOP and the mean time between outage, the mean business recovery time, probability of spares sufficiency, on-orbit backup rate, satellite design life, satellite development time, reliability of launch, satellite on-orbit test time and so as STS (bus and payload), STU (bus and payload) and LT are considered in this paper. The per-slot of GEO, IGSO, MEO is $A_{slotGEO} = 0.9556$, $A_{slotIGSO} = 0.9648$, $A_{slotMEO} = 0.9723$, and the service availability is $A_{SERVICE} = 0.9914$.

(2) The optimization strategies of optimizing the interrupt design, optimizing the design life and optimizing the on-orbit testing time are compared. The simulation shows service availability is 0.9928 at 50% increase by optimizing the

design life. In addition, shorting the on-orbit testing time has more room for improvement, and it has great significance for service availability.

(3) The method proposed by this paper, can be applied in BeiDou Satellite Navigation System for construction and operation.

Acknowledgements Project (61503077) supported by the National Science Foundation of China.

References

1. Durand JM, Caseau A (1990) GPS, availability, part II: evaluation of state probabilities for 21 satellite and 24 satellite constellations. Navig. J. Inst. Navig. 37(3):285–296
2. Quyen H (1997) Availability: what is availability? Availability of what? In: Proceedings of the national technical meeting 1997, Santa Monica, California, pp 831–838
3. Slattery R, Kovach K (1999) New and improved GPS satellite constellation availability model. In: Proceedings of ION GPS 1999, Nashville, TN, 14–17 Sept 1999, pp 2103–2112
4. Li G, Li J, Jiao W et al (2010) Analysis of PDOP availability of navigation constellation based on satellite service intermittence. Geomatics Inf Sci Wuhan Univ 35(7):841–845
5. Wang E, Zhang Q, Qu P et al (2017) GNSS constellation availability evaluation method based on Markov chain. Syst Eng Electron 35
6. Department of Defense of United States of America (2009) Global positioning system standard positioning service performance standard, 4th edn. Sept 2008 [S/OL]. http://pnt.gov/public/docs/2008/spsps2008.pdf

Beidou Receiver IFB Calculation and Stability Analysis

Yongxing Zhu, Laiping Feng, Xiaolin Jia, Yuxi Liu and Xianqiang Cui

Abstract The Beidou Navigation Satellite System (BDS) has been providing navigation services for customers all over the whole Asia-Pacific region since Dec, 2012. The international GNSS Monitor and Assessment System (iGMAS) has constructed more than 20 Multi-GNSS monitoring stations around the world, which provides the research foundation for the high precision application of BDS. The Inter-Frequency Bias (IFB) of the receiver is the main error affecting the high precision application of the navigation satellite system. In this paper, a variety of calculation methods for the receiver GPS IFB are analyzed. Using the measured data, we accessed the stability and accuracy of the single-station receiver IFB calculation method for BDS, and proposed some suggestions for the application of this method.

Keywords Beidou navigation satellite system (BDS) · International GNSS monitor and assessment system (iGMAS) · Inter-frequency bias (IFB) Stability

Y. Zhu (✉)
Information Engineering University, Zhengzhou 410052, Henan, China
e-mail: 18049269002@126.com

Y. Zhu · L. Feng · X. Jia · Y. Liu
Xi'an Research Institute of Surveying and Mapping, Xi'an 710054, Shanxi, China

Y. Zhu
State Key Laboratory of Geo-Information Engineering, Xi'an 710054, Shanxi, China

X. Cui
School of Geosciences and Info-Physics, Central South University, Changsha 410083, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_23

1 Introduction

Difference frequency signals has different time delay on the transmission media. Multi-frequency signals of the satellite navigation system has time delay deviation on different frequency channels of the satellite and the receiver. The deviation, ranging from a few nanoseconds to tens of nanoseconds, is referred to as the Differential Code Bias (DCB) to the satellite, and the inter-frequency bias (IFB) to the receiver [1–4]. The hardware time delay deviation is the main error affecting high-precision ionospheric Total Electron Content (TEC) extraction, precision single-point positioning, and baseline calculation, so it must be corrected in high-precision applications [4–8]. At present, high-precision navigation satellite systems include GPS, GLONASS galileo and BDS, etc. Among them, GPS and GLONASS have been working for a long period, so the solutions of their DCB and receiver IFB are multiple and mature [1–6]. While for the high-precision applications of BDS, we should find its own methods to solving its high-precision DCB and receiver IFB [9].

Similar to that of other GNSS systems, the BDS satellite DCB is relatively stable and relates only to the satellite frequency and its transmission channel. Users can obtain this from the International GNSS Service (IGS) Analysis Center, iGMAS Analysis Center or Broadcast Ephemeris [7]. However, different stations have different receiver IFB, and users have to solve the IFB by themselves. Because BDS have obviously different constellation, frequency and ranging code with GPS, we cannot simply use the solution of GPS receiver. In this study, we proposed a solution for BDS receiver based on the mature methods and the existing research results from other systems. In addition, the iGMAS as the first monitoring and evaluation integrated service platform proposed by China, has built more than 20 monitoring stations covering four major navigation satellite systems of the world, providing a large number of necessary research data.

For GPS and GLONASS receiver IFB is tranditionally testimated by the global or regional ionospheric modeling. Satellite DCB, receiver IFB, and VTEC are solved together as parameters. The accuracy of the obtained receiver IFB is strongly correlated with the selected ionospheric model. In particular, in small areas, its accuracy may be affected by the ionospheric model error [3]. Li et al. proposed a method to estimate the receiver IFB of BDS based on the single-difference between stations, which depends on a reference station with a known receiver IFB and close distance [8]. Yao et al. proposed a method for estimating Beidou satellite DCB and receiver IFB based on the Global Ionospheric Map (GIM) constraint, which depends on the external post-GIM product [1]. For the combined GPS/GLONASS precise single-point positioning, Liu et al. proposed to introduce multiple independent receiver IFB parameters and calculate them during the positioning [4]. But the algorithm is complex. An et al. proposed a single station receiver IFB solution for GPS, which is simple, practical and with high precision [10]. In this paper, we analyzed the stability and accuracy of the BDS receiver IFB based on single-station IFB solution method, using the IGS and IGMAS station datas.

2 Single Receiver IFB Calculating Method

2.1 Receiver IFB Extraction

The receiver IFB is generally coupled with the ionospheric TEC and satellite DCB. And several extraction methods have been developed [11–14]. In general, the ionospheric refraction delay is inversely proportional to the square of the signal frequency. Neglecting the effects of higher-order ionospheric terms, we get the vertical path ionospheric delay (VTEC), satellite DCB^{b^s} and receiver IFB^{B_j}, as follows:

$$VTEC + B_j + b^s + \varepsilon = f(z)\frac{f_1^2 f_2^2}{40.28\left(f_1^2 - f_2^2\right) \times 10^{16}}\left(\tilde{P}_2 - \tilde{P}_1\right) = f(z)\alpha_1\left(\tilde{P}_2 - \tilde{P}_1\right) \tag{1}$$

where $\tilde{P}_1$ and $\tilde{P}_2$ are the two frequency-smoothed pseudoranges, $f(z)$ is the projection function, and ε is the combination of multipath error and noise in the two frequency-smoothed pseudorange observations. Simple conversion of the above equation leads to the receiver IFB estimation formula

$$B_j = f(z)\alpha_1\left(\tilde{P}_2 - \tilde{P}_1\right) - VTEC - b^s + \varepsilon \tag{2}$$

where the value of b^s can be obtained through a variety of ways, B_j and VTEC are the unknowns.

2.2 Calculating Method

The solution of single station receiver IFB calculating is based on two assumptions [10]: the receiver IFB is stable over a certain period of time (e.g. one day) and the value of VTEC at all ionospheric pierce points of a single station within a certain period is almost the same. For the first assumption, the conventional methods has got consensus. However, the second assumption cannot be valid because of temporal and spatial changes. But the standard deviation of each point VTEC relative to the mean VTEC should be the smallest after correctly modified of satellite DCB and receiver IFB.

Assuming the data of a station were collected in N periods, and each period has M ionospheric pierce points, the standard deviation σ of all epochs is calculated by

$$\sigma = \sum_{i=1}^{N}\sqrt{\frac{1}{M}\sum_{j=1}^{M}\left(VTEC_j - m_i\right)} \tag{3}$$

where i is the number of periods, j is the number of ionospheric pierce points in a period, and m_i is the mean VTEC in the period. In this paper, each epoch is defined as a period. Taking the satellite DCB as a known value and defining the search range of the receiver IFB, different standard deviations can be obtained. When the standard deviation is the minimum, the receiver IFB can be obtained.

3 Experiment and Analysis

3.1 Test Data and Methods

In order to verify the single-station IFB calculation method for BDS, data of three iGMAS monitoring stations (i.e., bjf1, kun1 and wuh1) and two MGEX tracking stations (i.e., cut0 and jfng) were selected (Table 1). These data were collected in one week (day of year of 113–119 of 2016). We calculated the BDS receiver IFB of frequency combination B1B2 and B1B3, and analyzed the stability of the single-day calculation results. The results were then compared with the multi-day receiver IFB of the multi-GNSS global ionospheric model [15].

3.2 Results and Analysis

The single-station receiver IFB calculation method was used to solve the BDS receiver IFB values of B1B2 and B1B3, and the calculation time was set as one day. The results are shown in Tables 2 and 3. Taking the average value of 7 days for each station as a reference, we analyzed the stability of the receiver IFB solution. The results (Figs. 1 and 2) show that the discrete value of the BDS receiver IFB of the frequency-combination is about 2.0 ns, and the solution accuracy of IGS stations is comparable to that of iGMAS stations, which is slightly worse than that of the frequency-combination of GPS [10]. The calculation stability at station cut0 is slightly worse than that of other stations. This could be caused by that the station is located at the edge of the BDS coverage and has few visible satellites.

There are significant differences between the BDS constellation and GPS constellation. On the one hand, BDS has more GEO satellites but the same number of

Table 1 Parameters of the selected stations

	Name	Coordinates	Source
1	bjf1	115.892, 39.608	iGMAS
2	kun1	102.797, 25.030	iGMAS
3	wuh1	114.491, 30.516	iGMAS
4	cut0	115.895, −32.004	MGEX
5	jfng	114.491, 30516	MGEX

Table 2 Receiver IFB of B1B2 (unit: ns)

DOY	bjf1	kun1	wuh1	cut0	jfng
113	8.20	72.80	14.40	43.20	43.00
114	8.40	71.90	15.60	43.00	44.30
115	8.60	72.90	15.00	44.70	42.60
116	8.50	70.80	15.80	44.00	44.00
117	8.20	71.60	14.00	44.20	42.60
118	8.70	73.40	14.10	44.30	41.90
119	9.40	73.40	15.80	44.00	43.50
Mean	8.57	72.40	14.96	43.91	43.13

Table 3 Receiver IFB of B1B3 (unit: ns)

DOY	bjf1	kun1	wuh1	cut0	jfng
113	2.60	35.00	19.20	23.90	20.60
114	2.00	32.20	21.00	24.20	22.30
115	1.80	33.90	19.30	26.00	20.40
116	2.20	31.50	20.70	25.10	21.80
117	2.40	33.00	18.70	24.70	19.80
118	2.80	34.60	17.50	24.90	19.40
119	1.50	34.30	18.80	25.10	22.20
Mean	2.19	33.50	19.31	24.84	20.93

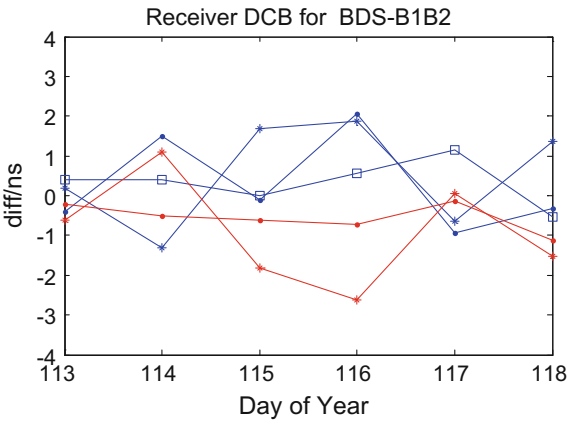


Fig. 1 Receiver IFB stability of Beidou B1B2 combination

ionospheric datas, resulting in less ionospheric pierce points in the same period than the GPS. On the other hand, BDS, as a regional system, has fewer satellites at the edge of its coverage (e.g., the cut0 station) and less ionospheric pierce points, resulting in poor solution results. Therefore, using single-station receiver IFB calculation method to solve the BDS receiver IFB needs to consider the location of the station and try to ensure a uniform distribution of the ionospheric pierce points.

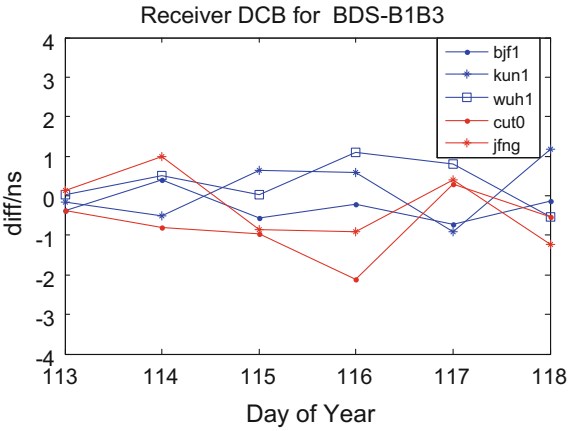


Fig. 2 Receiver IFB stability of Beidou B1B3 combination

Table 4 Comparison of the receiver IFB value obtained by two methods

Scheme	bjf1	kun1	wuh1	cut0	jfng
S1	8.78	72.39	15.16	44.05	42.90
S2	8.57	72.40	14.96	43.91	43.13
DIFF	0.21	−0.01	0.20	0.14	−0.23

To further assess the accuracy of the results, taking the receiver IFB of B1B2 as the reference, we compared the mean value of single-station receiver IFB in 7 days (denoted by S2) with the multi-GNSS global ionospheric modeling multi-day receiver IFB (denoted by S1). The results (Table 4) show that the BDS receiver IFB difference obtained by the two methods (denoted by DIFF) is about 0.2 ns, which achieves a high accuracy. It proves that the proposed method is feasible for BDS receiver IFB calculation. In addition, this method is simple. In order to ensure the accuracy of the solution, the mean receiver IFB in several days (e.g., one-week) should be considered.

4 Summary

The rapid development of the BDS and the construction of monitoring stations provide important information for studies on satellite navigation applications, space environment monitoring and ionospheric delay correction in the Asia-Pacific region, and challenge the BDS receiver IFB solution at the same time. In this study, we used some iGMAS and IGS stations data to analyze the feasibility of single-station IFB solution method for BDS. The following conclusions are drawn:

BDS, the stability of single-station receiver IFB calculation method is analyzed. The receiver IFB discrete value of frequency-combinations is about 2.0 ns. The accuracy of IGS stations is roughly equivalent to that of iGMAS stations. Because of the characteristics of the BDS constellation, when use the proposed method to solve the BDS receiver IFB, the ionospheric pierce points should be evenly distributed as far as possible.

Second, the accuracy of single-station receiver IFB calculation method is analyzed. Comparison between the mean value of single-station receiver IFB in 7 days and the results of multi-GNSS global ionospheric modeling show that the difference is about 0.2 ns, which achieves higher accuracy. The results prove that the proposed method is suitable for BDS receiver IFB calculation, but it is recommended to use the mean value of some days to ensure the accuracy.

Acknowledgements This work was supported by National Key Research Program of China "Collaborative Precision Positioning Project" (No.2016YFB0501900), The National Natural Foundation of China (No.41604024, 41674012).

References

1. Yao Y, Liu L, Kong J et al (2017) Estimation of BDS DCB combining GIM and different zero-mean constraints. Acta Geodaetica et Cartographica Sinica 46(02):135–143 (in Chinese)
2. Wang N, Yuan Y, Zhang B et al (2016) Accuracy evaluation of GPS broadcast Inter-Signal Correction (ISC) parameters and their impacts on GPS standard positioning. Acta Geodaetica et Cartographica Sinica 45(08):919–928 (in Chinese)
3. Xie Y, Wu J, Chen J, Liu W (2014) Global ionospheric TEC modeling using measured GPS and GLONASS. Geomantics Inf Sci Wuhan Univ 39(08):930–934 (in Chinese)
4. Liu Z, Wang J, Duan B (2015) Estimation of GLONASS code inter-frequency biases with multiple parameters based on a single station and its impact on combined precise point positioning. Acta Geodaetica et Cartographica Sinica 44(02): 150–159 (in Chinese)
5. Shu B, Liu H, Zhang M, Wu D (2016) Evaluation and analysis of BDS instrumental biases. Geomantics Inf Sci Wuhan Univ 41(02):279–284 (in Chinese)
6. Ruan R, Wu X, Feng L (2016) Comparison of observation models and ionospheric elimination approaches for single frequency precise point positioning. Geomantics Inf Sci Wuhan Univ 41(02):279–284 (in Chinese)
7. Li H, Li B, Wang X, Xu T (2016) A method for estimating BeiDou inter-frequency satellite clock bias. Acta Geodaetica et Cartographica Sinica 45(02):140–146 (in Chinese)
8. Li X, Guo J, Zhou C et al (2016) An accurate method for the BDS receiver DCB estimation in a regional network. Acta Geodaetica et Cartographica Sinica 45(08):929–934 (in Chinese)
9. Yang Y, Li J, Xu J et al (2011) Contribution of the compass satellite navigation system to global PNT users. Chin Sci Bull 56(21):2813–2819 (in Chinese)
10. An J, Wang Z, Qu X, Du Y (2010) Methods for estimation of differential code biases by using single station GPS measurements. J Geodesy Geodyn 30(02):86–90 (in Chinese)
11. Nie W, Hu W, Pan S, Song Y (2014) Extraction of regional ionospheric TEC from GPS dual observation. Geomantics Inf Sci Wuhan Univ 39(09):1022–1027 (in Chinese)

12. Zhang Q, Zhao Q, Zhang H, Chen G (2016) BDS differential code bias estimation using Beidou experimental tracking stations. Geomantics Inf Sci Wuhan Univ 41(12):1649–1655 (in Chinese)
13. Zhang B, Ou J, Yuan Y, Li Z (2011) Calibration of slant total electron content and satellite-receiver's differential code biases with uncombined precise point positioning technique. Acta Geodaetica et Cartographica Sinica 40(04):447–453 (in Chinese)
14. Rossbach U (2011) Positioning and navigation using russian satellite system GLONASS. University of the Federal Armed Forces Munich, Munich (in Chinese)
15. Zhu Y, Jia X, Feng L, Ruan R (2015) Analysis of ionosphere modeling accuracy based on multi-GNSS Data. In: China satellite navigation conference 2015 proceeding, pp 545–554

Part II
User Terminal Technology

Research on Miniaturized Receiving Antenna for Satellite Navigation

Haiguang Zhang, Jia Zhang and Kui Xu

Abstract In this paper, a new circular polarization microstrip antenna for satellite navigation receiving is proposed. The antenna layout employs electromagnetic bandgap structures over a Rogers TMM10i ($\varepsilon_r = 9.8$) substrate to achieve the effective permittivity enhangcement. At the edge of the radiation patch, four coupled shorting strips are designed to achieve the goals of size miniaturization and better circular polarization property. The measured and simulated results show that the effective bandwidth can cover the whole B3 band. The axial ratio for the zenith is less than 2 dB and less than 8 dB at the elevation 20°. The proposed antenna fabricated over a substrate with the size of 25 mm * 25 mm * 5 mm that has approximately 45% size reduction compared with the conventional microstrip antenna over the same substrate. The measured realized gain is 3.3 dBic over a 80 mm * 80 mm ground plane with 24 MHz 3 dB realized gain bandwidths. The proposed receiving antenna has important values in satellite navigation.

Keywords Satellite navigation · Receiving antenna · Electromagnetic bandgap structures · Miniaturization

1 Introduction

Relatively weak signal level of the Chinese Beidou navigation system on the ground makes it vulnerable by man-made intentional or nature noise jammers. In order to solve this problem military system employing multielement BD arrays to generate radiation pattern nulls in the directions of the jamming signals [1]. It has great meaning to investigate miniaturized receiving antenna to form portable anti-jamming arrays in a great number of platforms that have limited space. To achieve the demand above, technique [2] has been applied in literature for miniaturization of GPS antenna operating at the L2 band utilizes capacitance and

H. Zhang (✉) · J. Zhang · K. Xu
Xi'an Research Institute of Navigation Technology, Xi'an 710068, China
e-mail: zhg54321@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_24

inductance loaded that implements over a permittivity $\varepsilon_r = 9.8$ substrate. The antenna can realized the size of 30 * 30 mm^2. Electromagnetic Band-Gap (EBG) structures have some special electromagnetic properties such as slow-wave effect, when utilize in the antenna can miniaturize the antenna size [3]. But the literature don't analyze the effect of EBG structure parameters on slow-wave wavelength.

According to the EBG structure equivalent circuit model, the resonance frequency of the structure is calculated in this paper. When the EBG structure perform before the resonance frequency, the slow-wave effect occurs. Make use of this phenomenon to achieve the antenna size reduction. Analyze the influence of EBG structure parameters on slow-wave wavelength by simulate S parameter performance in the wave-guide model. Then calculate the effective permittivity. Consider the antenna bandwidth and the surface wave effect to assign the optimal EBG structure size. Four coupled strips connected to the ground by vertical conducting vias are designed around the squareness patch to achieve the further size reduction. The proposed antenna in this paper has a size reduction of 45% compared to a traditional patch antenna fabricated on the same substrate with the size of 25 * 25 * 5 mm^3.

2 Antenna Design

2.1 Analyze the EBG Structure

As the "mi" structure illustrated in Fig. 1. When the structure interacts with electromagnetic waves, currents are induced in the top of the patch, so the currents accumulate charges on the ends of the two patches, which can be described as a capacitance. As the charges flow around the patches and the ground through the shorting pin, which can be described as a inductance. So it can be equivalent as a LC parallel resonance circuit that is shown in Fig. 2 [4]. The capacitance C and inductance L equivalent equations are:

$$C = \frac{\varepsilon_0(1+\varepsilon_r)a}{\pi}\cosh^{-1}\left(\frac{a+g}{g}\right) \tag{1}$$

$$L = \mu_0 \times (H1 + H2) \tag{2}$$

where ε_0 is the permittivity of the free space and μ_0 is the permeability of the free space, a is the length of the "mi" structure, g is the gap between the EBG structures.

According to the equation $\omega = 1/\sqrt{LC}$ can calculate the resonance of the EBG structure.

Simulate the S parameter performance of the EBG structure in the wave-guide model, the initial parameters are given as follow: the top layer and sub layer both utilize high permittivity Rogers TMM10i ($\varepsilon_r = 9.8$, $\tan\sigma = 0.002$), the element gap

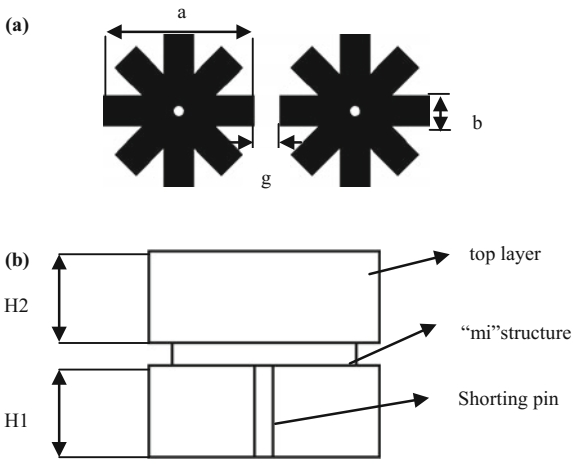


Fig. 1 Geometry of the "mi" structure

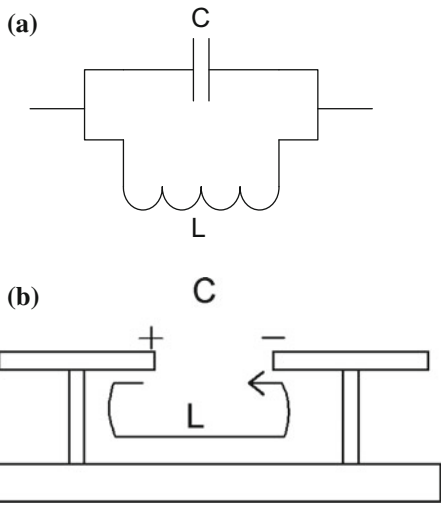


Fig. 2 The circuit model of the "mi" structure

g is equal to 0.5 mm, the length of the patch is equal to 3 mm, the width of the patch is equal to 0.6 mm, the height of the sub layer H1 = 2 mm, the height of the top layer H2 = 2 mm, the diameter of the shorting pin d = 0.5 mm. As the simulated mode illustrated in Fig. 3, assign the top and sub faces as PEC, assign the front and back faces as PMC. The simulated magnitude and phase results as shown in Fig. 4.

As shown in Fig. 4a the resonance of the "mi" structure is 1.42 GHz, we can also observe that at 1.72 GHz electromagnetic wave forbidden band is emerged. Figure 4b depict the mutation of the S21 phase performance. In order to analyze this phenomenon scientific, utilize S parameter formula proving to figure out the

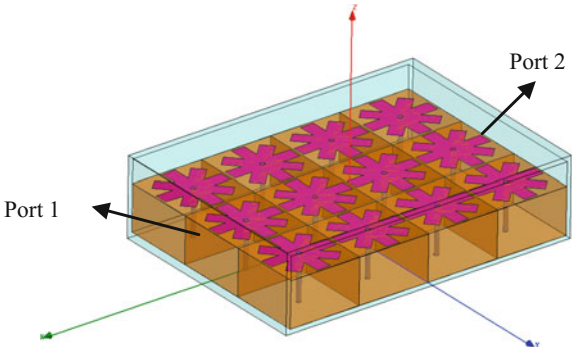


Fig. 3 Computational model of the EBG structure

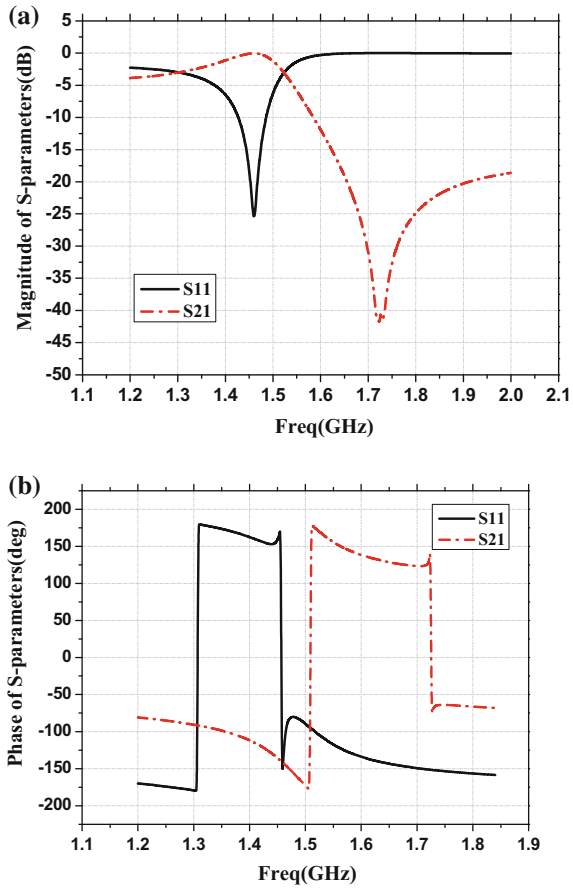


Fig. 4 Simulated S11 performances of the EBG structure: **a** magnitude; **b** phase

permittivity and permeability of the substrate that loaded with "mi" EBG structures [5]. S21 and S11 parameters can be expressed as:

$$S_{21} = \frac{1}{\cos(nkd) - (i/2)(z + Z^{-1})\sin(nkd)} \tag{3}$$

$$S_{11} = \frac{i}{2}(z^{-1} - z)\sin(nkd)S_{21} \tag{4}$$

where n denotes the refractive index of the substrate, z denotes the wave impedance, $k = \omega/c$ is the wave vector of the incident wave. Then n and z parameters can be expressed as:

$$n = \frac{1}{kd}\cos^{-1}[\frac{1}{2S_{21}}(1 - {S_{11}}^2 + {S_{21}}^2)] \tag{5}$$

$$z = \pm\sqrt{\frac{(1 + S_{11})^2 - {S_{21}}^2}{(1 - S_{11})^2 - {S_{21}}^2}} \tag{6}$$

Transform the function of $\cos^{-1}$ as:

$$\mathrm{Im}(n) = \pm\mathrm{Im}\{\frac{\cos^{-1}[\frac{1}{2S_{21}}(1 - {S_{11}}^2 + {S_{22}}^2)]}{kd}\} \tag{7}$$

$$\mathrm{Re}(n) = \pm\mathrm{Re}\{\frac{\cos^{-1}[\frac{1}{2S_{21}}(1 - {S_{11}}^2 + {S_{22}}^2)]}{kd}\} + \frac{2\pi m}{kd} \tag{8}$$

After figure out the n and z parameters, ε and μ can be expressed as:

$$\varepsilon = n/z \quad \mu = nz \tag{9}$$

Figure 5 depict the permittivity and permeability of the substrate that loaded with "mi" EBG structures. The results show that the real part of the efficiency permittivity become bigger with the frequency increasing, whereas efficiency permittivity become less than zero at the band of 1.72 GHz. Moreover the real part of the efficiency permeability increased within the whole band, and decreased rapidly around the band of 1.72 GHz.

Slow-wave effect occurs when the resonance of the antenna operates before the electromagnetic wave forbidden band of the "mi" EBG structures. Analyze the influence of height of EBG structure H1 and height of the top layer H2 on slow-wave effect and efficient permittivity. The initial value of H1 and H2 are both set up with 2, while keeping other parameters unchanged. The results are shown in Tables 1 and 2.

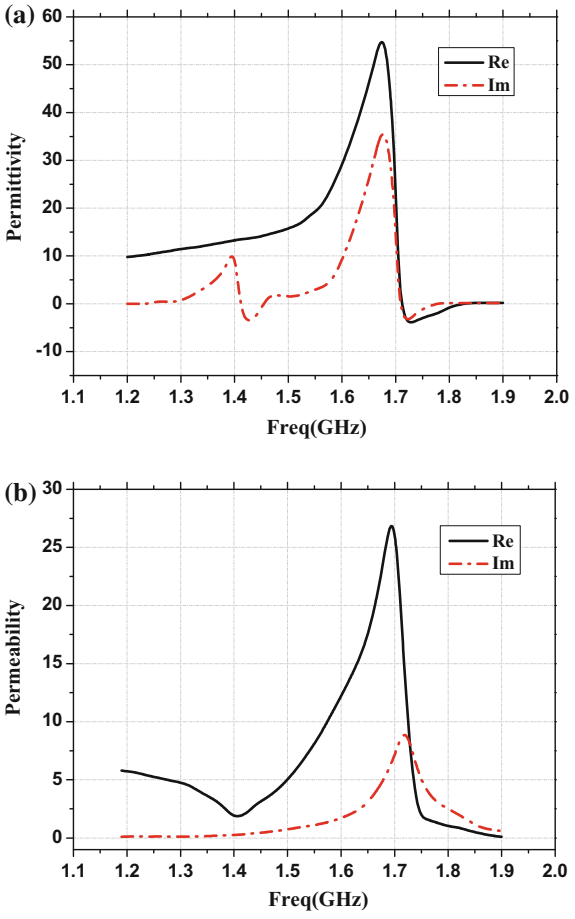


Fig. 5 Performances of the EBG structure: **a** permittivity; **b** permeability

Table 1 H1 = 2 mm, the permittivity and slow-wave wavelength range with different H2 values

H2 (mm)	Slow-wave wavelength (mm)	Efficient permittivity
1	58.85	16.16
1.5	60.74	15.17
2	62.45	14.35
2.5	64.04	13.65
3	64.85	13.31

The results show that the slow-wave wavelength gradually decreased with the height of EBG structure H1 increased, while the efficient permittivity increased either. The slow-wave wavelength gradually increased with the height of top layer H2 increased, whereas the efficient permittivity decreased.

Table 2 H2 = 2 mm, the permittivity and slow-wave wavelength range with different H1 values

H1 (mm)	Slow-wave wavelength (mm)	Efficient permittivity
1	75.2	9.9
1.5	68.54	11.91
2	62.45	14.35
2.5	57.46	16.95
3	52.39	20.4
3.5	50.2	22.2

2.2 *Antenna Design*

As is known in Sect. 2.1, in order to reduce the slow-wave wavelength and increase the efficient permittivity, the height of "mi" EBG structure should be increased otherwise reversed to the height of the top layer. Assign the optimal value of H1 = 3.5 mm, H2 = 1.5 mm after consider the bandwidth and surface wave effect of the antenna.

The squareness radiation patch of the microstrip antenna realized over the "mi" EBG structures substrate. In order to obtain a further miniaturization, four coupled strips connected to the ground by vertical conducting vias that around the radiation patch are designed. The geometry of the radiation patch is shown in Fig. 6.

The structure of the coupled shorting-termination strips can be equivalent as a LC series resonance circuit in which the capacitance and the inductance generated by squareness radiation patch coupled with short-termination strips and the conducting vias, respectively. The quality factor Q of the struture is:

$$Q = (1/R)\sqrt{L/C} \tag{10}$$

Assign the coupled shorting-termination strips adjacent to the squareness radiation patch for a capacitance enhancement in order to decrease the Q, while to expand the bandwidth.

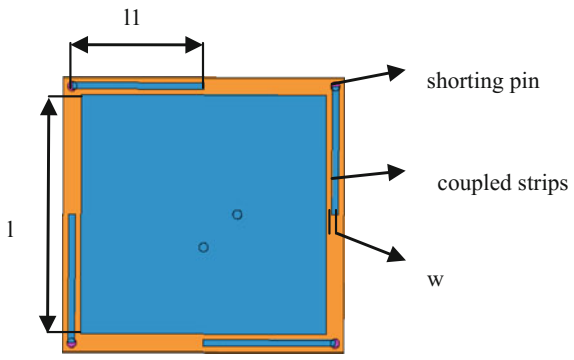


Fig. 6 Geometry of the radiation patch

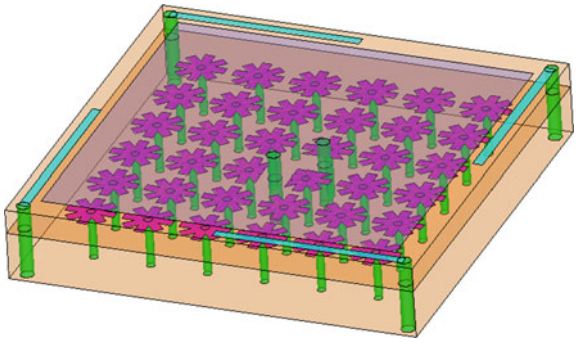


Fig. 7 Geometry of the proposed antenna

Table 3 Parameters of the antenna

Parameter	a	b	g	H1	H2	d	l	l 1	w
Value (mm)	3	0.6	0.5	3.5	1.5	0.5	21.8	11.7	0.6

The size of the proposed antenna is 25 * 25 * 5 mm^3, 6 * 6 "mi" EBG structure arrays fabricated on the sub layer. The computation model of the antenna is shown in Fig. 7 (Table 3).

The feeding network is implemented over a FR4 substrate ($\varepsilon_r = 4.4$, $\tan\sigma = 0.02$) and consists of a surface mount quadrature hybrid coupler (RN2, RCP1500Q03). The antenna is fed through vertical copper pins to realize circular receiving.

3 Results and Discussion

The proposed antenna fabricated over a ground plane with the size of 80 * 80 mm^2. Simulate the antenna over the same substrate with conventional antenna, only loaded with "mi" structures, only loaded coupled shorting strips and the proposed antenna in this paper, respectively. As is shown in Fig. 8, the resonance of the conventional antenna is 2.14, 1.56 GHz for the only loaded with coupled shorting strips, 1.4 GHz for the only loaded with "mi" structures and 1.266 GHz for the proposed antenna. The S11 < −10 dB bandwidths of 22 MHz.

The antenna prototype fabricated using the dimensions of the computational model that is shown in Fig. 9. The measured VSWR performance is shown in Fig. 10. Due to the isolated resistance assigned in the feeding network, the VSWR of the antenna is less than 1.5 over the whole B3 band.

Figure 11 presents the measured and simulated gain and axial ratio at the elevation angel of 90° when the antenna centered on a 80 * 80 mm^2 ground plane. The measured and simulated AR are both less than 2 dB over the whole band.

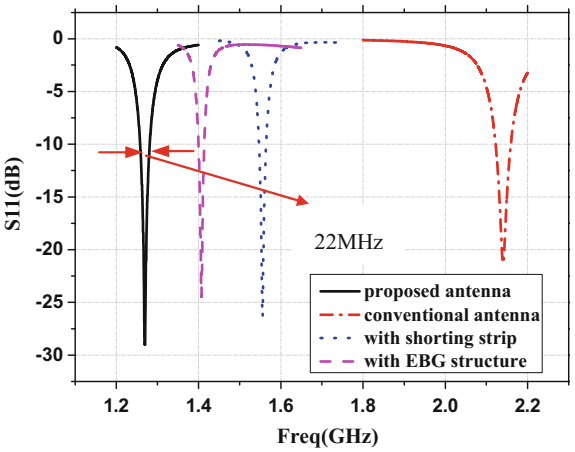


Fig. 8 Simulated S11 performances of the antennas

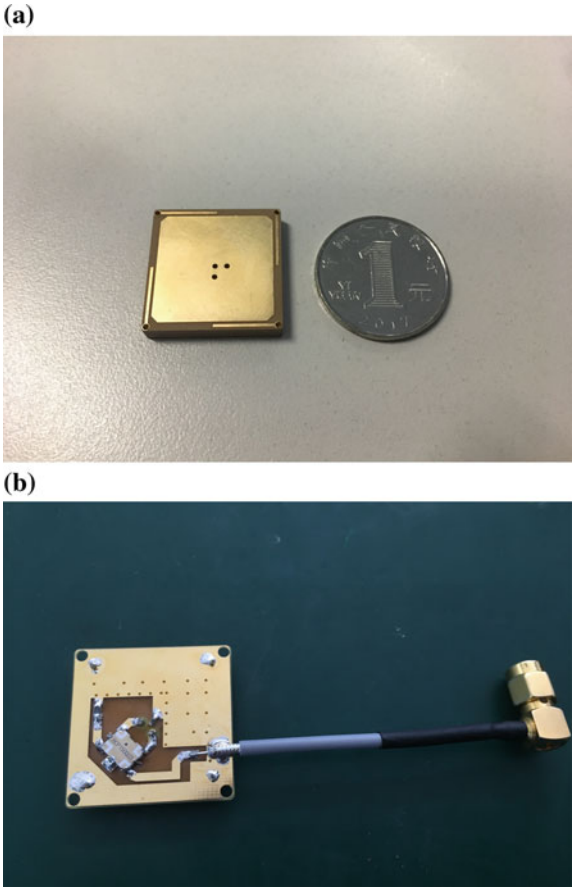


Fig. 9 Fabricated compact B3 band antenna prototype: **a** top view; **b** feed structure

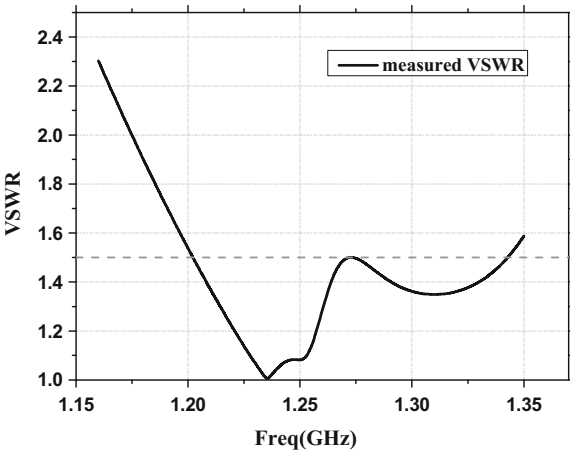


Fig. 10 Measured VSWR of the antenna

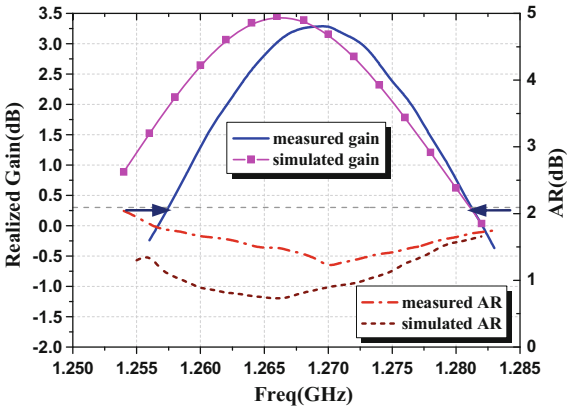


Fig. 11 Simulated and measured gains and AR within the band

The frequency of the measured and simulated maximal gain has a little distinction that can be attributed to the feeding network, fabrication tolerances. The miniature BD receiving antenna performed 3.3 dBic realized gain at 1.27 GHz with 24 MHz 3 dB gain bandwidths. Figure 12 shows the axial ratio at the frequency of 1.268 GHz, the axial ratio at elevation 20° is less than 8 dB. Figure 13 depicts the X-Z plane radiation patterns at the 1.268 GHz, in the main radiation direction the cross polarization levels were 15 dB lower than the polarized gain level. At elevation 20° the realized gain at 1.268 GHz is −5.2 dBic. The simulated results are in good agreement with the measured results. Due to the symmetry of the antenna and ground plane, the simulated and measured radiation patterns in the Y-Z plane are almost equal to those in X-Z plane, therefore not shown in this paper.

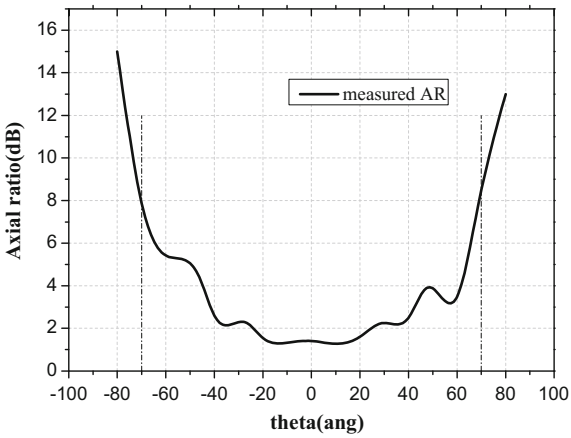


Fig. 12 Measured AR at 1.268 GHz

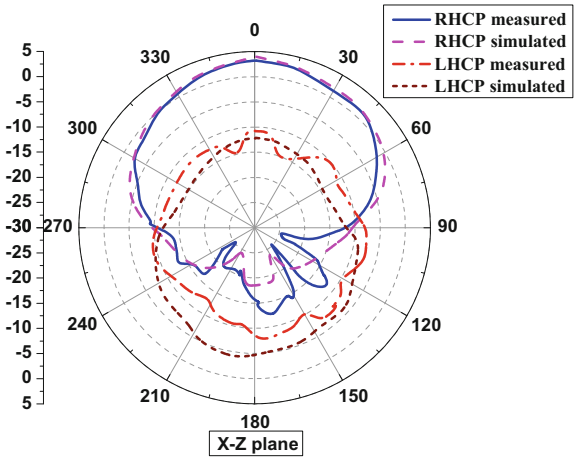


Fig. 13 Simulated and measured radiation patterns (1.268 GHz)

4 Conclusion

A miniature BD receiving antenna was proposed in this paper. To achieve the goal of size miniaturization, the antenna simultaneously incorporated slow-wave effect occurs when the resonance frequency of the antenna operates before the "mi" EBG structures forbidden band, four coupled shorting-termination strips. The proposed antenna fabricated over a $\varepsilon_r = 9.8$ substrate with the size of 25 mm * 25 mm * 5 mm that has approximately 45% size reduction comparing with the conventional microstrip antenna over the same substrate. The measured axial ratio at the elevation 90° is less than 2 dB, the maximal realized gain is 3.3 dBic with 24 MHz 3 dB gain bandwidth. At the elevation 20° the axial ratio is less than 8 dB, the

realized gain is −5.2 dBic. Solve the disadvantages of narrow band, low gain of the conventional miniaturized microstrip antenna. The proposed receiving antenna can act as the element of the portable anti-jamming arrays that has important values in satellite navigation.

References

1. Gupta I, Lee T, Griffith K, Slick C, Reddy CJ, Bailey M, DeCarlo D (2010) Non-planar adaptive antenna arrays for GPS receivers. IEEE Antennas Propag 52:35–51
2. Gupta S, Mumcu G (2013) Dual-band miniature coupled double loop GPS antenna loaded with lumped capacitors and inductive pins. IEEE Trans Antennas Propag 61:2904–2910
3. Suntives A, Abhari R (2010) Design of a compact miniaturized probe-fed patch antenna using electromagnetic bandgap structures. In: Antennas and propagation society international symposium, pp 1–4
4. Sievenpiper D, Zhang L, Broas RFJ, Alexopolous NG, Yablonovitch E (1999) High-impedance electromagnetic surfaces with a forbidden frequency band. IEEE Trans Microwave Theor Tech (47):2059–2074
5. Smith DR, Schultz S, Markos P, Soukoulis CM (2002) Determination of negative permittivity and permeability of meta-materials from reflection and transmission coefficients. Phys Rev B 65(19):1–5

Main Beam Pointing and Null Constraint Optimization for Satellite Navigation Nulling Antenna

Rong Shi

Abstract The power inversion algorithm is usually adopted in the traditional satellite navigation nulling antenna. The null can be formed on the direction of jamming with high power so as to decrease its power into the navigation receiver. The anti-jamming is realized, but receiving performances for navigation satellite signals have not been considered at all in above methods. In this paper, the published navigation satellite almanac and the parameters output from INS in the platform are fully utilized. The direction of signal arrival about each currently visible navigation satellite can be calculated in real time in the nulling antenna coordinate system. According to the LCMP (Linear Constrained Minimum Power) algorithm, the main beam pointing condition is restricted for each visible satellite in order to improve the receiving antenna gain for this direction. This is the first beam forming optimization. Then under the condition that the nulling freedom degree is within the allowable range, the null constraint conditions of antenna beams are imposed to the direction of other navigation satellites signals, in which the antenna gain are relatively larger in the firstly beam forming. This is the second beam forming optimization in order to reduce the multiple access interference of CMDA system. Through above two optimizations, it greatly reduces the external oppressive jamming and mutual interference between different navigation satellites. Finally, the feasibility and availability of this method are verified by simulations. It is an important reference for the improvement of anti-jamming performance and engineering application for satellite navigation nulling antenna.

Keywords Satellite navigation · Nulling antenna · Almanac · INS
Main beam pointing constraint · Null constraint · Oppressive jamming
CDMA interference

R. Shi (✉)
Science and Technology on Electronic Information Control Laboratory,
Chengdu, China
e-mail: jasminewyx@sina.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_25

1 Introduction

The power of navigation satellite signals received on the ground is very small. For example, the power of L1 C/A code signal of GPS is about 160 dBW and those of L1 and L2 P(Y) code are about −163 and −166 dBW respectively [1]. The signals are so weak and vulnerable to external jamming at the same frequency. Especially in the modern warfare, the military navigation receiver is inevitable to be jammed with high power. Therefore, the nulling antenna in satellite navigation is more and more widely used. It can form antenna beam null automatically to the direction of arrival for jamming signals and the antenna gain is reduced in this direction. The received jamming signal is greatly attenuated so as to achieve the purpose of anti-jamming.

The jamming is suppressed without receiving performance optimization by current nulling algorithms, whether a minimum mean square error model based on reference signals or LCMP (Linearly Constrained Minimum Power) model [2, 3]. When its nulls have been formed, the width of null is very large to decrease the antenna gain about the direction of satellite signal, and the received signal is also affected at the same time. Although the space-time processing methods have been put forward in some literatures by the intensity constraint on the satellite signals, the problem of mutual interference between CDMA signals is not considered in these approaches. The performance improvement of receiver is still limited.

In this paper the directions of satellite signals arrival at certain time are firstly calculated according to the almanac for the platform. The accurate directions in the nulling antenna coordinate system can be converted through the attitude parameters output by INS. Then two optimization processes are carried out to ensure that the independent main beam is formed for each desired satellite navigation signal, the gains of other satellites signal direction are low and the nulls are still retained to the jamming direction at the time. In this way the mutual interferences among different satellite signals and jamming are heavily reduced together. It greatly improves the performance of the satellite navigation receiver.

2 Problem for Traditional Nulling Model

In the satellite navigation receiving array which is consisted of M antenna units, the input signal vector and the weighted vector are denoted as $\boldsymbol{X} = (x_1, x_2, \ldots, x_M)^{\mathrm{T}}$, $\boldsymbol{W} = (w_1, w_2, \ldots, w_M)^{\mathrm{T}}$ respectively. The output signal of this array can be expressed by the following formula:

$$y = \boldsymbol{W}^{\mathrm{H}}\boldsymbol{X} \tag{1}$$

where H represents conjugate transpose. Usually satellite navigation signals are spectrum spread and their powers are very weak under the noise floor in the receiver

on the ground, while the jamming signal is often above them. So the power inversion algorithm is utilized to form the null effectively in the direction of jamming signal arrival. According to the theory of array signal processing, the LCMP is equal to above model [4]. The LCMV (Linear Constrained Minimum Variance) model is also called in some literatures. Its non-distortion constraint conditions are as follows:

$$\boldsymbol{W}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{PI}} = 1 \tag{2}$$

where $\boldsymbol{C}_{\mathrm{PI}} = (1, 0, 0, \ldots, 0)^{\mathrm{T}}$ is the constraint vector with M dimensions. According to Lagrange multiplier method, the optimal weighted vector can be expressed as

$$\boldsymbol{W}_{\mathrm{con}} = \underset{\boldsymbol{W}}{\operatorname{argmin}}\left[\left\|\boldsymbol{W}^{\mathrm{H}}\boldsymbol{X}\right\|^2 + \beta_1^*\left(\boldsymbol{W}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{PI}} - 1\right)\right] \tag{3}$$

where β_1 is a Lagrange multiplier, and * means conjugate operation. Its result is:

$$\boldsymbol{W}_{\mathrm{con}} = \left(\boldsymbol{R}_{xx}^{-1}\boldsymbol{C}_{\mathrm{PI}}\right)/\left(\boldsymbol{C}_{\mathrm{PI}}^{\mathrm{H}}\boldsymbol{R}_{xx}^{-1}\boldsymbol{C}_{\mathrm{PI}}\right) \tag{4}$$

where $\boldsymbol{R}_{xx} = \mathrm{E}\left(\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}\right)$ is the autocorrelation matrix of signals received by all antenna units. The antenna null can be formed in the direction of the jamming when the optimal weighted vector $\boldsymbol{W}_{\mathrm{con}}$ is introduced into the beam pattern synthesis function. The greater the jamming power is, the deeper the null is.

The only one constraint vector $\boldsymbol{C}_{\mathrm{PI}}$ is utilized in above traditional model, and the autocorrelation matrix is mainly composed of jamming signals and noise. When the null is formed to the direction of jamming, the gains to the direction of normal satellite signals arrival are not controlled confidently. The null width is so large in the suppression of jamming signals that the gains of normal signal direction are reduced at the time. The system performance decreases heavily.

3 Reference Direction of Arrival Based on Almanac and INS

The direction of navigation satellite signal arrival at certain time and certain place can be calculated from the almanac and INS to form the new constraint condition. The functions of the almanac and INS in this process are briefly explained as follows.

(1) The accurate location of satellites known from almanac
The almanac of navigation satellite is fully open and the new almanac can be downloaded from internet. For example, the GPS almanac broadcasted by satellites everyday is open in the webpage www.navcen.uscg.gov. The validity of almanac is more than half a year. The navigation satellites locations,

$\boldsymbol{p}_{\mathrm{s}}^{(i)} = \left(x_{\mathrm{s}}^{(i)}, y_{\mathrm{s}}^{(i)}, z_{\mathrm{s}}^{(i)}\right)^{\mathrm{T}}$, $i = 1, 2, \ldots, L$, in the earth fixed coordinate system can be calculate from the almanac, where the superscript $^{(i)}$ represents the serial number of satellite and L is the total number of satellites visible for receivers.

(2) The rough position of the platform and the attitude parameters of the nulling antenna array are obtained by INS

Although there are many defects of INS, such as lower positioning accuracy than that of the GNSS, it is not affected by the external electromagnetic jamming. It becomes an indispensable role in the integrated navigation application, and now lots of platforms are equipped with different kinds of INS. The position $\boldsymbol{p}_{\mathrm{p}} = \left(x_{\mathrm{p}}, y_{\mathrm{p}}, z_{\mathrm{p}}\right)^{\mathrm{T}}$ of the platform can be obtained by INS, although the error of $\boldsymbol{p}_{\mathrm{p}}$ is relatively large, the accuracy is enough for the direction estimation of navigation satellite signals. On the other hand, the INS outputs the heading, pitch and roll attitude parameters, which also determined the 3×3 dimensional rotation matrix $\boldsymbol{S}_{\mathrm{p}}$ between nulling antenna coordinate and earth fixed coordinate. A detailed calculation about the matrix $\boldsymbol{S}_{\mathrm{p}}$ is described in many documents [5]. It is not necessary to explained here again.

(3) The unit vector of direction of navigation satellite signal arrival in the nulling antenna coordinate can be calculated.

The unit vector $\boldsymbol{n}_{\mathrm{s}}^{(i)}$ of direction of navigation satellite signal arrival in the nulling antenna coordinate can be calculated through the satellites location $\boldsymbol{p}_{\mathrm{s}}^{(i)}$ obtained by almanac, the platform location $\boldsymbol{p}_{\mathrm{p}}$ obtained by INS and the rotation transform matrix $\boldsymbol{S}_{\mathrm{p}}$.

$$\boldsymbol{n}_{\mathrm{s}}^{(i)} = \left(\boldsymbol{p}_{\mathrm{s}}^{(i)} - \boldsymbol{p}_{\mathrm{p}}\right) / \left\|\boldsymbol{p}_{\mathrm{s}}^{(i)} - \boldsymbol{p}_{\mathrm{p}}\right\| \cdot \boldsymbol{S}_{\mathrm{p}} \tag{5}$$

The maximum error of satellite position from almanac within half a year is not more than 50 km. The maximum error of platform position output from INS within several hours will not exceed 20 km. The distance between the platform and navigation satellite is about 20,000 km. Even if there are above errors, the angle error $\Delta\varepsilon$ of unit vector pointing by formula (5) is expressed as follows:

$$\Delta\varepsilon < \mathrm{a}\tan((50 + 20)/20{,}000) \approx 0.2^{o} \tag{6}$$

So the unit vector of direction of navigation satellite signal arrival in the nulling antenna coordinate can be calculated accurately.

4 Beam Forming via Two Step Constrained Optimization

The antenna array with M antenna units has $M-1$ degrees of freedom, and $M-1$ constraints can be added at most. The only one constraint as shown in Eq. (2) is in the traditional nulling antenna model, and the other $M-2$ degrees of freedom are left for the subsequent adaptive null constraints in order to synthesize the beam nulls automatically to the direction of jamming arrival. With the decrease of cost and increase of hardware integration, the antenna unit number becomes more than before gradually, and 7, 9, 19 units are commonly used in the array. So the number of jamming signals is often far less than $M-2$ in many actual applications. The maximum number of jamming is denoted as N_{J}. The $M-2-N_{\mathrm{J}}$ degrees of freedom are not effectively restrained after the N_{J} nulls are formed in the antenna array. The beam pattern of the antenna array is not completely determined under the influence of different noise, resulting in the negative impact of receiving gain decrease and null area expansion in the direction of the normal navigation satellite signal arrival. In order to solve above problems, it is necessary to improve the constraint conditions about beam forming.

(1) Main beam constraint and the first optimization of beam forming

The unit vector $\boldsymbol{n}_{\mathrm{s}}^{(i)}$ of direction of navigation satellite signal arrival in the nulling antenna coordinate is obtained from the almanac and INS. Therefore, the constraint condition in formula (2) can be modified as the main beam constraint condition for the direction about the satellite.

$$\boldsymbol{W}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{s}}^{(i)}=\gamma \tag{7}$$

where γ is a gain coefficient. $\boldsymbol{C}_{\mathrm{s}}^{(i)}=\left(c_1^{(i)},c_2^{(i)},\ldots,c_M^{(i)}\right)^{\mathrm{T}}$ is the direction constraint vector for the ith navigation satellite, in which each component is expressed by the following formula:

$$c_m^{(i)}=\exp\left(\mathrm{j}2\pi(\boldsymbol{p}_{\mathrm{m}}-\boldsymbol{p}_1)\cdot\boldsymbol{n}_{\mathrm{s}}^{(i)}/\lambda\right) \tag{8}$$

where $\boldsymbol{p}_{\mathrm{m}}$, $m=1,2,\ldots,M$, is position in the rectangular coordinate system of nulling antenna. $\cdot$ is the dot product between two vectors. λ is the wave length of navigation signal. The above constraints ensure that the gain of the full array can be obtained in the direction of the interested satellite navigation signal. After the substitution condition (2) with constraint (7), the optimal weighted vector $\boldsymbol{W}_{\mathrm{s}}^{(i)}$ can be obtained as follows:

$$\boldsymbol{W}_{\mathrm{s}}^{(i)}=\gamma\boldsymbol{R}_{xx}^{-1}\boldsymbol{C}_{\mathrm{s}}^{(i)}/\left(\boldsymbol{C}_{\mathrm{s}}^{(i)^{\mathrm{H}}}\boldsymbol{R}_{xx}^{-1}\boldsymbol{C}_{\mathrm{s}}^{(i)}\right) \tag{9}$$

Although the weighted vector $\boldsymbol{W}_{\mathrm{s}}^{(i)}$ can ensure that the beam forming has a gain to the direction of the ith navigation satellite and nulls automatically formed in the direction of jamming, the gains at directions of other satellites are not controlled so as to cause these gains to be increased too. Because the satellite navigation systems are mostly CDMA system, this will increase the multiple access interference between the satellite navigation signals. Therefore, it is necessary to apply the remaining degrees of freedom for the null constraint in the direction of other satellite signals arrival.

(2) Null constraint to the direction of other satellites and the second optimization of beam forming

After the first optimization, beam forming is carried out by using the weighted vector $\boldsymbol{W}_{\mathrm{s}}^{(i)}$. The antenna gain in the direction of other satellite signal arrival is $\mathrm{G}^{(j)}$, and $M-2-N_{\mathrm{J}}$ direction vectors $\boldsymbol{n}_{\mathrm{s}}^{(k)}$ with larger gain are selected and null constraint conditions are applied in these directions as follows:

$$\boldsymbol{W}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{n}}^{K}=(0,0,\ldots,0) \tag{10}$$

where the dimension of right part in Eq. (10) is $M-2-N_{\mathrm{J}}$, $\boldsymbol{C}_{\mathrm{n}}^{K}$ is direction constraint vector with $M\times(M-2-N_{\mathrm{J}})$ dimensions, whose components are similar to the Eq. (8). The constraint condition in the second beam forming optimization from (7) and (10) can be expressed as follows:

$$\boldsymbol{W}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{T}}^{K}=\boldsymbol{g}^{\mathrm{T}} \tag{11}$$

where $\boldsymbol{g}=(\gamma,0,0,\ldots,0)^{\mathrm{T}}$ is the vector with $M-1-N_{\mathrm{J}}$ dimensions, $\boldsymbol{C}_{\mathrm{T}}^{K}=\left[\boldsymbol{C}_{\mathrm{s}}^{(i)},\boldsymbol{C}_{\mathrm{n}}^{K}\right]$ is the $M\times(M-1-N_{\mathrm{J}})$ matrix. According to the constrained Lagrange multiplier method [6], the optimal weighted vector for the ith navigation satellite signal can be obtained.

$$\boldsymbol{W}_{\mathrm{T}}^{(i)}=\left[\boldsymbol{g}^{\mathrm{H}}\left(\boldsymbol{C}_{\mathrm{T}}^{(i)\mathrm{H}}\boldsymbol{R}_{xx}^{-1}\boldsymbol{C}_{\mathrm{T}}^{(i)}\right)^{-1}\boldsymbol{C}_{\mathrm{T}}^{(i)\mathrm{H}}\boldsymbol{R}_{xx}^{-1}\right]^{\mathrm{H}} \tag{12}$$

The vector $\boldsymbol{W}_{\mathrm{T}}^{(i)}$ makes the nulls formed to the direction of jamming and main beam formed to the ith navigation satellite signal, and the beam nulls are formed to the other satellite directions so as to further reduce the mutual interference among different satellite navigation signals.

After the completion of the optimal beam forming to a navigation satellite, the optimal weighted coefficient of the other satellite can be calculated in accordance with the same process. It is a parallel implementation. In this way, the best beam reception is achieved for all navigation satellites within the visible range.

5 Simulations

The simulation scene is set up as a UAV (Unmanned Aerial Vehicle) flying on space of Changsha, China, at 8 a.m. on September 1, 2017. The UAV is equipped with GPS nulling antenna array and INS. The position output by INS is east longitude 113°, north latitude 28° and height 602 m. The attitude parameters output by INS are heading angle 92.8°, pitch angle 1.5° and roll angle 1.3°. The UAV loads the GPS almanac data of August 20, 2017 downloaded from the internet. The angles of GPS signal arrival at the present time can be calculated from the almanac and INS data, as shown in Table 1. The PRN in the table represents the pseudo code number of the GPS satellite.

As shown in Table 1, the UAV has received signals from 10 GPS satellites at this time and place. There are two navigation jammers respectively from azimuth 358°, pitch angle 10°, and azimuth 200°, pitch angle 10° to radiate the large power suppression jamming signal. According to the traditional power inversion nulling algorithm shown in Eq. (4), the gain of the antenna beam after optimization is shown in Fig. 1.

Table 1 The direction of GPS satellite signals from the UAV view

No.	PRN	Azimuth (°)	Pitch angle (°)
1	2	304.6018	60.7886
2	5	224.9354	63.9567
3	6	357.5378	28.7272
4	12	139.9462	19.8054
5	13	81.1993	41.5944
6	15	111.3744	17.4211
7	19	57.1425	16.4409
8	20	172.6456	30.6036
9	25	176.2393	16.4805
10	29	226.9076	27.2697

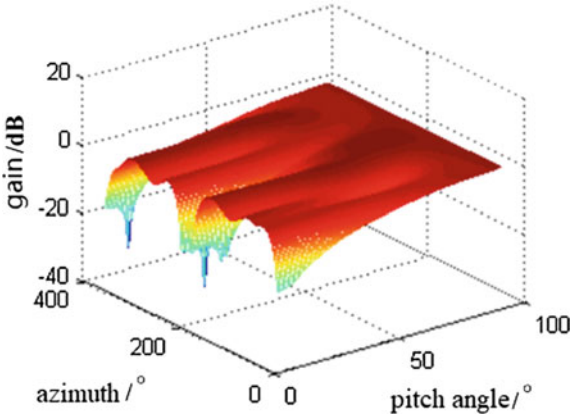


Fig. 1 Antenna beam gain diagram formed by conventional power inversion nulling algorithm

Table 2 Antenna gain for the direction of GPS satellite signals from the UAV view

No.	PRN	Antenna gain (PI algorithm) (dB)	Antenna gain (1st optimization) (dB)	Antenna gain (2nd optimization) (dB)
1	2	1.98	−5.04	−29.01
2	5	0.26	−0.21	−18.88
3	6	−6.25	4.76	4.93
4	12	1.45	4.08	−12.06
5	13	−1.12	2.17	−6.69
6	15	−0.55	5.40	−10.61
7	19	0.44	−5.54	−17.38
8	20	−4.37	6.79	−9.66
9	25	−13.13	3.65	−10.06
10	29	0.33	−5.73	−1.24

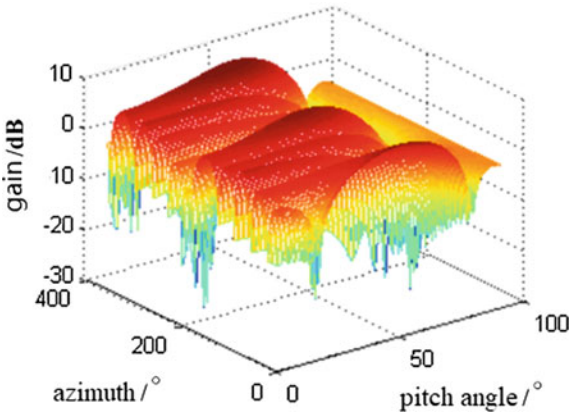


Fig. 2 Antenna beam gain diagram with main beam constraint

As shown in Fig. 1, the two nulls are accurately formed in the antenna beam pattern to the direction of two jammers. After the power inversion process, the antenna gains for each GPS satellite direction are shown in the third column of Table 2. The gains in some directions of GPS satellites decrease, for example, the gain of the GPS satellite with PRN = 6 is only −6.25 dB. The azimuth angle of one of the jammers is very close to that of the GPS satellite, so the traditional power inversion nulling algorithm leads to the decrease of signal reception performance in some directions at the same time. In order to solve this problem, the constraints are modified to the main beam forming constraint in direction of GPS satellite with PRN = 6, and then the antenna beam gain diagram is shown in Fig. 2.

After the first optimization, the antenna gains in the direction of each GPS satellite are shown in the fourth column of Table 2. Although the gain to direction of the GPS satellite with PRN = 6 is increased to 4.76 dB, the gains to direction of the other satellites have also been increased, such as the gains to direction of GPS

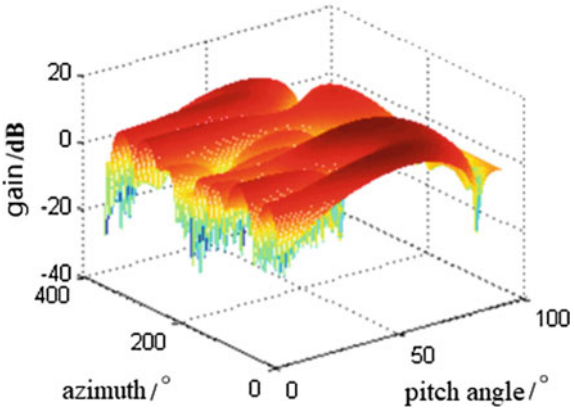


Fig. 3 Antenna beam gain diagram after two optimizations

satellites with PRN = 20, 15 and 12 are 6.79, 5.40, 4.08 dB etc. Since GPS satellite system is a code division multiple access system, there exists inter correlation interference between GOLD codes. In order to further optimize the receiving performance, null constraint conditions are applied to the direction of the other satellites by using the left freedom degree according to Eq. (11), and the second optimization is carried out. The antenna beam gain diagram is shown in Fig. 3.

After the second optimization, the antenna gains in the direction of each GPS satellite are shown in the fifth column of Table 2. The gain of the antenna is 4.93 dB in the direction of the GPS satellite with PRN = 6, and two nulls are formed in the direction of the jammers. At the same time, the antenna gains in the direction of other GPS satellites are lower than 0 dB. This greatly reduces the external jamming and the internal multiple access interference of the system itself, and improves the receiving performance. Similarly, the above two steps beam forming optimizations are carried out respectively for the other GPS satellite. The high performance can be achieved and it benefits the following position greatly.

6 Conclusions

In the application of satellite navigation, the traditional power inversion nulling algorithm may cause the antenna gain to decrease in the direction of the satellite, when the nulls are formed in the direction of the jamming with high-power. The problem has been solved in this paper. The current directions of visible navigation satellites are calculated in real time from data of the almanac and INS. For each satellite, the main beam pointing constraint and the null constraint of other satellites are independently added, and the weighted vector is obtained by the two steps optimization process. Not only the external jamming can be reduced, but also the multiple access interference of the satellite navigation system is reduced and the receiving performance is improved. The UAV GPS nulling antenna is used as an

example. The feasibility and effectiveness of the proposed method are verified by simulations. It is an important reference for the practical engineering application.

References

1. Teunissen P, Montenbruck O (2017) Handbook of global navigation satellite systems. Springer, Germany
2. Zhao X, Ma X, Qu W (2009) Application of power inversion arithmetic in anti-interference design of GPS receiver antenna. Telecommun Eng 49(7):45–48
3. Wu R, Sun Q, Hu T et al (2010) Realization of GPS anti-jamming real-time system based on power inversion algorithm. J Civ Aviat Univ China 28(3):45–48
4. Shi R, Deng K, Li Z et al (2014) Equivalence analysis on two kinds of power inversion models for zero-trip in antenna array. GNSS World China 39(4): 4–7, 15
5. He X, Hu X, Luo B (2015) Algorithm for BD/MIMU integrated navigation system. National Defense Industry, Beijing
6. Trees HLV (2002) Optimum array processing part IV of detection, estimation, and modulation theory. Wiley, USA, New York

A Novel Multipath Mitigation Method Based on Fast Orthogonal Search (FOS) for Short-Delay Multipath with Zero Doppler Shift Difference

Jun Mo, Zhongliang Deng, Buyun Jia, Xinmei Bian, Wen Liu and Yuan Sun

Abstract Multipath is the main factor affecting the positioning accuracy degradation in the Time and Code Division-Orthogonal Frequency Division Multiplexing (TC-OFDM) system. There is no effective mitigation method for short–delay multipath with zero Doppler shift difference in indoors. A novel multipath mitigation method is proposed to solve the problem, which is based on Fast Orthogonal Search (FOS). The candidate functions are created to represent the TC-OFDM reference correlation functions with several multipath delays and amplitudes, and the FOS tries to build a fitting model between the measured correlation functions and the selected candidate functions. Hence, the receiver can obtain accuracy direct and multipath estimations. The theoretical expressions of the proposed method are derived, and the multipath mitigation performances of n-Early Minus Late (nEML), Early Late Slope (ELS), Multipath Estimating Delay Lock Loop (MEDLL) are simulated and compared with the proposed method. The simulation results indicate that the performance of the proposed method is better than the classical techniques under the short-delay multipath scenarios with zero Doppler shift difference.

Keywords Multipath mitigation · FOS · TC-OFDM

1 Introduction

Global Navigation Satellite Systems (GNSS) have been widely used in our daily life. However, the GNSS signal availability and the positioning accuracy dramatically deteriorate in challenging environments, such as indoor environments and urban canyons [1, 2]. Recently, terrestrial radio positioning systems and their enhancements to the GNSS have been got increasing attention. This paper mainly

J. Mo (✉) · Z. Deng · B. Jia · X. Bian · W. Liu · Y. Sun
School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China
e-mail: mojun@bupt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_26

study a novel terrestrial radio positioning system called Time and Code Division-Orthogonal Frequency Division Multiplexing (TC-OFDM). The TC-OFDM system is based on the China Mobile Multimedia Broadcasting (CMMB) system, which multiplexes the CMMB signal and Pseudorandom Noise (PRN) codes in the same frequency band [3]. The positioning part of the TC-OFDM system is a direct-sequence spread spectrum code division multiple access (DSSS-CDMA) system employing binary phase shift keying (BPSK) modulation, and Ref. [3] describes the TC-OFDM system in detail. Compared with the GNSS, the TC-OFDM system has some potential advantages: the TC-OFDM signal transmission power is stronger and the frequency band is U Band, which contribute to better indoor propagation than the GNSS of L Band. Due to the reception sensitivity of the CMMB receiver is −95 dBm [4] and the CMMB system utilizes single frequency network for signal coverage [5], thus the TC-OFDM receiver demodulates PRN codes for high-sensitive positioning.

Multipath is the major factor affecting the positioning accuracy degradation in the TC-OFDM system, and the error caused by multipath can reduce the positioning accuracy of the receiver more than 60 m in severe cases. The multipath mitigation techniques in the positioning receiver can be divided into two categories: the multipath cancellation and the multipath estimation. The multipath cancellation technique mitigates the impact of multipath on the correlation function by improving the code loop discriminator, typically n-Early Minus Late (nEML) [6], Early Late Slope (ELS) [7]. nEML is the basic multipath mitigation technique, and its implementation is simple. The median-delay and long-delay multipath are mitigated by narrowing the correlator interval, so nEML is widely used in GNSS receivers. ELS utilizes the characteristics of the correlation curve slope to mitigate the impact of multipath, but it is susceptible to noise interference resulting large pseudorange errors or the tracking loop unlocked. Hence, there is no obvious breakthrough compared with the nEML. MEDLL is a typical multipath estimation technique [8, 9]. MEDLL uses a large number of correlators to detect the shape of the distorted correlation curve, and splits the synthetic curve based on the reference correlation functions, and obtains the optimal combination parameters of the direct and multipath signals (including amplitude, delay, phase and multipath number). However, it cannot mitigate the short multipath below 0.1 chip.

In indoors, the moving speed of the receiver is slow, and the received multipath Doppler is usually zero. The median-delay and long-delay multipath are mitigated by the above methods, so the short-delay multipath (below 0.25 chip) with zero Doppler shift difference is the main reason to reduce the positioning accuracy. This paper proposes a novel multipath mitigation method based on Fast Orthogonal Search (FOS) to solve the problem. The FOS is used to model the measured correlation function and obtain accurate direct and multipath signal estimations. The candidate functions are created to represent the TC-OFDM reference correlation functions with several multipath delays and amplitudes. The Gram-Schmidt orthogonalization is utilized to create a series of orthogonal functions with the candidate functions. The proposed method tries to build a fitting model between the

measured correlation function and the selected candidate functions. Simulation results show the performance of the proposed method is better than these classical techniques under the short-delay multipath scenarios with zero Doppler shift difference.

The following of this paper is organized as follows. Section 2 presents the multipath signal model. The proposed method based on FOS is introduced in Sect. 3. Section 4 presents the simulation results of the comparison between different multipath mitigation techniques. Finally, Sect. 5 concludes the work.

2 Multipath Signal Model

PRN codes are used for positioning in the TC-OFDM system. After the radio frequency module and the analog-digital conversion (ADC) module, the received direct signal from *i*th base station can be written as:

$$r^{(i)}(nT_s, \tau^{(i)}) = A^{(i)}c^{(i)}(nT_s - \tau^{(i)})\mathrm{e}^{j2\pi(f_{IF}+f_d^{(i)})nT_s+\varphi_0^{(i)}} + \omega^{(i)}(n) \tag{1}$$

where T_s denotes the sampling duration, $\tau^{(i)}$ is the propagation delay, $A^{(i)}$ is the amplitude, $\mathrm{c}^{(i)}$ is the PRN code, f_{IF} denotes the intermediate frequency, $f_d^{(i)}$ is the residual carrier frequency, $\varphi_0^{(i)}$ is the initial phase of the carrier and $\omega^{(i)}(n)$ is the additive white Gaussian noise (AWGN) with mean zero and variance, and $r^{(i)}(nT_s, \tau^{(i)}) = r^{(i)}(n, \tau^{(i)})$.

In the TC-OFDM system, multipath signals are usually the attenuation and delay of direct signals. Ignoring the navigation message, the received signal including multipath can be expressed as:

$$r_{multipath}^{(i)}(n) = r^{(i)}(n, \tau^{(i)}) + \sum_{k=1}^{K} \alpha_k^{(i)} r^{(i)}(n, \tau_k^{(i)}) \tag{2}$$

where K is the number of multipath signals, $\alpha_k^{(i)}$ is the attenuation of multipath signals, $\tau_k^{(i)}$ is the delay of multipath signals. The autocorrelation function of $r^{(i)}(n, \tau^{(i)})$ can be written as

$$R^{(i)}(\tau) = \frac{1}{NA^{(i)}} \sum_{n=0}^{N-1} r^{(i)}(n, \tau^{(i)}) c_{\mathrm{L}}^{(i)}(n) = \begin{cases} 1 - |\tau|/T_c & |\tau| \le T_c \\ 0 & other \end{cases} \tag{3}$$

where N is the number of the periodic PRN codes, $c_{\mathrm{L}}^{(i)}$ is the local replica code, and T_c is PRN code chip duration. When the multipath signal exists, the received signal correlation curve is distorted, and the distorted correlation curve is the superimposition of the direct signal and the multipath signal correlation curve, as shown in

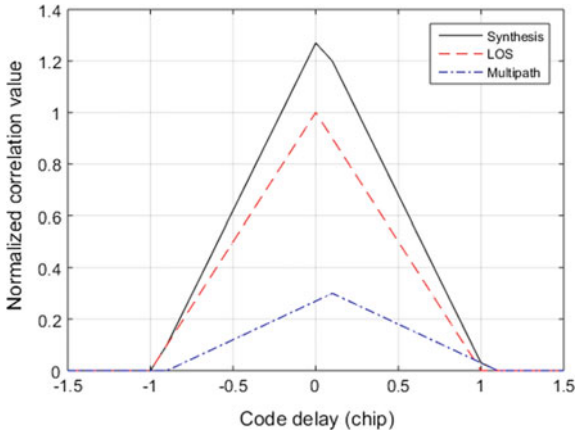


Fig. 1 The distorted correlation function under the unlimited bandwidth

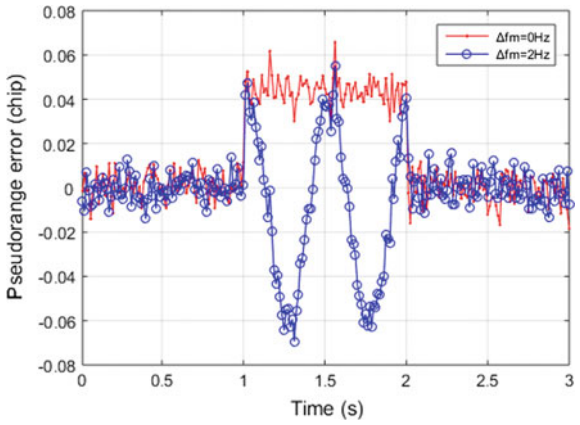


Fig. 2 Pseudorange error caused by the multipath Doppler shift difference

Fig. 1. Under the impact of multipath signals, $R^{(i)}(\tau)$ is distorted and the pseudorange error caused by the code discrimination increase. In addition, the multipath Doppler shift difference also distorts $R^{(i)}(\tau)$, resulting in pseudorange errors with a rectangular distribution, as shown in Fig. 2. In Fig. 2, when SNR = −20 dB, one multipath interference signal with $\alpha_1 = 0.7$ and $\tau_1 = 0.0625$ is added in the first 1–2 s. In the presence of the multipath Doppler shift difference, the multipath phase changes periodically with Δf_m. The distored curve of the direct and multipath signal will show periodic changes, and the adjustment of the code phase also changes periodically, eventually leading to periodic jitter pseudorange errors. Due to the multipath Doppler shift difference is usually zero in indoors, this paper mainly studies the short-delay multipath with zero Doppler shift difference.

3 A Novel Multipath Mitigation Method Based on FOS

The FOS algorithm has a lot of practical applications including spectral model estimations, time series analysis and nonlinear system control [10–12]. In this paper, the FOS is used in the multipath mitigation method to solve the short-delay multipath problem. The structure of the proposed method is shown in Fig. 3. The corresponding output $y(n)$ of all correlators is as input to the FOS model. The flowchart of the proposed method is shown in Fig. 4. The proposed method mainly consists of two parts: the construction of candidate functions and the selection of appropriate candidate functions. The Gram-Schmidt orthogonalization is utilized in the selection to build an accuracy multipath estimation model.

3.1 Candidate Functions

The first step of the proposed algorithm is to generate s series of candidate functions. The candidate functions is a set of TC-OFDM ideal autocorrelation functions $R_{ideal}(\tau)$ with several multipath delays and amplitudes.

$$P_v(n) = \alpha_v R_{ideal}(\tau_v) \tag{4}$$

where α_v and τ_v are the relative amplitude and the relative delay of one multipath with respect to the direct signal respectively. In practice, a series of P_v is usually generated after the acquisition is completed and stored in the corresponding memory. The corresponding multipath parameters can be defined as follows:

$$\begin{cases} -1 \le \alpha_v \le 1 \\ 0 \le \tau_v \le 2T_c \end{cases} \tag{5}$$

In this paper, the spacing of α_v is one tenth and the spacing of τ_v is the correlator spacing.

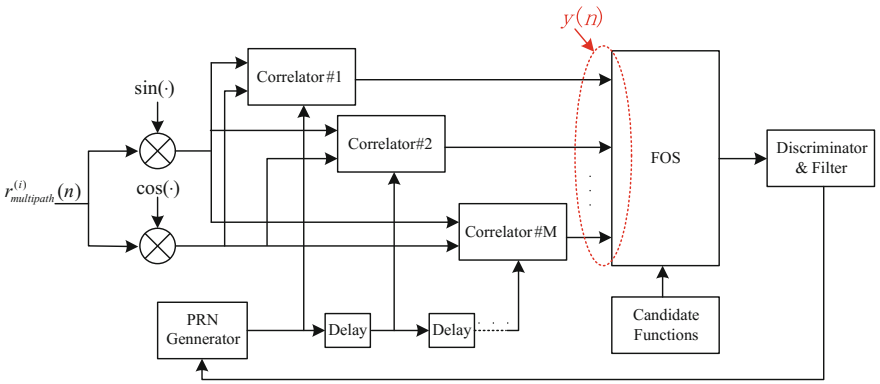


Fig. 3 The structure of the proposed method

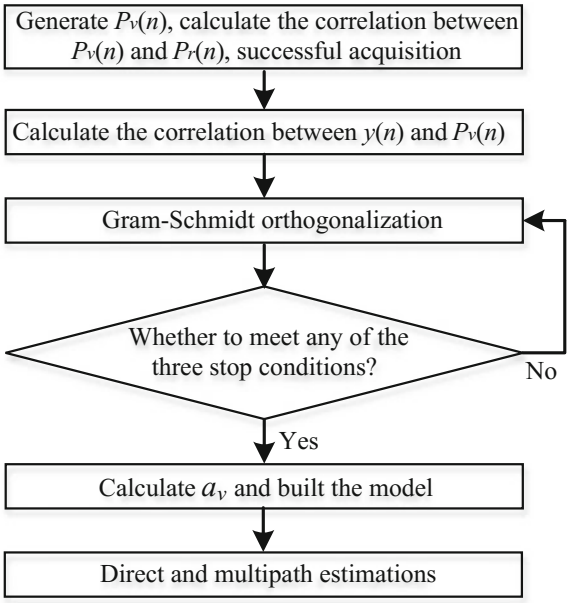


Fig. 4 The flowchart of the proposed method

3.2 Multipath Estimation

The main goal of the proposed method is to create a functional expansion of $y(n)$ by choosing the best basis functions from all candidate functions $P_v(n)$, in order to minimize the mean squared error (MSE) between the input and the functional expansion. $y(n)$ can be linear combinations of non-orthogonal candidate features of the following form:

$$y(n) = \sum_{v=0}^{V} a_v P_v(n) + \omega(n) \tag{6}$$

where V is the number of selected candidate functions, a_v are the weights of the functional expansion, and $\omega(n)$ is the model error between the estimation correlations and the measured corrlations.

In order to yield a unique solution for Eq. (6), the orthogonal transform is applied in the FOS. The new created functional expansion is given by

$$y(n) = \sum_{v=0}^{V} g_v w_v(n) + e(n) \tag{7}$$

where g_v are the weights of the orthogonal functions $w_v(n)$, $e(n)$ denotes the residual error. $w_v(n)$ is a set of orthogonal functions derived from $P_v(n)$ by utilizing the Gram Schmidt orthogonalization.

$$\overline{w_i(n)w_j(n)} = 0, \quad i \neq j \tag{8}$$

The calculation process is as follows.

$$\begin{aligned} w_1(n) &= P_1(n) = 1 \\ w_2(n) &= P_2(n) - \beta_{21} w_1(n) \\ &\vdots \\ w_v(n) &= P_v(n) - \sum_{r=1}^{v-1} \beta_{vr} w_r(n) \end{aligned} \tag{9}$$

And the the Gram-Schmidt coefficients β_{vr} is

$$\beta_{vr} = \frac{\overline{P_v(n)P_r(n)}}{\overline{w_r^2(n)}} \tag{10}$$

The FOS uses implicit orthogonal method for the convenient calculation, which only needs to calculate β_{vr} instead of calculating w_v point-by-point. In order to intuitively represent the relationship between the following equations, the intermediate functions $C(v, r)$ and $D(v, r)$ are designed.

$$\begin{cases} D(0,0) = 1 \\ D(v,0) = \overline{P_v(n)} \\ D(v,v) = \overline{w_v^2(n)} \\ D(v,r) = \overline{P_v(n)w_r(n)} = \overline{P_v(n)P_r(n)} - \sum_{i=0}^{r-1} \alpha_{ri} D(v,i) \end{cases} \tag{11}$$

$$\alpha_{vr} = \frac{\overline{P_v(n)P_r(n)}}{\overline{w_r^2(n)}} = \frac{D(v,r)}{D(r,r)} \tag{12}$$

$$\begin{cases} C(0) = \overline{y(n)} \\ C(v) = \overline{y(n)w_v(n)} = \overline{y(n)P_v(n)} - \sum_{r=0}^{v-1} \alpha_{vr} C(r) \end{cases} \tag{13}$$

$$g_v = \frac{C(v)}{D(v,v)} \tag{14}$$

where $v = 1, \ldots, V$; $r = 1, \ldots, v$.

At last, a_v can be found recursively using

$$a_v = \sum_{i=v}^{V} g_i u_i \tag{15}$$

where

$$u_i = \begin{cases} 1 & i = v \\ -\sum_{r=v}^{i-1} \beta_{ir} u_r & v < i \le V \end{cases} \tag{16}$$

In order to find an efficient model with few terms, each candidate function is fitted as the first candidate function in the orthogonal model. The candidate function that reduces the MSE of the model is chosen to be the first estimation in the model. Next, all the remaining candidate functions are tested as the second estimation in the model. Again, the estimation function with maximum reduction in MSE is next fitted in the orthogonal model.

The MSE of the orthogonal function expansion is

$$\overline{\varepsilon^2(n)} = \overline{y^2(n)} - \sum_{v=0}^{V} g_v^2 \overline{w_v^2(n)} \tag{16}$$

It then follows that the MSE reduction given by the vth candidate function is given by

$$Q_v = g_v^2 \overline{w_v^2(n)} = g_v^2 D(v,v) \tag{17}$$

The proposed method is stopped in one of the following three cases:

(1) certain maximum number of terms is fitted;
(2) the ratio of MSE to the mean squared value of the original signal is below a predetermined threshold;
(3) adding another term to the model reduces the MSE less than adding WGN.

In this paper, we assume that the total number of multipath components is no more than three, the direct signal has the shortest arrival time and the residual carrier Doppler shift can be correctly estimated by the carrier loop. Finally, the output of the proposed algorithm is the direct and multipath estimation.

4 Simulation and Analysis

In order to verify the performance of the proposed multipath mitigation method, the proposed method and other multipath mitigation algorithms are compared in two path static channels with different path amplitude values. Due to the short-delay multipath scenario with zero Doppler shift difference is generally the most challenging situation in indoors, the performance comparison of these algorithms mainly focuses on the scenarios. All the simulations are implemented by MATLAB R2015a, and repeated 100 times to obtain good statistical properties. The simulation profile is summarized in Table 1, and Fig. 5 shows the multipath error envelopes with different multipath delays for two path static channels. In Fig. 5, the curves above the abscissa correspond to the case when the reflected signal was in-phase with respect to the direct signal, whereas the curves below the abscissa correspond to the out-of-phase case. The proposed method outperforms other algorithms, and the performance of the MEDLL algorithm is better than both the nEML and ELS algorithm. In addition, the proposed method also has a better capability to mitigate the median-delay and long-delay multipath.

Then, we further study the effect of the short-delay multipath with zero Doppler shift difference on the pseudorange measurement of the proposed method. At the end of the first second, an in-phase multipath signal with delay 0.0625 chip and amplitude value 0.7, and the multipath signal lasts for one second. The remaining simulation parameters are shown in Table 1. Figure 6 shown the pseudorange error comparison under the short-delay multipath with zero Doppler shift deference. It can be seen that the short-delay multipath mitigation effect of the proposed method is obviously better than other algorithms, and the short-delay multipath mitigation effect of MEDLL is second.

Table 1 Simulation profile description

Parameter	Value
Channel model	Rayleigh fading channel
Path amplitude values	1 and 0.7
Path delay	0–1 chip
Path phase	0 or 180°
SNR	−20 dB
Coherent integration	25 ms
Code tracking loop bandwidth	1 Hz
Code tracking loop order	2nd order
Maximum number of FOS	3
nEML spacing	0.1 chip
ELS spacing	0.1 and 0.2 chip
Correlator spacing of the proposed method	0.1 chip

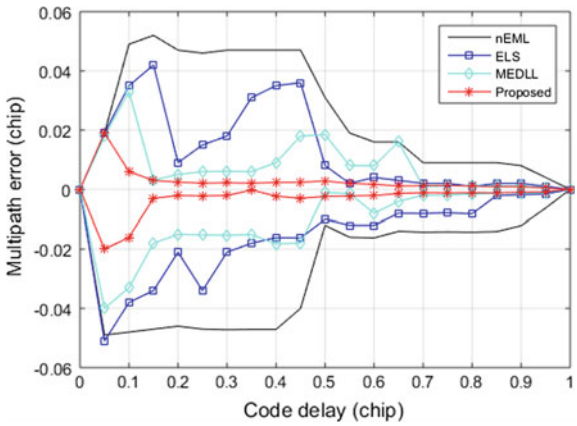


Fig. 5 Multipath error envelopes with different multipath delays for two static path channels

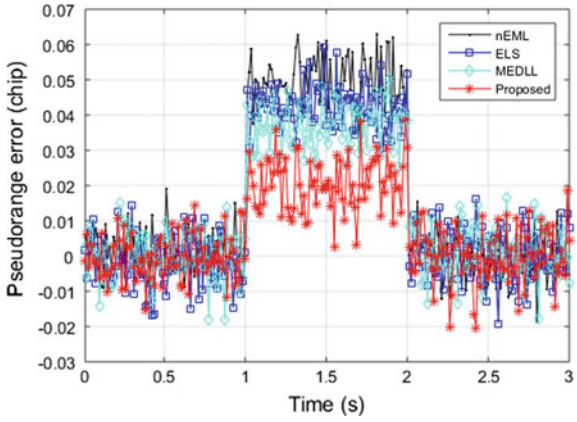


Fig. 6 Pseudorange error comparison under the short-delay multipath with zero Doppler shift difference

5 Conclusion

In this paper, a novel multipath mitigation method base on FOS is proposed, which can estimate the direct and multipath signals under the short-delay multipath with zero Doppler shift difference. The proposed method builds an accuracy estimation model to estimate direct and multipath signals in the TC-OFDM system. The candidate functions are created to represent the TC-OFDM reference correlation functions with several multipath delays and amplitudes, and the FOS tries to build the fitting model between the measured correlation function and the selected candidate functions. The simulation results demonstrate that the proposed method outperforms other algorithms under the short-delay multipath scenarios. Although the proposed algorithm can mitigate the impact of the multipath on the pseudorange measurement, it requires large computational loads. Therefore, our next step will focus on how to reduce the computational load of the FOS.

Acknowledgments Foundation project: The National Key Project of Research and Development Plan of China (2016YFB0502003).

References

1. Deng Z, Yu Y, Yuan X, Wan N, Yang L (2013) Situation and development tendency of indoor positioning. China Commun 10(3):42–55
2. Liang C et al (2014) Delay estimation for DVB-T signals in adverse multipath scenarios. In: 2014 International conference on localization and GNSS, Helsinki, Finland
3. Deng Z, Mo J, Jia B, Bian X (2017) An acquisition scheme based on a matched filter for novel communication and navigation fusion signals. Sensors 17(8):18
4. GY/T 220.7-2008 (2008) Mobile multimedia broadcasting (PR China) Part 7: technical specifications for receiving and decoding terminal
5. GY/T220.1-2006 (2006) Mobile multimedia broadcasting (PR China) Part 1: framing structure, channel coding and modulation for broadcasting channel
6. Van Dierendonck AJ, Fenton P, Ford T (1992) Theory and performance of narrow correlator spacing in a gps receiver. Navigation 39:265–283
7. Townsend BR, Fenton PC (1994) A practical approach to the reduction of pseudorange multipath errors in a ll gps receiver. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, Salt Lake City, USA, pp 143–148
8. Van Nee RDJ, Siereveld J, Fenton PC (1997) The multipath estimating delay lock loop: approaching theoretical accuracy limits. In: Proceedings of the national technical meeting of the institute of navigation, pp 207–215
9. Mohamed T, Aboelmagd N, Korenberg MJ, Kamel AM (2016) A new high-resolution gps multipath mitigation technique using fast orthogonal search. J Navig 69(4):794–814
10. Korenberg M, Paarmann L (1989) Applications of fast orthogonal search: time-series analysis and resolution of signals in noise. Ann Biomed Eng 17:219–231
11. Ahmed El-S, Aboelmagd N, Don M, Aini H (2012) Fast orthogonal search (FOS) versus fast Fourier transform (FFT) as spectral model estimations techniques applied for structural health monitoring (SHM). Struct Multidiscip Optim 45:503–513
12. McGaughey DR, Tarbouchi M, Nutt K, Chikhani A (2006) Speed sensorless estimation of ac induction motors using the fast orthogonal search algorithm. IEEE Trans Energy Convers 21:112–120

Design of Spaceborne GNSS-R Receiver Based on Multi-channel Main Frequency Parallel Rotation Algorithm

Xingyuan Han, Yanguang Wang, Lin Han, Bo Qu and Longlong Li

Abstract GNSS-R uses the reflected signals of navigation satellites to reflect the physical characteristics and parameters of the reflector. In this paper, a spaceborne GNSS-R receiver is designed, which is introduced from the aspects of system design, signal processing flow and method, test verification and so on. The receiver integrates the functions of direct signal navigation and positioning, reflection point calculation, and reflection signal fast processing algorithm. In this paper, the Newton iteration method is introduced for the first time to calculate the reflection point and deduced by trigonometric function. Compared with the angle bisector iteration method, the computational efficiency is greatly improved, and the average calculation time is reduced by more than 10 times. In the process of reflection signal processing, this paper adopts the algorithm based on the multi-channel main frequency parallel rotation, which improves the computational efficiency of reflected signal processing from code dimension and carrier dimension, the analysis shows that when using two main frequency parallel search, the calculation efficiency ratio is improved by about 3.88 times, compared with the parallel code search algorithm. Finally, this paper uses the receiver to receive the navigation satellite signal for reception test, the test results show that the spaceborne GNSS-R receiver can reliably generate the Doppler (DDM), which lays the foundation for the next step in carrying test onboard.

Keywords Spaceborne GNSS-R receiver · Direct signal navigation and positioning · Reflection point calculation · Reflection signal processing DDM

X. Han (✉) · Y. Wang · L. Han · B. Qu · L. Li
Academy of Space Electrical Information Technology,
No.504 in East Chang'an Street, Chang'an District,
Xi'an 710100, China
e-mail: beihang_hxy@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_27

1 Introduction

Global Navigation Satellite System-Reflection (GNSS-R) is a new type of remote sensing technology that uses the reflected GNSS navigation signal to invert the physical characteristics and parameters of the reflector [1]. Inversion using GNSS reflected signals is one of the current research hotspots in the field of GNSS applications.

Typical applications of GNSS-R include sea surface wind inversion [2], soil moisture detection [3], sea ice detection [4], sea surface height [5, 6], sea surface oil exploration [7], ESA and SSTL are committed to the development of spaceborne GNSS-R Receiver. NASA's successful launch of CY-GNSS for hurricane observations is a multi-joint, on-board receiver system. China's relevant universities and research departments [1] also carried out research on GNSS-R equipment technology, and made important progress.

The main function of the spaceborne GNSS-R receiver is to generate a time-delay Doppler (DDM), which provides observational data for parameter inversion. This paper designs a kind of spaceborne GNSS-R receiver, which introduces the receiver from the aspects of system design, signal processing flow and method, experiment verification and so on. The receiver integrates the functions of direct signal navigation and positioning, reflection point calculation, and reflection signal fast processing algorithm.

The traditional reflection point calculation uses the angle bisecting iteration method [8], the calculation method is simple, but the calculation efficiency is low, many times need to iterate more than 10,000, the time cost is large, and can't meet the reflection point real-time requirements. The traditional reflection signal processing method is mainly parallel code phase search method [9], because GNSS-R technology requires a wide frequency search range, generally about 5 $\sim$ 10 kHz, parallel code phase search method is not good at speed fast Search, which causes the search to be inefficient.

In order to solve the above two problems, this paper introduces the Newton iterative method to the reflection point calculation process for the first time, which greatly improves the computational efficiency compared with the angle bisector iteration method. In the reflection signal processing, the algorithm based on multi-channel main frequency parallel rotation is used in this paper to improve the computational efficiency of the reflected signal processing from the code dimension and the carrier dimension.

The successful development of the GNSS-R receiver, improved the algorithm of reflection point and reflection signal processing of the GNSS-R receiver, improved the computational efficiency and promoted the development of the satellite GNSS-R receiver technology. This laid the foundation for satellite carrying test in the next step.

2 System Design

The GNSS-R receiver is mainly composed of two antennas, two low noise amplifiers, one RF front end, one baseband processing board and so on. One of the antennas is installed on the zenith direction and the other antenna is installed on the opposite direction. Respectively through the low noise and RF front end, the direct signal and the reflected signal is amplified and converted to IF signal, and is outputted to the baseband processing board, the baseband processing board completes the processing of the direct signal and reflected signal. The system components of the spaceborne GNSS-R receiver are shown in Fig. 1.

3 Signal Processing Flow and Method

The signal processing of the spaceborne GNSS-R receiver includes direct signal processing and reflected signal processing. The main purpose of the direct signal processing is to complete the navigation and positioning, and to provide the current position information of the local satellite for the reflection point calculation. The direct signal processing includes capture, tracking, positioning, the process is shown in Fig. 2.

The reflection signal processing includes the satellite selection of reflection signal, the reflection point calculation and the reflection signal processing quickly, first selects a navigation satellite as the object of the reflected signal processing; secondly uses the navigation satellite, the local satellite, the earth's relative geometric relation to calculate the reflection point coordinate and reflection signal parameter; Finally, uses the reflection signal parameters to estimate the value of the rapid processing of reflected signals to generate DDM. Reflected signal processing process is shown in Fig. 3.

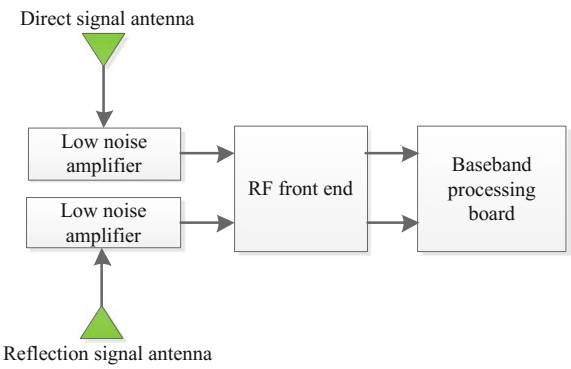


Fig. 1 The composition of the spaceborne GNSS-R receiver

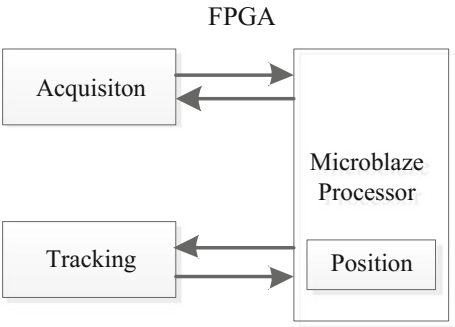


Fig. 2 Spaceborne GNSS-R receiver direct signal processing

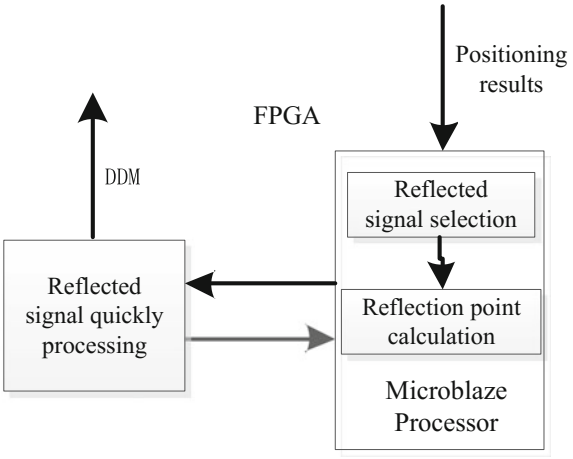


Fig. 3 Spaceborne GNSS-R receiver reflection signal processing

3.1 Reflecting Point Calculation Method Based on Newton Iteration

The reflection point, also known as the specular reflection point, is the point at which the distance between the transmitter and the earth's surface-receiver is the shortest point. The reflection point calculation refers to the process of determining the signal transmission time of the navigation satellite under the condition of knowing the receiving time of the reflection signal so as to determine the coordinates of the reflection point, the reflection signal path delay, and the reflection signal Doppler.

The core idea of the reflection point calculation is that the time interval between the launch time of the navigation satellite and the reception time of the local satellite is equal to the propagation delay of the reflected signal. The calculation process is as follows:

(1) initialize the launch time *t_tr* of the navigation satellite to the receiving time *t_re* of the local satellite;

(2) Using the ephemeris data of the navigation satellite, calculate the inertial coordinate *X_sat* of the navigation satellite at the time of transmission *t_tr*, As shown in point T in Fig. 4; locate the inertia coordinate *X_re* of the local satellite at the receiving time *t_re* by direct signal navigation, As shown in point R in Fig. 4; Calculate $\left|\overrightarrow{RT}\right|$, $\left|\overrightarrow{OR}\right|$, $\left|\overrightarrow{OT}\right|$, and then use the cosine theorem to solve $\angle ROT = \theta$;

(3) In ΔORP, presumed $\angle POR = \alpha$, we can use the cosine theorem to obtain: $\left|\overrightarrow{PR}\right| = \sqrt{\left|\overrightarrow{OP}\right|^2 + \left|\overrightarrow{OR}\right|^2 - 2\left|\overrightarrow{OP}\right|\left|\overrightarrow{OR}\right|\cos\alpha}$, which $\left|\overrightarrow{OP}\right|$ is equal to the radius of the earth R_e. Presumed $\angle MPR = \varphi$, we can use the sine theorem to obtain $\frac{\left|\overrightarrow{PR}\right|}{\sin\alpha} = \frac{\left|\overrightarrow{OR}\right|}{\sin\varphi}$.

(4) in ΔOTP, presumed $\angle POT = \beta$, and there is $\beta = \theta - \alpha$, we can use the cosine theorem to obtain: $\left|\overrightarrow{PT}\right| = \sqrt{\left|\overrightarrow{OP}\right|^2 + \left|\overrightarrow{OT}\right|^2 - 2\left|\overrightarrow{OP}\right|\left|\overrightarrow{OT}\right|\cos\beta}$, According to the principle of specular reflection, there is $\angle MPT = \angle MPR = \varphi$, and we can use the sine theorem to obtain: $\frac{\left|\overrightarrow{PT}\right|}{\sin\beta} = \frac{\left|\overrightarrow{OT}\right|}{\sin\varphi}$;

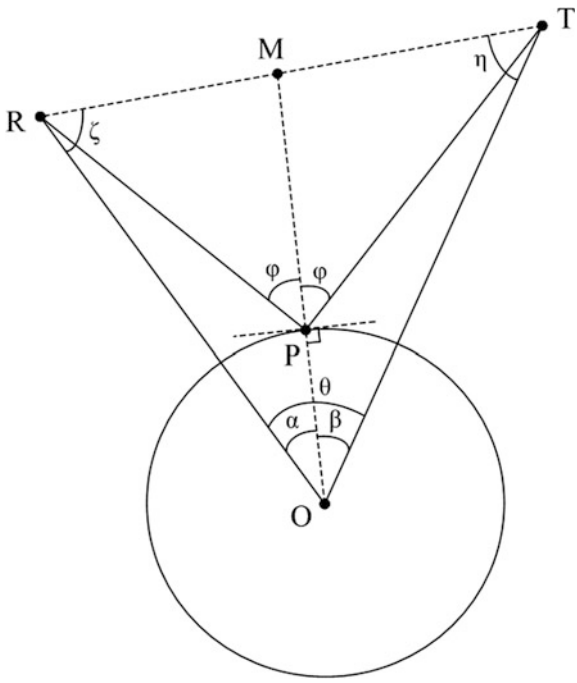


Fig. 4 Reflecting point calculation diagram based on newton iteration method

(5) Combine steps (3) and (4), then $\sin\varphi = \frac{\left|\overrightarrow{OR}\right|\sin\alpha}{\left|\overrightarrow{PR}\right|} = \frac{\left|\overrightarrow{OT}\right|\sin\beta}{\left|\overrightarrow{PT}\right|}$, define the function $f(\alpha) = \left|\overrightarrow{OR}\right|\left|\overrightarrow{PT}\right|\sin\alpha - \left|\overrightarrow{OT}\right|\left|\overrightarrow{PR}\right|\sin\beta$, $\beta = \theta - \alpha$ will be substituted into $f(\alpha)$, and derivating $f(\alpha)$:

$$f'(\alpha) = \left|\overrightarrow{OR}\right|\left[\left|\overrightarrow{PT}\right|\cos\alpha - \frac{\left|\overrightarrow{OP}\right|\left|\overrightarrow{OT}\right|\sin(\theta-\alpha)\sin\alpha}{\left|\overrightarrow{PT}\right|}\right] + \left|\overrightarrow{OT}\right|\left[\left|\overrightarrow{PR}\right|\cos(\theta-\alpha) - \frac{\left|\overrightarrow{OP}\right|\left|\overrightarrow{OR}\right|\sin\alpha\sin(\theta-\alpha)}{\left|\overrightarrow{PR}\right|}\right]$$

(6) Newton iteration formula can be written as: $\alpha_{k+1} = \alpha_k + \frac{f(\alpha_k)}{f'(\alpha_k)}$, Set the iteration of the initial value $\alpha_0 = \frac{\theta}{2}$, when $\frac{|\alpha_{k+1}-\alpha_k|}{\alpha_k} \le 10^{-10}$ or the number of iterations $N_\alpha \ge 100$, stop the iteration, assign $\hat{\alpha} = \alpha_{k+1}$.

(7) $\hat{\alpha}$ will be substituted into steps (3) and (4), then the reflection path delay $\hat{\tau} = \frac{\left|\overrightarrow{PR}\right| + \left|\overrightarrow{PT}\right|}{c}$, where c is the speed of light, modify the navigation satellite launch time $t\_tr = t\_re - \hat{\tau}$;

(8) repeat the steps (2) $\sim$ (7) of the process, if $\frac{|\hat{\tau}_{k+1}-\hat{\tau}_k|}{\hat{\tau}_k} \le 10^{-10}$ or the number of iterations $N_\tau \ge 5$, then stop the iteration, to solve this $\alpha = \hat{\alpha}$, $\tau = \hat{\tau}_{k+1}$;

(9) In ΔORM and ΔOTM, the sine theorem can be used: $\frac{\left|\overrightarrow{RM}\right|}{\sin\alpha} = \frac{\left|\overrightarrow{OR}\right|}{\sin\angle OMR}$, $\frac{\left|\overrightarrow{TM}\right|}{\sin\beta} = \frac{\left|\overrightarrow{OT}\right|}{\sin\angle OMT}$, assign $\lambda = \frac{\left|\overrightarrow{RM}\right|}{\left|\overrightarrow{TM}\right|} = \frac{\left|\overrightarrow{OR}\right|\sin\alpha\sin\angle OMT}{\left|\overrightarrow{OT}\right|\sin\beta\sin\angle OMR} = \frac{\left|\overrightarrow{OR}\right|\sin\alpha}{\left|\overrightarrow{OT}\right|\sin\beta}$, then the inertia coordinates of point M is $X\_m = \frac{X\_re + \lambda X\_sat}{1+\lambda}$, so the inertia coordinates of point P can be obtained: $X\_p = \frac{X\_m}{|X\_m|}\left|\overrightarrow{OP}\right|$;

(10) The reflected signal Doppler can be obtained indirectly by dividing the reflected signal path by a differential, that is $\frac{|\tau(t_1)-\tau(t_0)|}{t_1-t_0} = -\frac{\delta f}{f_0}$, then the reflected signal Doppler $\delta f = -\frac{|\tau(t_1)-\tau(t_0)|}{t_1-t_0} f_0$, the coordinates of reflection point X_p, the reflection signal path delay τ, the reflected signal Doppler δf all solved.

In this paper, we use numerical method of angular bisector and Newton iteration to test the data points of navigation satellite's five groups of inertial coordinates respectively. The data of coordinate points are shown in Table 1 (the Z coordinate is always zero). Regardless of changes in satellite position during reflection signal propagation, the number of iterations and calculation time are shown in Table 2.

Table 1 Five sets of coordinate data for the local satellite and the navigation satellite

Coordinate (m)	Group no					
	Group 1		Group 2	Group 3	Group 4	Group 5
Navigation satellite	X	42000000	0	23002175	18781197	13280312
	Y	0	42000000	13280312	18781197	23002175
local satellite	X	7378137				
	Y	0				

Table 2 Number of iterations and computation time of two methods

Iterative performance			Group no				
			Group 1	Group 2	Group 3	Group 4	Group 5
Angle bisector method	Number of iterations		2246	90171	8213	13139	24140
	Calculating time (s)		0.010452	0.294304	0.029697	0.046373	0.082954
Newton method	Number of iterations		2	6	5	5	5
	Calculating time	Time(s)	0.004717	0.004627	0.004579	0.005082	0.004787
		Multiple reduced	2.2158	63.6058	6.4855	9.1250	17.3290

It can be seen from Table 2 that the number of iterations of the Newton iteration method is about 5 times, the calculation time is about 0.005 s, reduced by 2.2 $\sim$ 63.6 times, which is 10 times lower than that of the angle bisector iteration method above, greatly improving the computational efficiency.

3.2 Reflected Signal Processing Algorithm Based on Multi—Channel Main Frequency Parallel Rotation

GNSS-R technology requires a narrow range of chip search, only about dozens of chips; frequency search range is wide, generally about 5 $\sim$ 10 kHz. The traditional parallel code phase search method can calculate the correlation values of all the code phases, and most of the calculation results need to be discarded, the calculation efficiency is limited, moreover, the method is not good at fast search of the frequency dimension, resulting in low search efficiency. In this paper, based on parallel code phase algorithm, the reflected data is block-correlated, which avoids the FFT calculation of excessive points. In parallel, the multi-channel main frequencies are searched and transformed respectively. The correlation value avoids a large number of carrier multiplication operations and improves the computational efficiency of the reflected signal processing from the code dimension and the carrier dimension.

The fast processing algorithm of the reflected signal used in this paper mainly includes two parts: one is based on the DBZP frequency domain convolution and the main frequency rotation transformation method [10]; the second is based on the multi-channel main frequency parallel search method as shown in Fig. 5, The block and zeroing operations in the DBZP algorithm ensure that the pseudo-code phase search is sufficient and the repeated calculation is avoided. As shown in Fig. 6, the main frequency rotation transform reduces the computational complexity of frequency search, and accelerates the calculation speed of frequency dimension search through multi-channel main frequency parallel search.

According to the Doppler frequency characteristics of the reflected signal, the required resolution of the DDM and the resources of the FPGA determines the Doppler frequency range F_d, the main frequency parallel search channel number N, the frequency rotation transformation step F_n and the Doppler frequency resolution F_m, F_d need to be an integer multiple of N * F_n, that the number of the main frequency search is $F_d/(N * F_n)$, each time parallel search N main frequency; F_n needs to be an integral multiple of F_m, indicating that a single main frequency needs to perform frequency rotation operations of (F_n/F_m) times.

Compared with the traditional single-channel parallel code phase algorithm, the fast algorithm reduces the code phase search range by tens of times, reduces the number of frequency search by F_n/F_m, and increases the number of parallel search branches by N times.

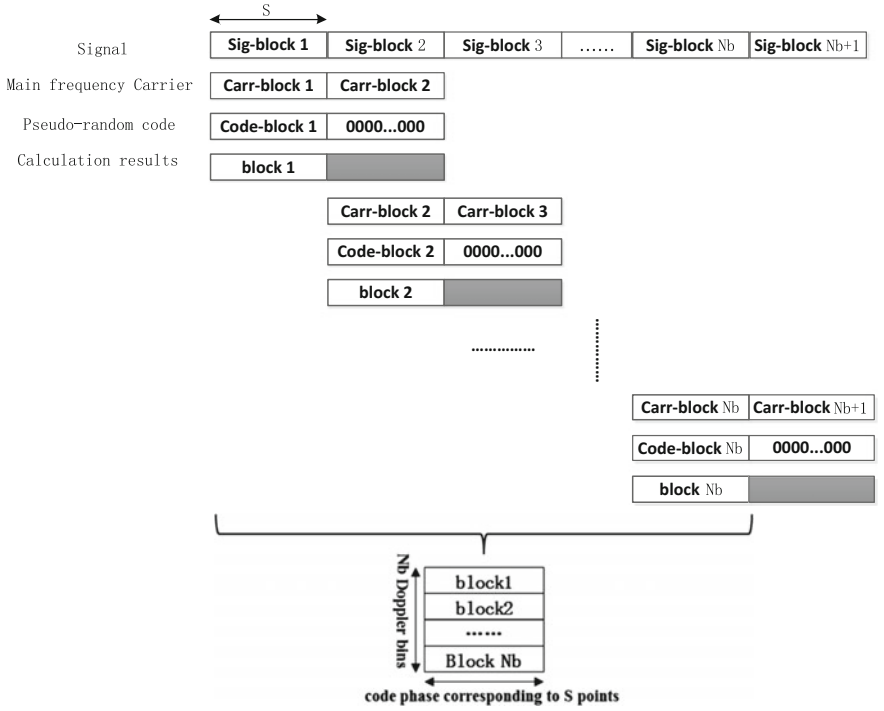


Fig. 5 DBZP correlation algorithm

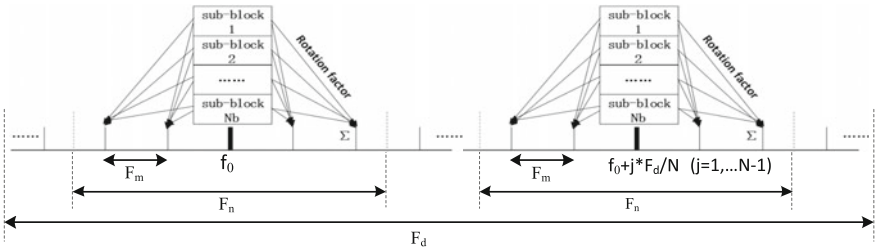


Fig. 6 Multi—channel main frequency parallel search method

Table 3 The computational complexity of the two algorithms

Algorithm		Complex additions	Complex multiplications	Total calculation
PCPS		7.3594×10^5	4.0791×10^5	1.1438×10^6
Fast algorithm	Two branches	4.2931×10^5	1.5974×10^5	5.8905×10^5
	Each branch	2.1466×10^5	7.987×10^4	2.9453×10^5
Improvement factor	Two branches	1.7142	2.5536	1.9418
	Each branch	3.4284	5.1072	3.8835

Corresponding to the design parameters of the GNSS-R receiver in Table 5, this paper compares the computational complexity of the fast algorithm and the parallel code phase algorithm. The results are shown in Table 3.

As can be seen from Table 3, compared with the parallel code phase algorithm, the computational complexity of the two branch main frequency parallel rotation algorithm is reduced by 1.94 times, the computational complexity of each branch is reduced by 3.88 times, effectively reducing the computational complexity and accelerating the calculation speed.

4 Experimental Verification

4.1 Simulation Analysis

In this paper, we use the fast algorithm to analyse the L1 C/A reflected signal data of No. 22 GPS satellites collected by the UK Disaster Monitoring Satellite (UK-DMC) in the Hawaiian Sea at 07:54:53 on November 16, 2004. Simulation parameters shown in Table 4, the simulation results shown in Fig. 7.

As can be seen from Fig. 7, the shape of the DDM is an arched shape, and the entire arch has a tailing of about 20 chips in the code dimension, which is consistent

Table 4 Simulation parameters

Data sampling rate f_s	5.714 MHz
Coherent integration time	6000 sampling points/f_s
Non-coherent accumulation times	1000
Main frequency parallel search path number N	1
Doppler frequency range F_d	±5 kHz
Frequency rotation step size F_n	500 Hz
Doppler frequency resolution F_m	100 Hz
Number of sub-blocks N_b	6
Sub-block length S	1000

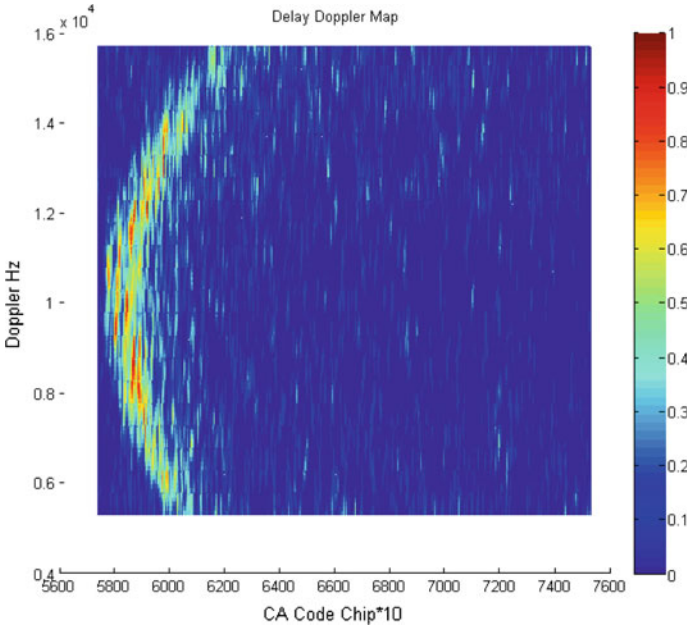


Fig. 7 Delay-doppler mapping

with the shape of the DDM described in some literatures, indicating that the fast algorithm used in this paper can effectively generate DDM.

4.2 Test Verification

At present, the GNSS-R receiver has completed the development of the engineering prototype and has the function of quick processing of the reflected signal.

Table 5 Spaceborne GNSS-R receiver design parameters

Data sampling rate f_s	5.0 MHz
Coherent integration time	4992 sampling points/f_s
Non-coherent accumulation times	1000 ms
Main frequency parallel search path number N	2
Doppler frequency range F_d	±10 kHz
Frequency rotation step size F_n	100 Hz
Doppler frequency resolution F_m	25 Hz
Number of sub-blocks N_b	39
Sub-block length S	128

The receiver adopts two main frequency parallel search design. The main design parameters are shown in Table 5. In this paper, L1 CA direct signal analog reflection signal, the signal power is about −130 dbm, with the receiver to test, the test results shown in Fig. 8.

It can be seen from Fig. 8 that the shape of the direct signal DDM is a spot and the spot has a slope from the lower left to the upper right because the carrier Doppler frequency has an effect on the code Doppler, so that the code phase changes. The results of the test in Fig. 8 are consistent with the results of the direct signal capture, indicating that the GNSS-R receiver has the exact function of generating DDM.

5 Conclusion

This paper designs a kind of spaceborne GNSS-R receiver, which introduces the system design, signal processing flow and method, and test verification.

The receiver integrates the functions of direct signal navigation and positioning, reflection point calculation, and reflection signal fast processing algorithm. In this paper, the Newton iteration method is introduced for the first time to calculate the reflection point and deduced by trigonometric function. Compared with the angle bisector iteration method, the computational efficiency is greatly improved, and the average calculation time is reduced by more than 10 times. In the process of reflection signal processing, this paper adopts the algorithm based on the multi-channel main frequency parallel rotation, which improves the computational efficiency of reflected signal processing from code dimension and carrier dimension, the analysis shows that when using two main frequency parallel search, the calculation efficiency ratio is improved by about 3.88 times, compared with the parallel code search algorithm.

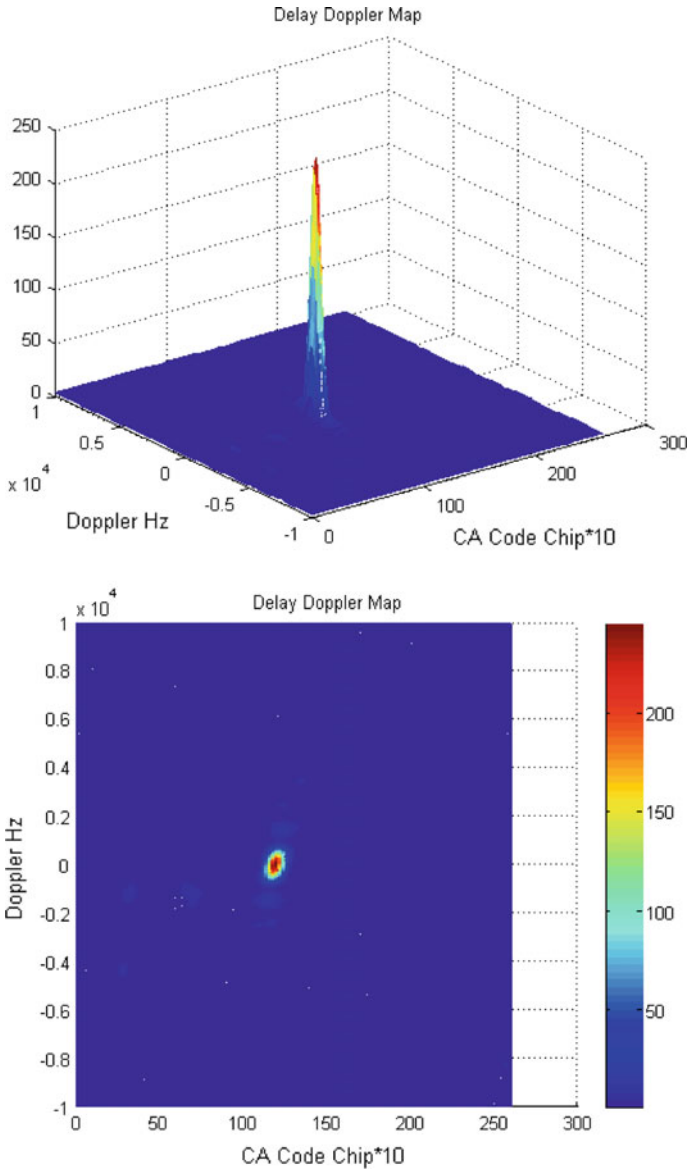


Fig. 8 Delay-doppler mapping: (Upper) oblique view, (bottom) top view

Finally, this paper uses the receiver to receive the GPS L1CA navigation signal. The test results show that the GNSS-R receiver can reliably generate the Doppler (DDM), which lays the engineering foundation for the next step in carrying test onboard.

References

1. 杨东凯 (2012) GNSS反射信号处理基础与实践. 电子工业出版社
2. Zavorotny VU, Voronovich AG (2000) Scattering of GPS signals from the ocean with wind remote sensing application. IEEE Trans Geosci Remote Sens 38(2):951–964
3. Rodriguez-Alvarez N, Bosch-Lluis X, Camps A et al (2009) Soil moisture retrieval using gnss-r techniques: experimental results over a bare soil field. IEEE Trans Geosci Remote Sens 47(11):3616–3624
4. Gleason S (2010) Towards sea ice remote sensing with space detected gps signals: demonstration of technical feasibility and initial consistency check using low resolution sea ice information. Remote Sens 2(8):2017–2039
5. Hajj GA, Zuffada C (2003) Theoretical description of a bistatic system for ocean altimetry using the GPS signal. Radio Sci 38(5):10–11
6. Martin-Neira M, D′Addio S, Buck C et al (2011) The PARIS ocean altimeter in-orbit demonstrator. IEEE Trans Geosci Remote Sens 49(6):2209–2237
7. Li C, Huang W (2013) Simulating GNSS-R delay-doppler map of oil slicked sea surfaces under general scenarios. Prog Electromagnet Res B 48(48):61–76
8. 桂文军 (2014) 基于导航卫星反射信号的海背景目标成像方法研究].华中科技大学硕士论文, 5
9. 谢钢 (2013) GPS原理与接收机设计. 电子工业出版社
10. 韩琳,蒙艳松,等 (2017) 一种基于DBZP的GNSS-R信号快速处理算法 第八届中国卫星导航学术年会, CSNC

Research of GNSS Spoofer Localization Using Information Fusion Based on Particle Filter

Shunshun Shang, Hong Li and Mingquan Lu

Abstract A spoofer localization algorithm using information fusion based on the particle filter (PF) is proposed. There are many researches about the jammer localization, but less researches about the spoofer localization. The jamming signal, whose structure is unknown to us, usually is above the thermal noise. According to the effect of jamming signal bandwidth on the measuring of time and frequency, we cannot obtain the precise measurements of TDOA and FDOA simultaneously. Thus, we usually cannot implement the information fusion of time and frequency in the jammer localization. Besides, we cannot locate the spoofer directly with the method of jammer localization, because the spoofing signal is usually below the noise. However, we can obtain the precise measurements of time and frequency of the spoofing signal at the same time because the spoofing signal needs to disguise as the true GNSS signal. Based on this, we propose a localization algorithm using the information fusion of TDOA and FDOA. We couple the position and velocity together by the PF, which will not introduce the linearization error. By the analysis of the CRLB, we find the information fusion can improve the system performance. To verify the algorithm, we simulate and analyze the results of PF and the weighted least squares (WLS). Compared to the WLS without information fusion, the PF has a better performance of locating and tracking, which can work properly under the circumstances where the spatial layout is poor, the number of base stations is small and the measuring errors are large.

Keywords GNSS · Information fusion · Particle filter · Localization
CRLB · TDOA · FDOA

S. Shang · H. Li (✉) · M. Lu
Department of Electronic Engineering, Tsinghua University,
100084 Beijing, China
e-mail: lihongee@tsinghua.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_28

1 Introduction

People pay more attention to the security issues of GNSS along with the devices based on GNSS increasing. A signal transmitter with low power can spoof the GNSS receiver because of the weakness of GNSS signal. There are many spoofing techniques which can mislead the receiver, such as self-consistent attack, nulling attack, estimate-and-replay attack and so on [1]. We need to detect and locate the spoofer in time to defense against the attack of spoofer. However, there are few researches on the localization of GNSS spoofer, although there are many researches on the localization of jammer and detection of spoofer. With regard to the localization of jammer, the measurements of time and frequency are contradictory on the level of the signal bandwidth, as a result of which, they cannot both be measured precisely at the same time. Besides, the spoofer cannot be located by the way of jammer localization, since the power of signal is below the noise. However, we can easily obtain the time, frequency, power, strength and phase of the spoofer signal because the signal needs to disguise as the true GNSS signal. After obtaining the measurements, we can locate the spoofer by the passive location techniques, such as time of arrival (TOA), time difference of arrival (TDOA), received signal strength, difference of received signal strength, frequency difference of arrival (FDOA), direction of arrival and so on [2–5].

We propose a GNSS spoofer location method, which is based on measuring TDOA and FDOA simultaneously and different from the jammer location. In order to locate and track the spoofer, we couple the measurements of time and frequency together by the particle filter (PF), which enables the estimations of position and velocity to improve each other. The main works of this paper are as follows.

- We establish the GNSS spoofer localization system. It can obtain the precise measurements of TDOA and FDOA at the same time, which makes the measuring information fusion by the PF practicable.
- We deduce the Cramer-Rao lower bound (CRLB) of position and velocity, based on which, we analysis and find that the information fusion can improve the estimated performance of the system. Besides, we prove the errors of position and velocity are only related to the measuring errors and the difference of geometric matrix when the spoofer is far from the base stations, which is about more than 50 meters. Also we prove the difference of geometric matrix has the same effect on the position and velocity errors. The CRLB is suitable for both 2-D space and 3-D space.
- We implement the particle filter with information fusion and weighted least squares without information fusion. After that, we compare and analyze their performances in different scenarios.

The rest of the paper is organized as follows. Section 2 proposes the model of the GNSS spoofer localization system. Section 3 deduces the CRLB of position and velocity. Section 4 provides the implementation details of PF. Section 5 simulates the performance of PF and WLS. Section 6 concludes the paper.

2 System Model

2.1 Model of the Spoofer and Base Stations

Firstly, we need to establish the model of the spoofer and base stations. We suppose there are N static base stations at the positions $\boldsymbol{p}_i$ $(i = 1, \ldots, N)$ used to locate the spoofer with a speed of $\boldsymbol{v}$ at the position $\boldsymbol{p}$. The unit vector $\boldsymbol{a}_i$ points from the i-th base station to the spoofer.

Figure 1 shows the configuration of the spoofer and base stations in the 2-D space. We need to obtain the estimations of position $\hat{\boldsymbol{p}}$ and velocity $\hat{\boldsymbol{v}}$ based on the known position of base stations and the measuring data.

2.2 The Spoofer Localization System

As shown in the Fig. 2, the spoofer localization system consists of three parts. The first part is about the spoofing detection. There are many techniques to detect the spoofer such as power monitoring, distribution analysis of the correlator output, code and phase rate consistency check, receiver autonomous integrity monitoring and so on [1, 6]. The second part is about feature extraction of the spoofer signal, which is about the time and frequency measurements in this paper. The time or frequency of the spoofer signal could be changed deliberately. So we need to eliminate the common parts which are brought in by the spoofer before we locate

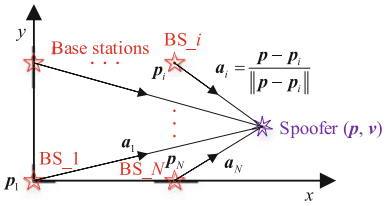


Fig. 1 Configuration of spoofer and base stations

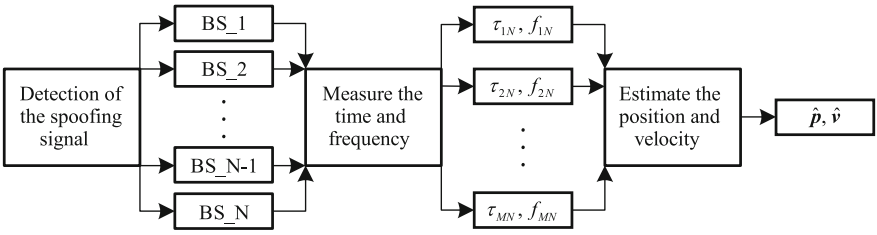


Fig. 2 Framework of spoofer localization

the spoofer. In Fig. 2, τ_{iN} $(i = 1, \ldots, M)$ is the difference between the TOA of the i-th base station and the N-th base station. The difference of FOA is f_{iN} $(i = 1, \ldots, M)$, $M = N - 1$. The third part is about the estimating of position and velocity. In the jammer location, we cannot get both of the precise measurements of time and frequency at the same time because of the signal bandwidth. Different from this, we can obtain the precise measurements of time and frequency simultaneously because of the special structure of the spoofer signal. Thus, we can locate the spoofer by TDOA and FDOA. Because of the complicated linear relation in the measuring equations, we adopt the PF to reduce linearization error. As comparison, we also implement the WLS [7].

2.3 Measurement Model of TDOA and FDOA

During the feature extracting of the spoofer signal, we first measure the TOA and FOA of the spoofer signal. Then we get the TDOA and FDOA after subtracting the measurements of the N-th base station from the TOA and FOA. We model the measuring error vessctor of TOA $\boldsymbol{w}_{TOA} \in \mathbb{R}^{N\times 1}$ and the measuring error vector of FOA $\boldsymbol{w}_{FOA} \in \mathbb{R}^{N\times 1}$ as independent Gaussian distribution. That is to say, $\boldsymbol{w}_{TOA} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_{TOA}^2 \boldsymbol{I}_N\right), \boldsymbol{w}_{FOA} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_{FOA}^2 \boldsymbol{I}_N\right)$ in which $\boldsymbol{I}_N$ is an N-order identity matrix. If the measuring error vectors of TDOA and FDOA are $\boldsymbol{w}_{TDOA}$ and $\boldsymbol{w}_{FDOA}$, then their distributions can be given as follows

$$\boldsymbol{w}_{TDOA} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_{TOA}^2(\boldsymbol{I}_M + \boldsymbol{E}_M)\right), \boldsymbol{w}_{FDOA} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_{FOA}^2(\boldsymbol{I}_M + \boldsymbol{E}_M)\right) \tag{1}$$

where $\boldsymbol{E}_M$ is a M-order square matrix with all elements are 1. We can see from Eq. (1) that the elements in $\boldsymbol{w}_{TDOA}$ or $\boldsymbol{w}_{FDOA}$ are not independent.

We need to establish the measuring model for the deduction of CRLB and implementation of PF and WLS. The measuring vector $\boldsymbol{y} \in \mathbb{R}^{2M\times 1}$ consists of measurements of TDOA with mean vector $\boldsymbol{\mu}_{TDOA}$ and FDOA with mean vector $\boldsymbol{\mu}_{FDOA}$. Thus, the mean vector of $\boldsymbol{y}$ can be expressed as $\boldsymbol{\mu} = \left(\boldsymbol{\mu}_{TDOA}^T, \boldsymbol{\mu}_{FDOA}^T\right)^T$. Based on Eq. (1), we have $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$ where $\boldsymbol{C}$ is a block diagonal matrix as below

$$\boldsymbol{C} = \begin{pmatrix} \sigma_{TOA}^2(\boldsymbol{I}_M + \boldsymbol{E}_M) & \\ & \sigma_{FOA}^2(\boldsymbol{I}_M + \boldsymbol{E}_M) \end{pmatrix} \tag{2}$$

The measuring equations are

$$y^{(i)} = \tau_{iN} = \mu_{TDOA}^{(i)} + w_{TDOA}^{(i)}, y^{(i+M)} = f_{iN} = \mu_{FDOA}^{(i)} + w_{FDOA}^{(i)} \tag{3}$$

where $y^{(i)}, i = 1, \ldots, M$ is the i-th element of $\boldsymbol{y}$, $w_{TDOA}^{(i)}$ and $w_{FDOA}^{(i)}$ are the i-th element of $\boldsymbol{w}_{TDOA}$ and $\boldsymbol{w}_{FDOA}$, respectively. Besides, $\mu_{TDOA}^{(i)}$ and $\mu_{FDOA}^{(i)}$ are the theoretical results of the i-th TDOA and FDOA, which go as below

$$\mu_{TDOA}^{(i)} = \|\boldsymbol{p} - \boldsymbol{p}_i\| - \|\boldsymbol{p} - \boldsymbol{p}_N\|, \mu_{FDOA}^{(i)} = f_0 c^{-1}\left((\boldsymbol{v}_i - \boldsymbol{v})^T \boldsymbol{a}_i - (\boldsymbol{v}_N - \boldsymbol{v})^T \boldsymbol{a}_N\right) \quad (4)$$

In Eq. (4), the unit of $\mu_{TDOA}^{(i)}$ is in meter. The f_0 is the center frequency of the spoofer signal, which is 1575.42 MHz for the C/A signal. The c is the speed of signal and $\boldsymbol{v}_i$ is the speed of the i-th base station. The FDOA in Eq. (4) is an approximate result, since the spoofer maybe introduce the frequency bias f_b [2]. Thus, the actual frequency of the i-th base station is

$$f_i = \left((\boldsymbol{v}_i - \boldsymbol{v})^T \boldsymbol{a}_i c^{-1} + 1\right)(f_0 + f_b) \quad (5)$$

Because the spoofer is trying to mislead the receiver to track the spoofer signal, the f_b should not exceed the normal Doppler frequency range ± 5000 Hz [7]. Moreover, the speed of signal is far greater than the velocity of spoofer or base stations. Then we get

$$f_b \ll f_0, (\boldsymbol{v}_i - \boldsymbol{v})^T \boldsymbol{a}_i c^{-1} \ll 1 \quad (6)$$

Thus, Eq. (5) can be approximated as

$$f_i \approx (\boldsymbol{v}_i - \boldsymbol{v})^T \boldsymbol{a}_i f_0 c^{-1} + f_0 + f_b \quad (7)$$

If we calculate the FDOA according to Eq. (7), then we can get the expression of FDOA in Eq. (4).

3 Deduction and Analysis of the CRLB

From the measuring model, we can see that the measuring vector $\boldsymbol{y}$ has a Gaussian distribution $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$. We suppose the estimated parameter is $\boldsymbol{\theta} = (\boldsymbol{p}^T, \boldsymbol{v}^T)^T$. According to the reference [8], the fisher information matrix (FIM) of the estimated parameters is $\boldsymbol{I}(\boldsymbol{\theta}) = \boldsymbol{J}^T \boldsymbol{C}^{-1} \boldsymbol{J}$ where $\boldsymbol{J}$ is the jacobian matrix of $\boldsymbol{\mu}$ about $\boldsymbol{\theta}$ expressed as $\boldsymbol{J} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^T}$. The block form of $\boldsymbol{J}$ goes as follow

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial \boldsymbol{\mu}_{TDOA}}{\partial \boldsymbol{p}^T} & \boldsymbol{0} \\ \frac{\partial \boldsymbol{\mu}_{FDOA}}{\partial \boldsymbol{p}^T} & \frac{\partial \boldsymbol{\mu}_{FDOA}}{\partial \boldsymbol{v}^T} \end{pmatrix} \triangleq \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{B} & \boldsymbol{D} \end{pmatrix} \quad (8)$$

If the i-line of $\boldsymbol{A}$ is expressed as $\boldsymbol{A}[i, :]$ $(i = 1, \ldots, M)$, according to the matrix derivation theory [9], then we have

$$\begin{aligned} \boldsymbol{A}[i,:] &= \frac{\partial \mu_{TDOA}^{(i)}}{\partial \boldsymbol{p}^T} = (\boldsymbol{a}_i - \boldsymbol{a}_N)^T,\ \boldsymbol{D}[i,:] = \frac{\partial \mu_{FDOA}^{(i)}}{\partial \boldsymbol{v}^T} = \frac{f_0}{c}(-\boldsymbol{a}_i + \boldsymbol{a}_N)^T \\ \boldsymbol{B}[i,:] &= \frac{\partial \mu_{FDOA}^{(i)}}{\partial \boldsymbol{p}^T} = \frac{f_0}{c}\left(\boldsymbol{F}_i(\boldsymbol{v}_i - \boldsymbol{v}) - \boldsymbol{F}_N(\boldsymbol{v}_N - \boldsymbol{v})\right)^T \end{aligned} \tag{9}$$

in which $\boldsymbol{F}_i$ is expressed as

$$\boldsymbol{F}_i = \|\boldsymbol{p} - \boldsymbol{p}_i\|^{-1}\boldsymbol{I} - \|\boldsymbol{p} - \boldsymbol{p}_i\|^{-3}(\boldsymbol{p} - \boldsymbol{p}_i)(\boldsymbol{p} - \boldsymbol{p}_i)^T \tag{10}$$

where $\boldsymbol{I}$ is an identity matrix. By the equation $\boldsymbol{I}(\boldsymbol{\theta}) = \boldsymbol{J}^T\boldsymbol{C}^{-1}\boldsymbol{J}$ and Eq. (8), we can deduce the inverse of $\boldsymbol{I}(\boldsymbol{\theta})$. If we denote the upper-left block matrix of $\boldsymbol{I}(\boldsymbol{\theta})^{-1}$ as $\boldsymbol{P}_{11}$ and the lower-right block matrix as $\boldsymbol{P}_{22}$, according to the matrix inverse theory [9], then we have

$$\boldsymbol{P}_{11} = \sigma_{TOA}^2\left(\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A} + \sigma_{TOA}^2\sigma_{FOA}^{-2}\boldsymbol{B}^T\left(\boldsymbol{G} - \boldsymbol{G}\boldsymbol{D}\left(\boldsymbol{D}^T\boldsymbol{G}\boldsymbol{D}\right)^{-1}\boldsymbol{D}^T\boldsymbol{G}\right)\boldsymbol{B}\right)^{-1} \tag{11}$$

$$\boldsymbol{P}_{22} = \sigma_{FOA}^2\left(\boldsymbol{D}^T\left(\boldsymbol{G} - \boldsymbol{G}\boldsymbol{B}\left(\sigma_{FOA}^2\sigma_{TOA}^{-2}\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A} + \boldsymbol{B}^T\boldsymbol{G}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^T\boldsymbol{G}\right)\boldsymbol{D}\right)^{-1} \tag{12}$$

where $\boldsymbol{G} = (\boldsymbol{I}_M + \boldsymbol{E}_M)^{-1}$.

The diagonal element of the inverse of $\boldsymbol{I}(\boldsymbol{\theta})$ is the lower bound of the variance for parameter estimation. By comparing the expression of $\boldsymbol{A}$ and $\boldsymbol{D}$, we get $\boldsymbol{D} = -f_0c^{-1}\boldsymbol{A}$. In Eq. (9), the elements of $\boldsymbol{B}$ are inversely proportional to the distance $\|\boldsymbol{p} - \boldsymbol{p}_i\|$ from the spoofer to the i-th base station. The $\boldsymbol{B}$ can be ignored when the spoofer is far away from the base stations, such as 50 m away. Thus, by ignoring $\boldsymbol{B}$ in Eqs. (11) and (12), we get the following approximation

$$\boldsymbol{P}_{11} \approx \sigma_{TOA}^2\left(\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A}\right)^{-1}, \boldsymbol{P}_{22} \approx \sigma_{FOA}^2\left(\boldsymbol{D}^T\boldsymbol{G}\boldsymbol{D}\right)^{-1} = \sigma_{FOA}^2c^2f_0^{-2}\left(\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A}\right)^{-1} \tag{13}$$

The approximation in Eq. (13) applies to most spoofing scenarios. The row vector of $\boldsymbol{A}$ is the difference between the two unit vectors, which are only related to the position of spoofer and base stations. So the matrix $\boldsymbol{A}$ can be called the difference of geometric matrix. From Eq. (13) we can see that the positions of spoofer and base stations have the same effect on the estimations of position and velocity. The position error is affected by the TOA measuring error and the velocity error is affected by the FOA measuring error. Besides, the estimated errors are not affected by the velocity of spoofer and base stations. The reference [7] lists the normal ranges of measuring errors. Since there are not errors of the ionosphere and troposphere, we can suppose the measuring error of TOA and FOA are 6 m and 2 Hz. According to Eq. (13), the position error is 15.8 times the velocity error.

According to the CRLB of position and velocity, the TOA measuring error will not affect the position error and the FOA measuring error will not affect the velocity error. Hence, when the position error increases because of the increasing TOA

measuring error, we can reduce position error with the assist of FOA. When the velocity error increases because of the increasing FOA measuring error, we can reduce velocity error with the assist of TOA. So the performance of the spoofer localization system can be improved by coupling the position and velocity together based on the information fusion of TDOA and FDOA. The way to couple the position and velocity is their physical relation. Even though the WLS without information fusion locates the spoofer by the TDOA and FDOA, its performance approaches to the method using only the TDOA. This is because the WLS does not utilize the physical relation, which results in the failure of coupling between the position and velocity. In contrast, the location algorithm based on the filter process contains the physical relation, which can couple the position and velocity together. Thus, the algorithm based on the filter process using measuring information fusion, such as PF, should have a better performance than WLS when we can obtain the TDOA and FDOA measurements at the same time.

4 Theory of Particle Filter

According to the Sect. 3, the filter algorithm integrates the physical relation to the state transition equations to couple the position and velocity together, which is called measuring information fusion. Since the measuring equations is nonlinear, we adopt the PF to locate and track the spoofer in order to improve the precision of the system model [10].

4.1 Model of Particle Filter

The model of PF consists of the state model and the measuring model. The state transition equations in the state model contain the physical relation of the spoofer. In the state model, the state vector is $\boldsymbol{x} = (\boldsymbol{p}^T, \boldsymbol{v}^T)^T \in \mathbb{R}^{2d\times 1}$ and d is the space dimension. We need to introduce the process noise to improve the convergence of PF. We suppose the position noise is $\boldsymbol{w}_p \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_p^2 \boldsymbol{I}_d\right)$, the velocity noise is $\boldsymbol{w}_v \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}_d\right)$ and the acceleration noise is $\boldsymbol{w}_a \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_a^2 \boldsymbol{I}_d\right)$. The $\boldsymbol{I}_d$ is a d-order identity matrix. If the measuring interval is T_s, then we have the discrete state transition equations as follow [10]

$$\boldsymbol{x}(n+1) = \begin{pmatrix} \boldsymbol{I}_d & T_s \boldsymbol{I}_d \\ \boldsymbol{0} & \boldsymbol{I}_d \end{pmatrix} \begin{pmatrix} \boldsymbol{p}(n) \\ \boldsymbol{v}(n) \end{pmatrix} + \begin{pmatrix} \frac{1}{2}\boldsymbol{a}(n) T_s^2 + \boldsymbol{w}_p(n) \\ \boldsymbol{a}(n) T_s + \boldsymbol{w}_v(n) \end{pmatrix} \tag{14}$$

in which n is the measuring time. The Eq. (14) indicates that the state at time $n + 1$ depends on the state and the process noise at time n. The measuring model in PF is

similar to the model in Sect. 2.1. The measuring vector is $\boldsymbol{y} \in \mathbb{R}^{2M\times 1}$ and the measuring equations are

$$\boldsymbol{y}(n) = \begin{pmatrix} \boldsymbol{\mu}_{TDOA}(n, \boldsymbol{x}(n)) + \boldsymbol{w}_{TDOA}(n) \\ \boldsymbol{\mu}_{FDOA}(n, \boldsymbol{x}(n)) + \boldsymbol{w}_{FDOA}(n) \end{pmatrix} \tag{15}$$

The measuring equations express the relation between the measurements and states at time n.

4.2 Algorithm of Particle Filter

The PF is the approximation of the Bayesian filter. If the measurements from time 1 to n is $\boldsymbol{y}(1:n)$, then the posterior probability in the Bayesian is $p(\boldsymbol{x}(n)|\boldsymbol{y}(1:n))$, whose solution is hard to obtain. So we can get its approximate solution by the PF. The Alg. (1) is the implementation of the PF, which consists of four parts [11]. The first part is about generating N_s new particles according to the state Eq. (14). The second part is about updating the particle weights according to the measuring Eq. (15). The equation of updating weights for the k-th particle is

$$\omega^{(k)}(n+1) = \omega^{(k)}(n) p_G(\boldsymbol{y}(n)|\boldsymbol{\mu}, \boldsymbol{C}) \tag{16}$$

where $\omega^{(k)}(n)$ represents the k-th particle's weight at time n and $p_G(\boldsymbol{y}(n)|\boldsymbol{\mu}, \boldsymbol{C})$ is the probability of Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{C})$ at $\boldsymbol{y}(n)$. The third part is about the resampling of particles in order to prevent the particles from gathering in the place with small weight [11]. Without the resampling, the particle degeneracy will appear and the weights of particles will not exactly approximate to the posterior probability. The Alg. (2) is the implementation of resampling [11]. The fourth part is about the estimation of the state parameters. Having obtained the approximation of the posterior probability, we can estimate the parameters by the Bayesian minimum mean square error. According to the reference [8], the estimation of the state is the mean of posterior probability. That's

$$\hat{\boldsymbol{x}}(n) = \sum_{k=1}^{N_s} \omega^{(k)}(n) \boldsymbol{x}^{(k)}(n) \tag{17}$$

Algorithm (1): Particle Filter	**Algorithm (2)**: Resampling of Particles
Input: $p(0)$, $v(0)$, particles' number N_s, threshold of resampling N_T, length of measurements N_m **Output**: Estimation of parameters $\hat{\boldsymbol{x}}(n)$ 1. Initialize $\boldsymbol{x}^{(k)}(0),\ \omega^{(k)}(0) = N_s^{-1}, k = 1, \ldots, N_s$	**Input**: State $\boldsymbol{x}^{(k)}(n)$, weight $\omega^{(k)}(n), k = 1, \ldots, N_s$ **Output**: Results of resampling $\boldsymbol{x}^{(k)}(n)$, $\omega^{(k)}(n)$

(continued)

(continued)

2. **for** $n = 1 \rightarrow N_m$ **do** 3. Generate new particles $\boldsymbol{x}^{(k)}(n)$ by Eq. (14) 4. Update weights $\omega^{(k)}(n)$ by Eq. (16) 5. Normalize: $\omega^{(k)}(n) \leftarrow \omega^{(k)}(n)\left(\sum_{k=1}^{N_s} \omega^{(k)}(n)\right)^{-1}$ 6. Calculate the number of equivalent particles $N_{eff} = \left(\sum_{k=1}^{N_s} \left(\omega^{(k)}(n)\right)^2\right)^{-1}$ 7. **if** $N_{eff} < N_T$ **then** 8. Resample according to Alg. (2) 9. **end if** 10. Get the output of PF $\hat{\boldsymbol{x}}(n)$ by Eq. (17) 11. **end for**	1. Calculate the cumulative distribution of weights $c^{(k)} = \sum_{i=1}^{k} \omega^{(i)}(n),\ k = 1, \ldots, N_s$ 2. Randomly select u_1 from $\left[0, N_s^{-1}\right]$ 3. **for** $j = 1 \rightarrow N_s$ **do** 4. $u_j = u_1 + (j-1)N_s^{-1}$, initialize $i = 1$ 5. **while** $u_j > c_i$ **do** 6. $i \leftarrow i + 1$ 7. **end while** 8. Generate the particles of resampling $\boldsymbol{x}_s^{(j)}(n) = \boldsymbol{x}^{(i)}(n),\ \omega^{(j)}(n) = N_s^{-1}$ 9. **end for** 10. Denote the resampling particles as $\boldsymbol{x}^{(k)}(n),\ \omega^{(k)}(n)$

5 Simulation Results and Analysis

The theory analysis in Sect. 3 and positioning algorithm in Sect. 4 are suitable for both 2-D space and 3-D space. For illustration purposes, we attempt to locate and track the spoofer in the 2-D space. For comparison, we implement both the PF and WLS for the spoofer localization. We implement the WLS by expending the pseudo-range measuring equations in [7] to the TDOA and FDOA measuring equations. We suppose there are N base stations distributed evenly in the 2-D space with interval 20 km. We denote the TOA measuring error as σ_{TOA}, the FOA measuring error as σ_{FOA}, the initial position of spoofer as $\boldsymbol{p}(0)$ and the spoofer speed as $\boldsymbol{v}$. As for the process noise, we set the position noise variance σ_p to 1 m, the velocity noise variance σ_v to 0.1 m/s and the acceleration noise variance σ_a to 0.05 m/s^2.

5.1 Distribution of the CRLB of Position and Velocity

According to the theory result of Sect. 3, we can obtain the distribution of CRLB in different locations. While simulating, we set $N = 4$, $\sigma_{TOA} = 6$ m, and $\sigma_{FOA} = 2$ Hz.

Figure 3 shows the distribution of CRLB in the 2-D space. The base station is represented as pentagram. Under the current simulation conditions, the error of position in the rectangle composed of the four base stations is less than 9 m and the error of velocity is less than 0.6 m/s. In the corner of the figure, the position error is about 100 m and the velocity error is about 6.2 m/s. The relation between the position error and velocity error coincides with the proportional relation of 15.8 times in Sect. 3.

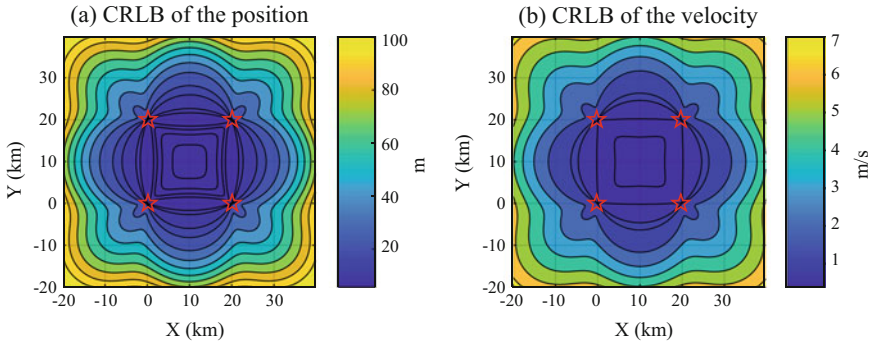


Fig. 3 CRLB of the position and velocity

5.2 *Performance of the Algorithm*

We implement the PF and WLS for comparison and analyse the errors of the algorithm under different conditions. According to the results of Sect. 3, we can obtain the effect factors of errors, including the measuring errors, $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{D}$. The matrix $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{D}$ are related to the position of the spoofer and base stations. Besides, $\boldsymbol{B}$ is still related to the velocity of the spoofer and base stations. Moreover, the number of the base stations affects the dimension of $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{D}$. Therefore, we should consider the effect factors of the algorithm's performance, including the relative position of spoofer and base stations, $\boldsymbol{v}$, σ_{TOA}, σ_{FOA} and N. There are 6 cases with different simulation parameters in the Table 1.

The performance of tracking by PF and WLS is simulated in the case 1. As shown in the Fig. 4, the errors of WLS jitter around the CRLB. The PF needs time to converge. When the PF is stable, the errors of PF are far less than WLS and below the CRLB. The n-th estimation of PF with information fusion is based on the first n measurements, which is practicable by measuring the TOA and FOA at the same time, but the WLS and CRLB are only based on the n-th measurement. Thus, the errors of PF are able to be less than WLS and CRLB. From the figure, we can also see that the errors of WLS jitter violently because of the independence of estimated process in time series. On the contrary, the errors of PF jitter slowly because of the filter process in which the states before and after depend on each other.

Table 1 Simulation parameters

Case	$\boldsymbol{p}$ (km)	$\boldsymbol{v}$ (m/s)	σ_{TOA}(m)	σ_{FOA}(Hz)	N
1	(10, 10)	(5, −5)	10	5	4
2	(10:2:30, 10)	(5, −5)	6	2	4
3	(10, 10)	(1:3:28, 0)	6	2	4
4	(10, 10)	(5, −5)	1:2:19	2	4
5	(10, 10)	(5, −5)	6	1:1:10	4
6	(21, 21)	(5, −5)	6	2	3:1:9

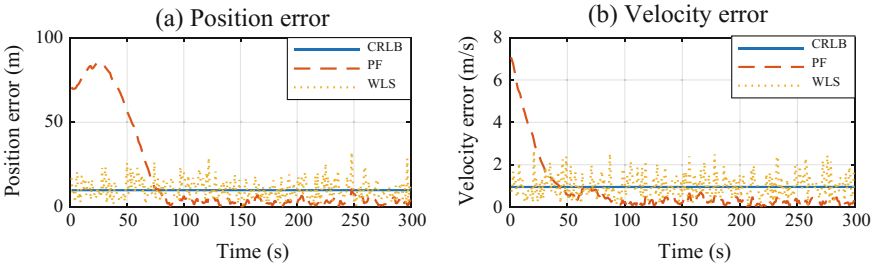


Fig. 4 Performance of tracking: PF versus WLS

The simulated results in the case 2 are shown in the Fig. 5a. It shows the errors of position and velocity of PF and WLS at different distance between the spoofer and base stations. The parameter 10:2:30 in the table represents the number selected from 10 to 30 with interval 2. Along with the increasing of distance, CRLB and the errors of WLS increase quickly. However, the errors of PF are always below WLS and increase slowly. Therefore, the PF is better to adapt to the poor geometric layout when the spoofer is far from the base stations. Besides, the CRLB of position and velocity have the similar variation trends, which coincides with Eq. (13). The Fig. 5b shows the errors of PF and WLS with different spoofer velocity in the case 3. We can see that the errors of position and velocity are unrelated to the spoofer velocity, which is shown in Eq. (13).

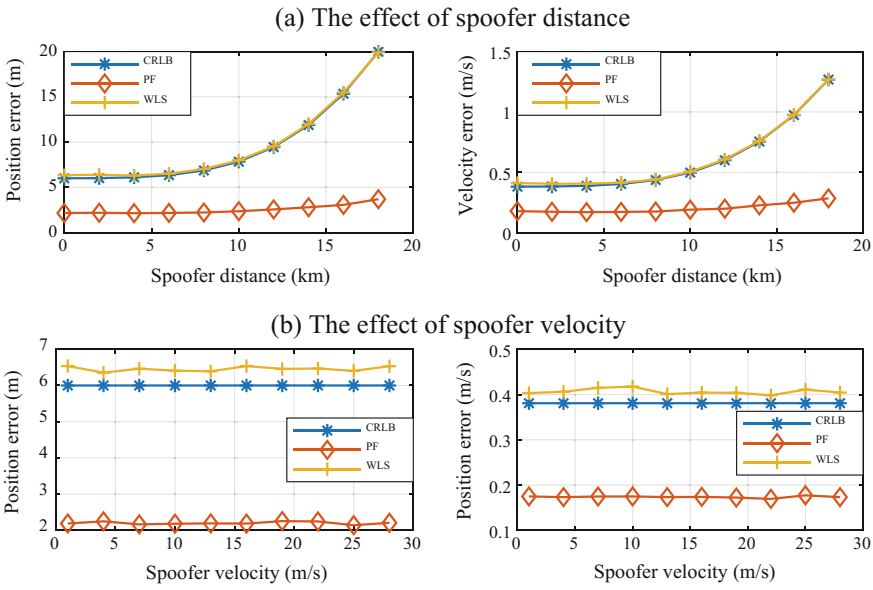


Fig. 5 The effect of distance and velocity of spoofer

The simulated results in the case 4 and 5 are shown in the Fig. 6. The Fig. 6a shows the effect of TOA error on the position and velocity and the Fig. 6b shows the effect of FOA error. We can see that TOA error is less likely to affect the velocity error and FOA error is less likely to affect the position error, which coincides with the Eq. (13). With regard to the WLS, the position error of WLS gradually deviates from CRLB when TOA error increases and the position error is close to CRLB when TOA error is small. So is the effect of FOA error on the velocity error of WLS. As can be seen, under the high SNR condition, the errors of WLS can approach to the CRLB. With regard to the PF, when the TOA error is large, we can obtain precise velocity estimation if the FOA error is small. Similarly, when the FOA error is large, we can obtain precise position estimation if the TOA error is small. Therefore, the PF improves the performance of the spoofer localization system by coupling the position and velocity together, which enables the errors increase slowly even though TOA or FOA measuring error increases sharply. The Fig. 6 verifies that we can observably improve the system location performance by measuring the time and frequency at the same time.

Figure 7 shows the simulated results in the case 6. From the figure, we can see that the position and velocity errors decrease with the number of base stations increasing. The number has a big effect on the WLS when it's less than 5. However, the number has a small effect on the PF. Thus, comparing to the WLS, the PF is more applicable to locate and track the spoofer when the number is small, which can reduce the resource of the base stations.

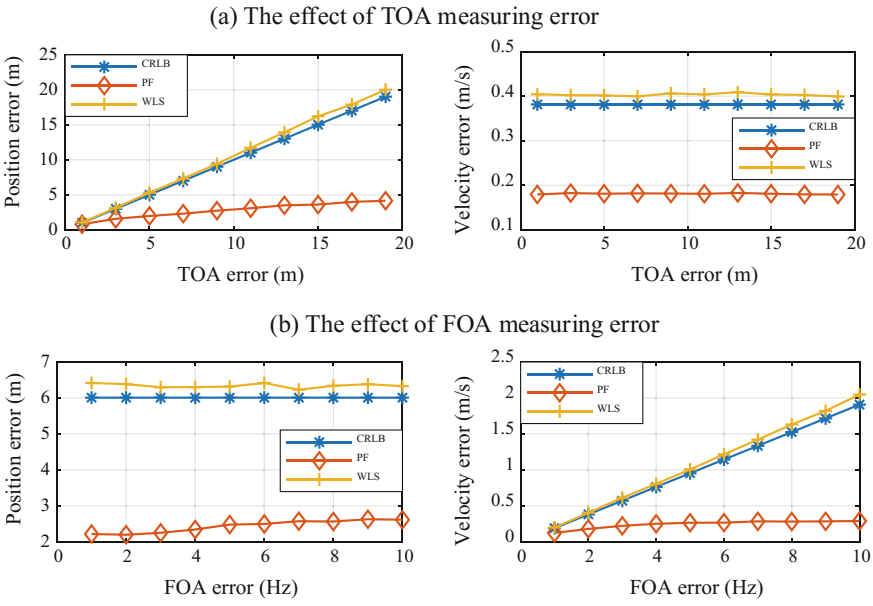


Fig. 6 The effect of measuring errors

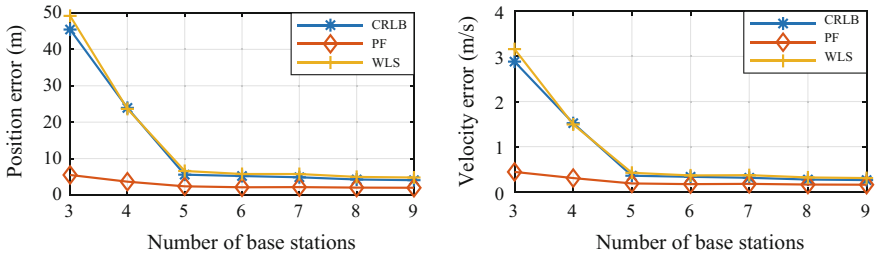


Fig. 7 The effect of the base stations' number

6 Conclusions

We propose an algorithm for the spoofer localization based on the measuring of TDOA and FDOA simultaneously. And we adopt the PF to improve the precision of the system model and implement the measuring information fusion. On the basis of establishing the spoofer localization system, we obtain the expressions of position and velocity errors according to the CRLB theory. Based on this, we prove the errors of position and velocity are unrelated to the velocity and only related to the measuring errors and the difference of geometric matrix when the spoofer is far from the base stations. We also find the measuring information fusion can improve the system performance. According to the error expressions, we obtain the effect factors of error on the position and velocity. Through the simulation under different conditions, we find that the performance of PF is prior to the WLS. The PF can track the spoofer stably and relies less on the geometric layout. Also the errors of PF increase slowly with the measuring errors increasing. The estimated performance of the system can be significantly improved by coupling the position and velocity through the filter process, which needs to measure the TDOA and FDOA at the same time. It is because the filter process estimates the state at time n based on the first n measurements. Besides, the PF can gain a better performance than the WLS even though the number of base stations is small. In the future, we will implement the proposed algorithm for the spoofer locating and tracking on hardware. We also need to pay attention to the effects of clock synchronization error and the position errors of base stations on the performance of the proposed algorithm.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 61571255).

References

1. Psiaki ML, Humphreys TE (2016) GNSS spoofing and detection. Proc IEEE 104(6):1258–1270

2. Amar A, Weiss AJ (2008) Localization of narrowband radio emitters based on doppler frequency shifts. IEEE Press 56(11):5500–5508
3. Bhatti JA, Humphreys TE, Ledvina BM (2012) Development and demonstration of a TDOA-based GNSS interference signal localization system. IEEE PLANS 2012:455–469
4. Dempster AG, Cetin E (2016) Interference localization for satellite navigation systems. Proc IEEE 104(6):1318–1326
5. Liu ZM, Guo FC (2015) Azimuth and elevation estimation with rotating long-baseline interferometers. IEEE Trans Signal Process 63(9):2405–2419
6. Jafarnia-Jahromi A, Broumandan A, Nielsen J, Lachapelle G (2012) GPS vulnerability to spoofing threats and a review of antispoofing techniques. Int J Navig Obs (9)
7. Xie Gang (2009) Principles of GPS and receiver design. Publishing House of Electronics Industry
8. Kay SM (1993) Fundamentals of statistical processing, vol I. Estimation Theory. PTR Prentice hall
9. Petersen KB, Pedersen MS (2008) The matrix cookbook. Version: Nov 14, 2008
10. Gustafsson F, Gunnarsson F, Bergman N et al (2002) Particle filters for positioning, navigation, and tracking. IEEE Trans Signal Process 50(2):425–437
11. Van Trees H, Bell K (2007) A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. Wiley-IEEE Press

Implementation of a Dual Estimate Tracking Based Multipath Mitigation Method on a Software Receiver

Chuhan Wang, Yang Gao, Xiaowei Cui and Mingquan Lu

Abstract Binary Offset Carrier modulated signal, which will widely be used in next generation Global Navigation Satellite Systems, has multi peaks in the auto-correlation function. In practical application condition, multipath and ambiguity are two main problems that can serious degrade the GNSS receiver ranging precision due to the feature of multi peaks. To solve the ambiguity problem, we typically adopt Dual Estimate Tracking (DET) that treats the subcarrier independently and tracks it by an additional loop. However, the DET multipath mitigation structure is not specially considered yet due to the complex features of two-dimension cross-correlation function. In this paper, we implement a DET based anti-multipath method for BOC modulated signals, which mitigates the direct influence by an earlier phase Pseudo-Noise code instead of the prompt for subcarrier locked loop and the indirect influence by the Double Delta method for delay locked loop. A software receiver further proves the achievability of the multipath mitigation method and the results show that this method can significant reduce the short and mediate range delay multipath error at the expense of increasing the long range delay multipath error a little. This method can also achieve the compromise between multipath and noise performance by adjusting the local Pseudo-Noise code phase for different application scenarios.

Keywords Mitigation · Binary offset carrier · Software receiver
Dual estimate tracking

C. Wang · X. Cui (✉) · M. Lu
Electronic Engineering Department, Tsinghua University, Beijing, China
e-mail: cxw2005@mail.tsinghua.edu.cn

Y. Gao
Beijing Satellite Navigation Center, Beijing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_29

1 Introduction

The next-generation Global Navigation Satellite Systems (GNSS) will generally use the Binary Offset Carrier (BOC) signal which provides better multipath mitigation performance and higher ranging accuracy than the conventional BPSK signal. In all kinds of BOC signals, the sin-BOC signal have good performance in the real environment and is therefore widely used.

Traditionally, the multi-peak ambiguity problem and multipath problem of BOC signals seriously hinder the receiver's performance. Therefore, there is a variety of methods proposed, including BPSK-like method [1], Bump-jump method [2] and PCF-based method [3] to solve the ambiguity problem, and Gating method [4] to improve the multipath performance. However, with the increasing order of the BOC signal, the autocorrelation function will include more correlation peaks, which makes the multi-peak problem and the multipath problem more severe, and the above method is no longer applicable.

In order to solve the problem of multi-peak ambiguity, Hodgart proposed a 2-D tracking method, that is, a Dual Estimation Tracking method (DET) [5, 6]. The DET method regards code and subcarrier as two independent parts. The delay locked loop is used to track the code part of the BOC signal to obtain a pseudo-range estimation value without ambiguity but with low accuracy. The subcarrier loop tracks the subcarrier part of the BOC signal to obtain a more accurate, but ambiguous estimation value. And then a more accurate and unambiguous estimate is obtained from the non-linear combination of those two estimation values. This technique has low complexity and is easy to implement on hardware and software platforms [7]. The experimental results show that it has better performance of unambiguous tracking for BOC (15, 2.5) signal, without losing accuracy [8], which can also be extended for tracking alternate binary offset carrier (AltBOC) signal [9]. However, at present, the design of anti-multipath method under the DET method is seldom studied, and the main difficulty of the research is the complexity of 2-D cross-correlation function (CCF).

In order to solve the multipath problem, based on the feature of the 2-D CCF of sin-BOC signal, Ref. [10] designed the anti-multipath algorithm to solve the anti-multipath problem of BOC signal, especially high-order BOC signals. Firstly, based on the ideal 2-D CCF at infinite bandwidth, two multipath effects for subcarrier tracking are analysed, namely direct and indirect effects. Then, the basic structure of the anti-multipath method is proposed and the corresponding parameter constraints are clarified to reduce the two effects of multipath.

The above studies mostly focus on the theory, and less involved in the implementation of the anti-multipath algorithm on the receiver. Based on the theory of DET anti-multipath algorithm, this paper designs the algorithm framework and gives the details. Then the ranging performance for the BOC (15, 2.5) signal is analysed by using self-made signal source and the software receiver platform. The noise and multipath noises in different parameters are analysed.

The content of this paper is arranged as below: we first gives the 2-D CCF of sin-BOC signal, and analyses the characteristics of the multipath effect on the DET structure in Sect. 2. In Sect. 3, we design the frame of multipath algorithm and the constraints of algorithm parameters. Section 4 evaluates the performance of the proposed algorithm on software receiver platform to prove the achievability and effectivity of the proposed method. Section 5 summarizes this paper.

2 DET and Its Anti-multipath Method

2.1 BOC Signal Ideal 2-D Cross-Correlation Function

Double estimation method was first proposed by Hodgart. The core view is to use the delayed independent $c(t-\tau_c)$, $s(t-\tau_s)$ instead of $p(t-\hat{\tau})$ for the CCF of received signal. The correlation function can be expressed as:

$$R(\tau-\tau_c,\tau-\tau_s)=\frac{1}{T}\int_0^T p(t-\tau)c(t-\tau_c)s(t-\tau_s)dt \tag{1}$$

where τ_c,τ_s represent the delay of the local code and subcarrier, respectively. For example, the function of $BOC_s(2,1)$ shown in Fig. 1.

It can be seen that the 2-D CCF presents different characteristics in two dimensions. In the subcarrier dimension, there are multiple peaks and extend indefinitely for the period $2T_s$; in the code dimension, there is only one peak and the peak width is $2T_c$. Therefore, the receiver can track any one of the peaks at which

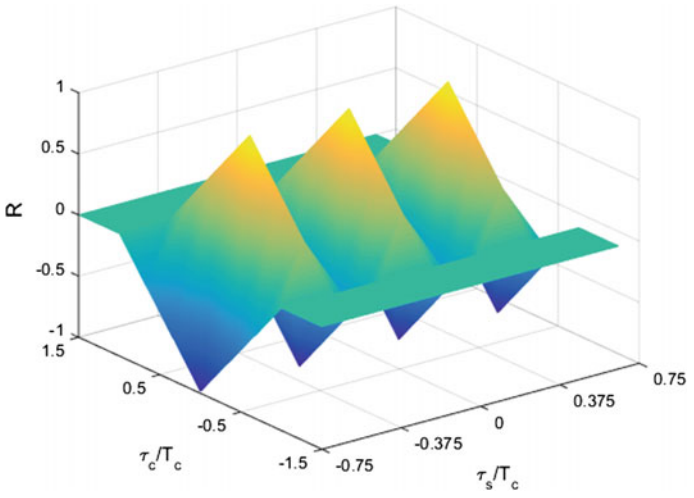


Fig. 1 2-D CCF of BOC(2, 1) modulated signal

the two dimensional delays satisfy $\tau_c = \tau$, $\tau_s = \tau + nT_s$, where the ambiguity of the subcarrier dimension (n) is an integer.

2.2 The Anti-multipath Method of DET

The DET anti-multipath algorithm is proposed in [10], and the core idea is to reduce the direct and indirect effect of multipath on the subcarrier tracking, so as to enhance anti-multipath performance.

First, we consider the direct effect of multipath. In existing methods, the acquisition of correlation values $A1$, $A2$ exploits the code phase of an punctual path; therefore, there is a direct effect when the multipath delay is within range $[0, T_c]$. The direct effect of reducing multipath can be equivalent to reducing the multipath delay range described above. However, due to hysteresis of multipath delay, the local code of early path can be used to obtain the correlation value, and these correlation values can be used to discriminate the subcarriers so as to reduce the multipath delay range that is directly influenced.

Second, we consider the indirect effects of multipath. To reducing the indirect effect is equivalent to reducing the effect of multipath on the code dimension. As can be seen from Fig. 1, in the DET structure, the code dimension can be first approximated as tracking the correlation peak of a BPSK signal. Therefore, in theory, anti-multipath algorithms designed for BPSK signals can be directly used. Here, the Double Delta (DD) method with better anti-multipath effect and lower complexity is selected [11].

In view of the above thought, correlation values used for phase discrimination of code and subcarrier are shown in Fig. 2.

In Fig. 2, phase discriminations of subcarrier are represented by C_1, C_2; D_1, D_2, D_3 and D_4 represents phase discriminations of code. These correlation values can be expressed as:

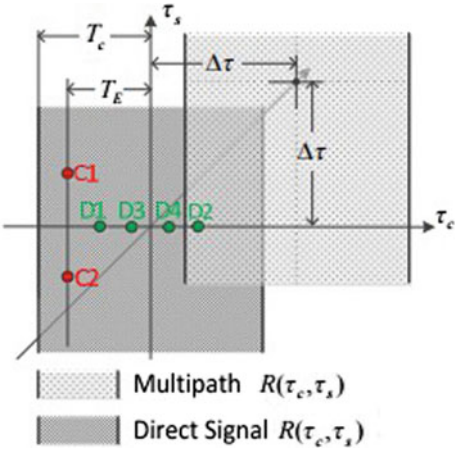


Fig. 2 Illustration of the proposed method

$$\begin{cases} C_1 = R(\tau_c - T_E, \tau_s + \dfrac{T_{DS}}{2}) \\ C_2 = R(\tau_c - T_E, \tau_s - \dfrac{T_{DS}}{2}) \end{cases} \tag{2}$$

$$\begin{cases} D_1 = R(\tau_c - T_{DC}, \tau_s) \\ D_2 = R(\tau_c + T_{DC}, \tau_s) \\ D_3 = R(\tau_c - \dfrac{T_{DC}}{2}, \tau_s) \\ D_4 = R(\tau_c + \dfrac{T_{DC}}{2}, \tau_s) \end{cases} \tag{3}$$

Here, T_E is the local early code phase for subcarrier phase discrimination. T_{DC} and T_{DS} represent E-L spacing of the code and subcarrier, respectively. The constraint parameters can refer to the literature [10].

3 Implement of the Anti-multipath Algorithm

The following figure shows the anti-multipath algorithm framework based on DET. Similar to the traditional DET architecture, the frame map can be roughly divided into two parts: the correlator on the left and the discriminator and filter on the right. In the software receiver, the correlator's algorithm is implemented on GPU, while the algorithm of discriminator and filter is implemented on CPU. The intermediate frequency signal r_{IF} is correlated with the local code, subcarrier, and then a total of 8 correlator outputs feed into the discriminator and filter unit, to correct the three NCOs, and finally form three independent tracking loops, namely, delay locked loop (DLL), subcarrier phase locked loop (SLL), carrier phase locked loop (PLL) (Fig. 3).

Different from the traditional DET architecture, we adopt the early local code of the subcarrier discriminator, so the code generator C_s takes the place of the C_P to generate the correlation value C_1, C_2 for the subcarrier discrimination. We use DD method, so the code generator increases the C_{VE} and the C_{VL}, resulting in correlation values D_1, D_2 for the code discrimination.

The filter and discrimination unit is also different from the traditional DET. The three loop equations are given below.

$$\begin{cases} \varepsilon_c = (D_3 - D_4) - 0.5 \cdot (D_1 - D_2) \\ \varepsilon_s = C_2 - C_1 \\ \varepsilon_\varphi = atan2(Q_P, I_P) \end{cases} \tag{4}$$

The carrier phase discriminator ε_φ uses a pure PLL four-quadrant arctangent discriminator, which can achieve a gain of 6 dB than the Costas loop. Code

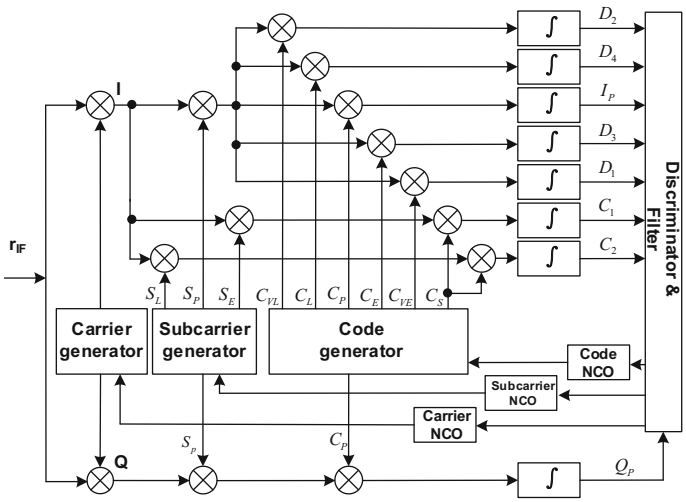


Fig. 3 The tracking architecture of the algorithm

discriminator ε_c and subcarrier discriminator ε_s avoid the square loss caused by non-coherent discriminator. The outputs of these discriminators are respectively fed into three loop filters to smooth the noise. The filtered outputs of discriminators are finally used to drive NCOs of code, subcarrier and carrier, respectively.

The accuracy of code locked loop measurement is low but unambiguous, and the accuracy of subcarrier locked loop measurement is high but ambiguities. So one natural idea is to combine the code measurement with the subcarrier measurement to form an unambiguous and high-precision measurement. Below we give the combined expression.

$$\hat{c}_s = \hat{s} + round(\frac{\hat{c} - \hat{s}}{\lambda}) \times \lambda \tag{5}$$

where λ represents the subcarrier wavelength, $\hat{s}$ represents the subcarrier measurement and $\hat{c}_s$ represents the combined measurement. When the precision of code measurement $\hat{c}$ is less than $\lambda/2$, we can achieve high-precision, unambiguous measurement.

4 Experiment and Discussion

This article uses Matlab 2014b software to generate BOC (15, 2.5) signal. In order to facilitate the algorithm debugging, this paper developed the corresponding real-time software receiver. The device configuration parameters used in this article

are: Intel (R) Core i5-4570 CPU, clocked at 4.00 GHz, memory 16.0 GB, NVIDIA GeForce GTX 750 Ti GPU.

Aiming at the BOC (15, 2.5) signal generated by the signal generator and the anti-multipath DET algorithm implemented in the software receiver, we design 9 experiments to analyze the noise performance and multipath performance of the new algorithm compared with the ordinary DET algorithm. Nine experiments are divided into two dimensions. The first dimension includes the multipath scene without noise, the noise scene without multipath, the noise and the multipath scene. The second dimension includes: the ordinary DET algorithm, 1/2 and 5/6 early code phase for the subcarrier discriminator. It should be noted that in the case of the ordinary DET algorithm, we use the narrow-correlation code discriminator instead of Double-Delta discriminator.

The parameters are configured as follows: carrier-to-noise ratio 45 dB/Hz, front-end bandwidth 40 MHz, multipath power ratio −6 dB, multipath error from 0 to 1.5 chips. The loop parameters include the PLL bandwidth 4 Hz, the SLL bandwidth 1 Hz, the DLL bandwidth 1 Hz, the code E-L spacing 1/12 chips, the subcarrier E-L spacing 1/12 chips, coherent integration time 10 ms.

Experiment one: noise-free, multipath power ratio −6 dB, respectively, T_E is 0, 1/2 and 5/6 chips. Other parameters can been seen in the previous section. Multipath error envelope of subcarrier is as follows.

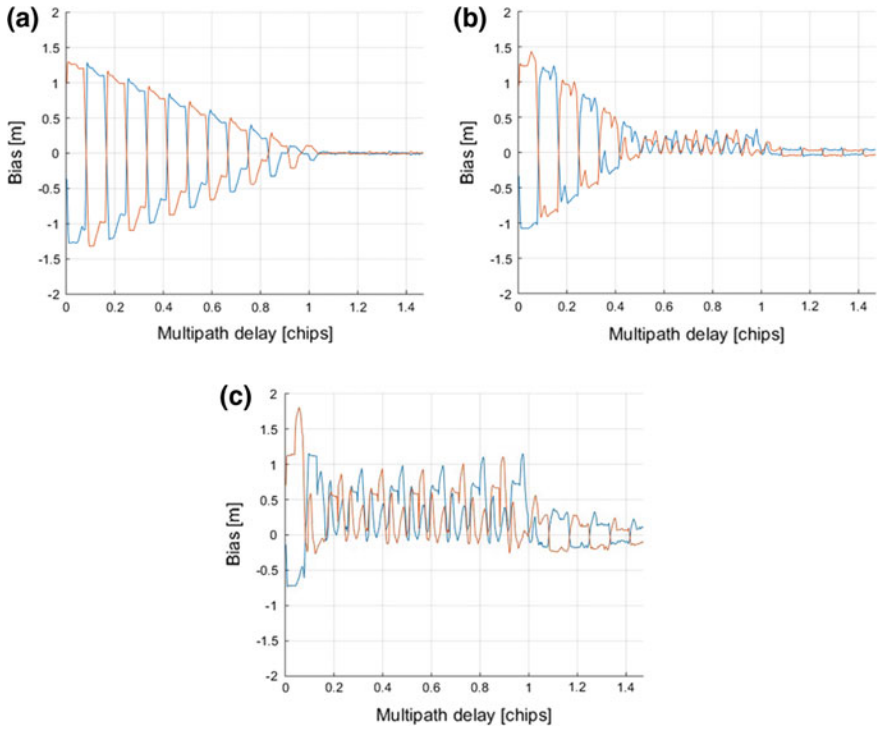


Fig. 4 Multipath error in different T_E **a** $T_E = 0$ **b** $T_E = \frac{1}{2}$ **c** $T_E = \frac{5}{6}$

As can be seen from the Fig. 4, the early local code can suppress multipath noise, which is reflected in two aspects: one is to eliminate multipath noise with a delay of more than 0.5 chips, the other is to compress multipath noise within 0.5 chip. On the other hand, we can compare the area and peak of the error envelope in different T_E. The specific results are as follows (Table 1).

As the T_E envelopes decreases, the area and the peak value of the multipath envelope decreases. However, when T_E is too small, although the peaks and areas of the envelopes do not change much, the anti-multipath performance will not be promoted. The above experiment shows that the new algorithm can effectively reduce SLL noise and effectively suppress multipath.

Experiment two: carrier to noise ratio of 45 dB/Hz, no multipath, respectively, T_E is 0, 1/2 and 5/6 chips. See the previous section for other parameters. The subcarrier measurements in different T_E are as follows (Fig. 5; Table 2).

Table 1 Results of experiment 1

	$T_E = 0$	$T_E = \frac{1}{2}$	$T_E = \frac{5}{6}$
Area	1.2182	0.6994	0.715
Peak/[m]	2.4753	2.3964	2.3304

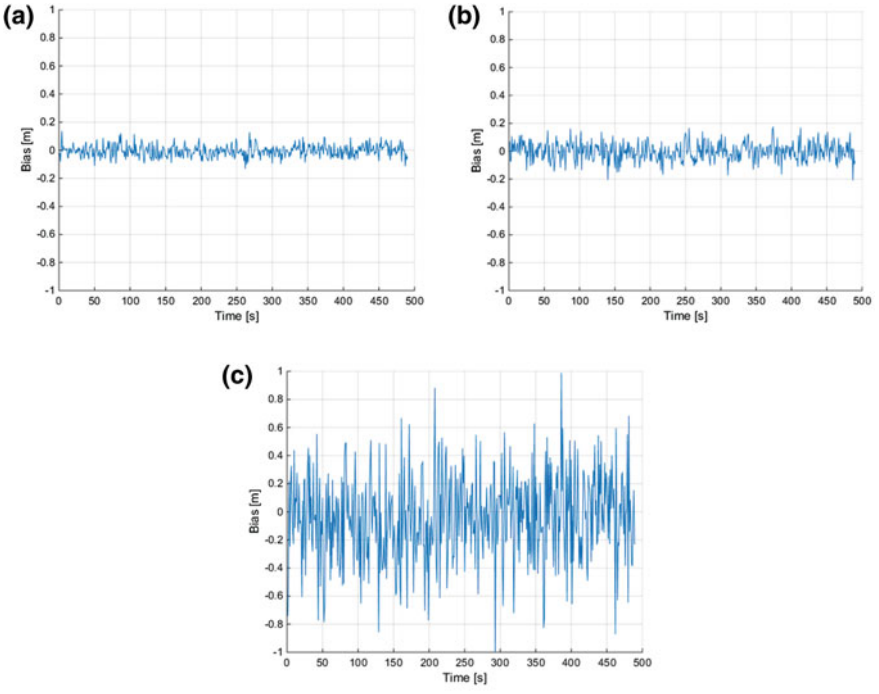


Fig. 5 Noise error in different T_E **a** $T_E = 0$ **b** $T_E = \frac{1}{2}$ **c** $T_E = \frac{5}{6}$

Table 2 Results of experiment 2

	$T_E = 0$	$T_E = \frac{1}{2}$	$T_E = \frac{5}{6}$
Standard deviation/[m]	0.0141	0.0215	0.1001

We can compare the standard deviation of the ranging errors in the three graphs. The specific results are as follows.

Three sets of data from experiment two show that the early local code will cause an increase in the error of phase discriminator, that is, the new algorithm adds the thermal noise error.

Experiment three: the carrier to noise ratio of 45 dB/Hz, multipath power ratio −6 dB, T_E 0, 1/2 and 5/6 chips, respectively. The subcarrier measurements in different T_E are as follows (Fig. 6).

Experiment three can be seen as a combination of experiment one and two, which is close to the real scene. We can also compare the error envelope area and the peak, and analyze anti-multipath performance in noisy conditions. The specific results are as follows (Table 3).

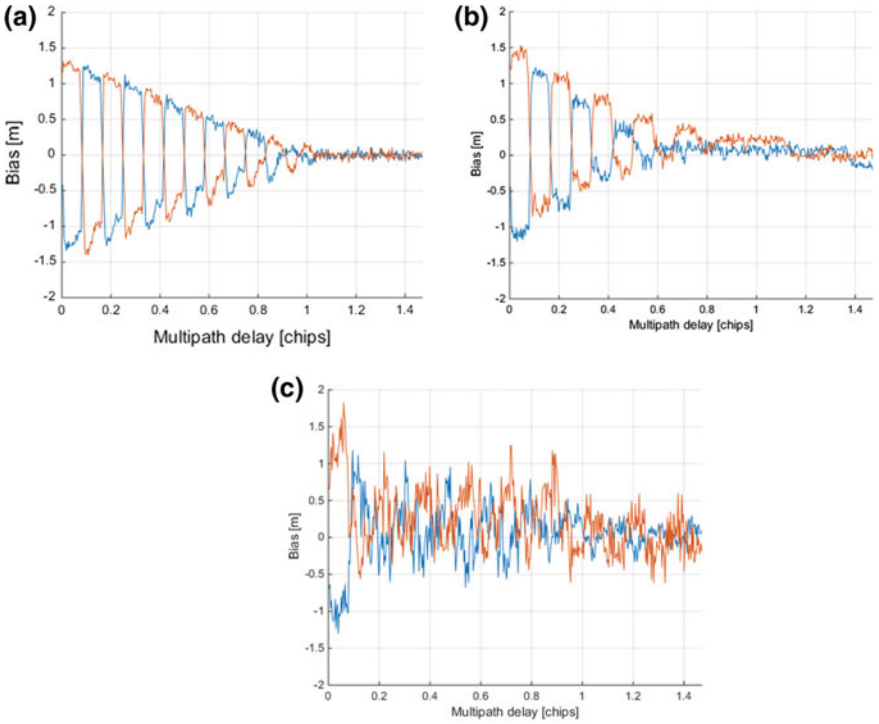


Fig. 6 Multipath and noise error in different T_E **a** $T_E = 0$ **b** $T_E = \frac{1}{2}$ **c** $T_E = \frac{5}{6}$

Table 3 Results of experiment 3

	$T_E = 0$	$T_E = \frac{1}{2}$	$T_E = \frac{5}{6}$
Area	1.3120	1.3134	1.2051
Peak/[m]	2.5341	2.5749	2.5341

The results show that: firstly the receiver can track the signal with 45 dB/Hz noise in the multipath scene; secondly, this algorithm replaces the thermal noise performance in exchange for multipath rejection. The real receiver needs to compromise between multipath and noise errors and take the appropriate parameters.

5 Summary

BOC signal processing faces two issues of ambiguity and multipath. This problem is especially noticeable in higher-order BOC signals. Based on DET anti-multipath algorithm, this paper refers to the characteristics of 2-D CCF for anti-multipath design, and realizes the algorithm implementation in receiver and the test of receiver performance.

The method above can improve the anti-multipath performance for BOC modulated signal, but with the position of the subcarrier discriminator constantly prepositioned, the receiving chance will sacrifice the thermal noise performance. The algorithm can also achieve the compromise between multipath performance and noise performance through the adjustment of the local code phase of the receiver so that the algorithm can be applied to different application scenarios.

Research on anti-multipath algorithms of high-order BOC or AltBOC signals is still in its infancy, and the next step can be to try to achieve other designs to deal with such signals to better meet different application requirements.

References

1. Fishman PM, Betz JW (2000) Predicting performance of direct acquisition for the M-code signal. 2000: Navigating New Millennium, pp 574–582
2. Fine P, Warren W (1999) Tracking algorithm for GPS offset carrier signals. In: Institute of navigation, national technical meeting ‘vision 2010: present and future’. San Diego, CA
3. Yao Z et al (2010) Pseudo-correlation-function-based unambiguous tracking technique for sine-BOC signals. In: IEEE transactions on aerospace and electronic systems, vol 46.4, pp 1782–1796
4. Nunes FD, Sousa FMG, Leitao JMN (2007) Gating functions for multipath mitigation in GNSS BOC signals. In: IEEE transactions on aerospace and electronic systems, vol 43.3
5. Hodgart MS, Blunt PD, Unwin M (2008) Double estimator—a new receiver principle for tracking BOC signals. In: Inside GNSS, vol 3.3, pp 26–36
6. Hodgart MS, Simons E (2012) Improvements and additions to the double estimation technique. In: 2012 6th ESA workshop on IEEE satellite navigation technologies and european workshop on GNSS signals and signal processing (NAVITEC)

7. Blunt P et al (2007, May) Demonstration of BOC (15, 2.5) acquisition and tracking with a prototype hardware receiver. In: ENC GNSS
8. Blunt P (2007) Advanced global navigation satellite system receiver design. Diss. University of Surrey, United Kingdom
9. Ren J et al (2012) Unambiguous tracking method for alternative binary offset carrier modulated signals based on dual estimate loop. In: IEEE communications letters, vol 16.11, pp 1737–1740
10. Gao Y et al (2014) A two-dimension tracking based multipath mitigation method for BOC signals. In: Proceedings of the international technical meeting of the institute of navigation, pp 876–883
11. McGaw GA (1999) GNSS multipath mitigation using gated and high resolution correlator concepts. In: Proceedings of national technical meeting of the satellite division of the institute of navigation, vol 3, pp 333–342

BDS-3 Spaceborne Receiver Design Applying to Space Service Volume

Chen Zhou, Shuren Guo, Yinan Meng, Dingxi Xiong and Chengeng Su

Abstract With the increase activities in space, the demand for navigation capability of space vehicles is also been stressed importantly. The application of spaceborne receiver can make space vehicles get rid of the limitation of geometric influence of ground observation, and have a high precision positioning result. However, the capture capability of the receiver is the key of the performance of the space service volume (3000–36000 km height). In this paper, we study the influence of the minimum receiving power of receiver on the performance of BDS-3 space service volume by simulation. Through the satellite visibility, visible time and position dilution of precision, the paper provides BDS-3 space service capability and demonstrates the importance of the exist of sidelobes. Finally, combining the BDS-3 and GPS navigation system, a minimum sensitivity of –175 dBW design is proposed, which can achieve a great navigation and positioning ability. The results of this paper provide important reference for the design of spaceborne receiver and antenna of space vehicle.

Keywords Satellite navigation · Space service · Volume · Receiver
BDS-3 · GPS

1 Introduction

With the potential economic benefits and scientific development, space activities and space missions are becoming more and more frequently. Through the National Space Administration statistics, in the next 20 years, about 40% of space missions need to be transported in the Medium Earth Orbit (MEO), geosynchronous orbit

C. Zhou (✉) · Y. Meng · D. Xiong · C. Su
Beijing Institute of Tracking and Telecommunication Technology,
100094 Beijing, China
e-mail: zhouchen1006@foxmail.com

S. Guo
National University of Defense Technology, 410073 Changsha, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_30

(GEO) and elliptical high orbit (HEO). As a result, the demand for navigation capabilities in the space service volume (SSV) is constantly improving. There are two kinds of navigation and positioning methods for space vehicles, which are ground observation technology and satellite borne navigation technology. The traditional method of ground observation is to build a set of fixed observations on a very long baseline. Sometimes influenced by the restriction of geography, we also have to utilize unfixed devices, such as measuring ships, measuring cars, etc. The distribution of the observations results to the quality of this technology. Taking into account the political and economic reasons, it is difficult for us to deploy overseas monitoring station. So sometimes ground observation technology is not very convenient and useful. The second is the satellite navigation system, like the global navigation satellite system, such as GPS of the United States, the Russian GLONASS, European GALILEO and Chinese Beidou navigation system. Through the extended application of those navigation satellites, the navigation space can be extended to the Space Service Volume (SSV) [1] 3000–36000 km height, which also effectively solve the problem of ground observation. The GNSS are well distributed in space, good in geometry and long visible time, which makes up for many defects of ground observation technology [2].

A lot of researches have been done on the definition of space service volume and space navigation experiments. Bauer [3] defines the service performance of the GPS space, and study the signal availability, pseudorange error and signal power of GPS; Jing [4] analyses Beidou signal availability, and based on the conclusion, simulate the lunar mission; Li [5] analyses and evaluates the GNSS service capabilities based on the main lobe of the antenna. Because the performance of the SSV is not only affected by the geometric visibility of the space, but also affected by the power of the signal. In terms of signal power, there are few studies on the design and requirements of the receiver for the Beidou SSV.

Chinese global navigation satellite system BeidouIII (simplified as BDS-3) is expected to be completed in 2020. It consists of 24 MEO, 3 GEO and 3 IGSO. BDS-3 Interface Control Document on signal B1C, B2a [6] is also announced in 2017 September. With the development of BDS-3 system and weak signal capture technology, the spaceborne receiver capability can gradually achieve the requirements from the LEO to the HEO [7–9]. Due to the different height of space vehicles and position to BDS-3 constellation, the requirements of receiver are different. In the case of LEO, the main lobe signal received by the antenna is strong, and the requirement of the receiver design is not that strict. In the case of HEO, the main lobe of the receiver is weak due to the large transmission loss. In this time, side lobe signals need to be taken into account. However, because of the weak side lobe signal, the visibility of the space vehicle has a very strict demand for the receiver (Fig. 1).

Therefore, in this paper, we discuss the performance of BDS-3 SSV through different receiver parameters design by software simulation and calculation under different environments. Then put forward effective designing suggestion, so as to achieve the purpose of optimizing resources and saving costs in engineering.

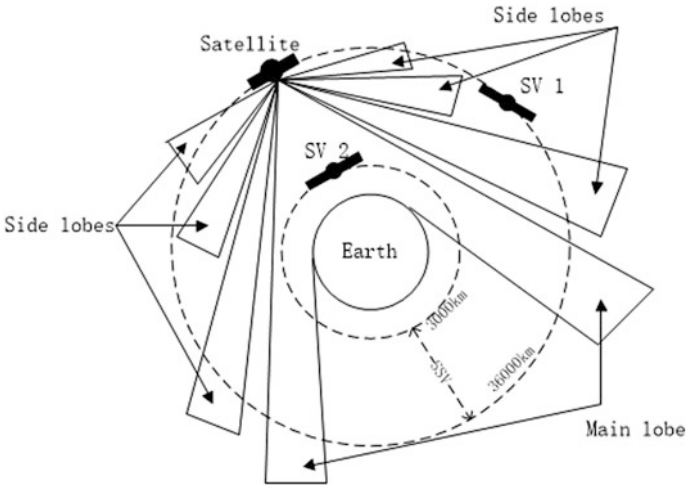


Fig. 1 Sketch map of space service volume

2 BDS-3 Constellation Modeling

2.1 Basic Parameters and Modeling

Utilize a simulation platform based on STK and MATLAB. Take STK as the core of link computing, and use MATLAB to analyse and simulate the model [10]. The simulation parameters and time settings are shown in Table 1.

Table 1 BDS-3 constellation simulation parameters

	BDS-3		
	24MEO	3GEO	3IGSO
Sources of data	Walker (24/3/1)	80E/110.5E°/140E°	Walker (3/3/2)
Inclination	55°	0°	55°
Orbit height	21528 km	35786 km	35786 km
Periodic regression	7/13 sidereal day	–	–
Number of satellites	30		
Simulation start time (UTCG)	2017-09-26 04:00:00		
Simulation stop time(UTCG)	2017-09-27 04:00:00		
Total time	24 h		

2.2 Antenna Calculation

The received power of antenna reflects the absolute intensity of the signal. The receiver's ability of signal acquisition and tracking sensitivity is usually measured by the minimum received signal power of the antenna. The transmission formula of signal transmitting in free space is Friis equation:

$$P_R = P_T + G_T + G_R + 20\log\left(\frac{c}{4\pi df}\right) - L_A \tag{1}$$

where P_R represents received power of the antenna, P_T is launching power of the satellite, G_T is gain of launching antenna of a specific angle, G_R is gain of receiving antenna of a specific angle, c represents velocity of light, d is the distance between satellite and receiver, f is frequency of the signal, L_A represents the atmospheric loss usually taken as 0.5 dB (Fig. 2).

The peak value of the antenna B1C signal can be obtained by calculating the spatial geometric relationship.

$$\frac{R_e}{\sin\alpha} = \frac{H_s + R_e}{\sin(\theta + 90^\circ)} \tag{2}$$

$$d = \frac{R_e\cos(\alpha+\theta)}{\sin\alpha} \tag{3}$$

$$P_{MEO} = P_R - G_T - G_R + L_A - 20\lg\left(\frac{c}{4\pi df}\right) = 16.5\ \text{dBW} \tag{4}$$

$$P_{GEO} = P_R - G_T - G_R + L_A - 20\lg\left(\frac{c}{4\pi df}\right) = 18.0\ \text{dBW} \tag{5}$$

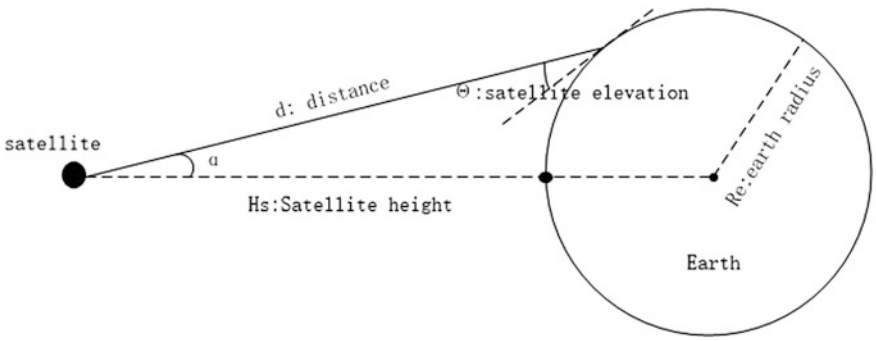


Fig. 2 Stellate elevation angle

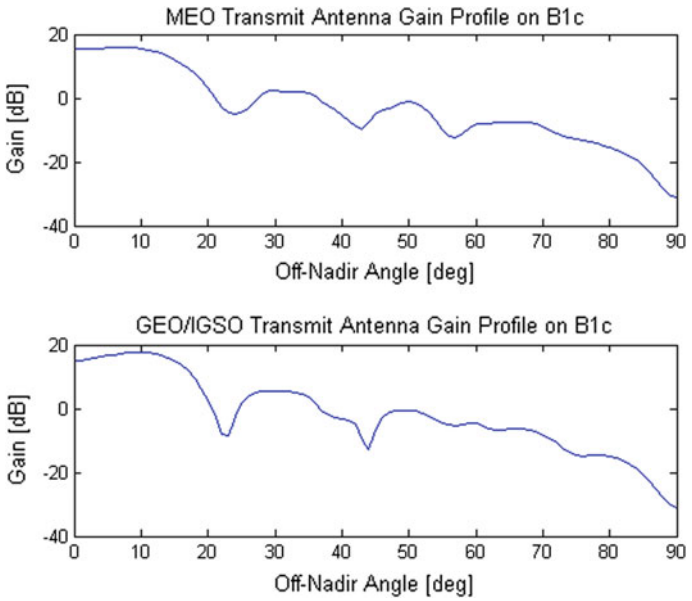


Fig. 3 Antenna pattern

The peak value of the B1C signal on the MEO and GEO heights is calculated, and the general antenna pattern of BDS-3 is obtained, as shown in Fig. 3.

Therefore, we can calculate the limit distance of antenna power transmission under the given receiver sensitivity, as shown in Fig. 4, under the sensitivity of −170 dBW.

Through the antenna data, main lobe angle and side lobes angles can be concluded in Table 2.

3 Receiver Design

3.1 Sensitivity Design

Suppose that a spacecraft flies at the altitude of 3000 km in equator, and the number of satellites visible in the simulation period is as follows (Fig. 5):

Under the sensitivity of −170 dBW, the average number of visible satellites in different height during the simulation period is calculated. It can be found that with the increase of height, satellite visibility increases slightly; when the height continues to increase; the number of visible satellites sharply reduce; at the height about 25,000 km, there will be some fluctuations in visibility because of the sidelobe signals; the visibility of the edge of SSV at 36,000 km will converge to the minimum (Fig. 6).

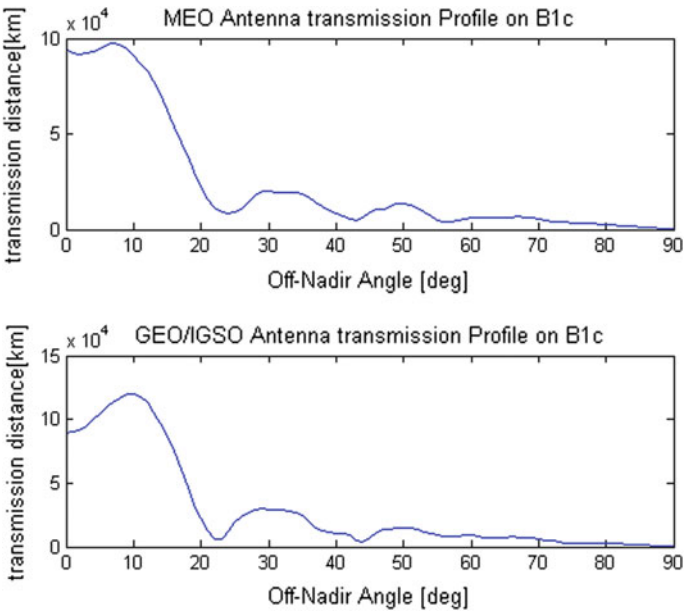


Fig. 4 Antenna transport distance

Table 2 Angle of main lobe and side lobes

	Main lobe	First side lobe	Second side lobe	Third side lobe
MEO antenna	0°–22°	26°–40°	45°–55°	60°–70°
GEO antenna	0°–21°	24°–36°	45°–55°	–

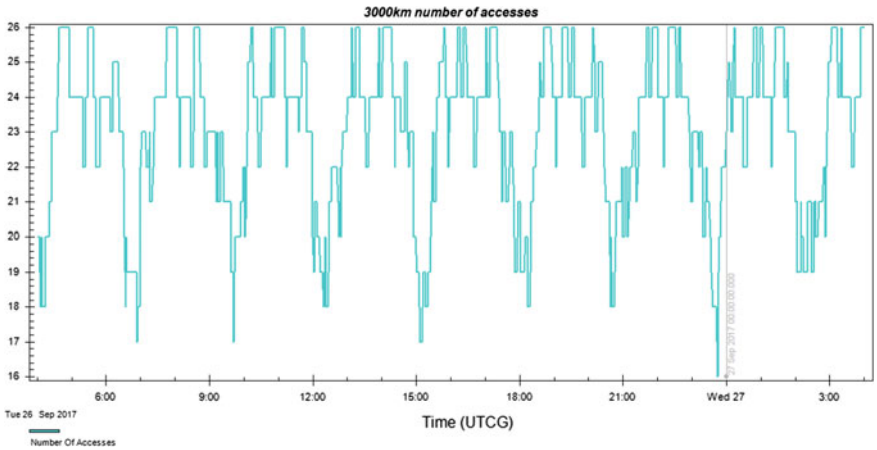


Fig. 5 3000 km number of accesses

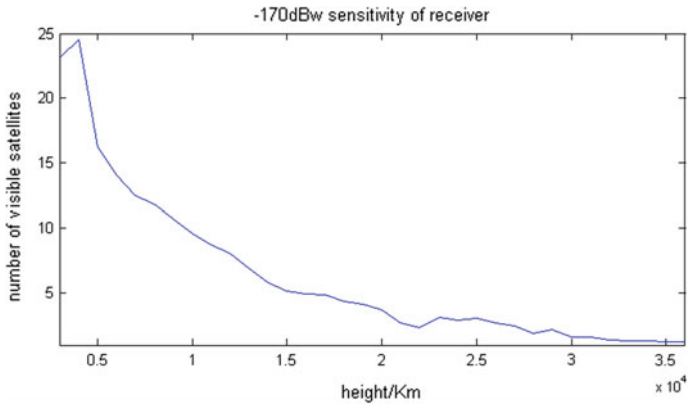


Fig. 6 Visibility varies with height

In order to study the limit of BDS-3 space service volume, we design the receiver sensitivity by taking the height at the boundary height of SSV. That is, the number of visible satellites at 36,000 km is greater than 4, and then we can tell the capability of SSV is excellent. By adjusting the receiver sensitivity, the relationship between the availability and sensitivity of the satellite can be obtained in Table 3.

Figure 7 shows that the greater the sensitivity is, the better the SSV capability is. With the increase of the sensitivity, −183 dBW shows the maximum availability, and then visible period no longer increase. Although the visible number of satellites is increasing, it is unnecessary to continue to increase sensitivity at the expense of a high price. At −173 to −180 dBW, the visibility does not increase with the increase of sensitivity, and it can be known that the performance improvement of this

Table 3 Visibility varies with receiver sensitivity

Sensitivity/dBW	Average visible satellites	Maximum visible satellites	Minimum visible satellites	Visible time of all simulation period/%	Maximum visible time/s	Maximum invisible time/s
−170	1.34	4	0	12.88	5381.07	45469.58
−171	2.87	6	0	25.53	8736.88	28753.37
−173	4.40	8	2	69.13	15535.72	4256.31
−175	4.40	8	2	69.13	15535.72	4256.31
−178	4.40	8	2	69.13	15535.72	4256.31
−180	4.40	8	2	69.13	15535.72	4256.31
−182	6.34	10	2	79.45	49574.67	2875.40
−183	8.17	12	2	99.69	86131.81	268.19
−185	11.90	17	5	100	86400	0
−187	12.64	19	5	100	86400	0
−190	14.24	21	5	100	86400	0

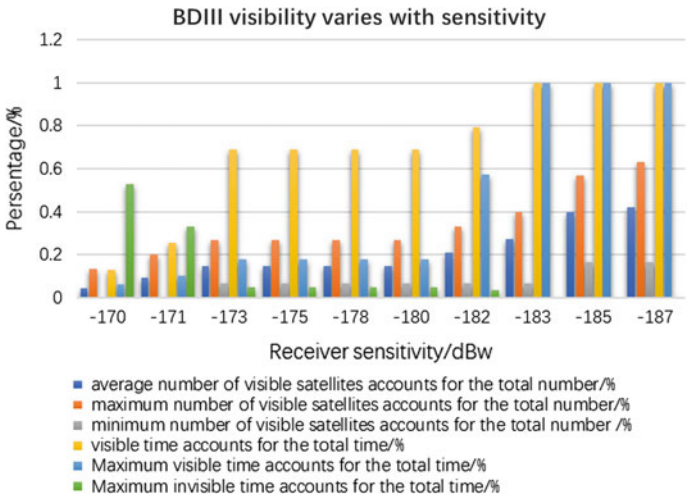


Fig. 7 Visibility varies with receiver sensitivity

segment can reflect below 36,000 km. Because the sidelobe signal does not extend to the GEO orbit, so the performance of the GEO does not increase.

Because of the increase of sensitivity, the sidelobe signals can be used, and the number of visible satellites becomes more. It is obvious that the side lobe signal is very important to the visibility of BDS-3. Now, we study the influence of the sidelobe signals from the visible time, the visible number of satellites and the position dilution of precision (PDOP) (Fig. 8).

When the main lobe signal and sidelobe signals are both used in the simulation in 36,000 km height, the service capability fills the entire space; when only use the main lobe signal, the satellite visibility is discontinuous, and it's unable to offer service. Even if the receiver sensitivity is improved, the problem cannot be solved.

Under the condition of −183 dBW sensitivity, Fig. 9 compares the average number of visible satellites using main lobe signal with both main lobe and sidelobe signals at GEO altitude. With the existence of sidelobe signals, the number of visible satellites is significantly improved.

In navigation, the concept dilution of precision (DOP) is generally used to represent the magnification of error. The DOP can be obtained from the coefficient matrix.

$$G = \begin{bmatrix} -\cos\theta^{(1)}\sin\alpha^{(1)} & -\cos\theta^{(1)}\cos\alpha^{(1)} & -\sin\theta^{(1)} & 1 \\ -\cos\theta^{(2)}\sin\alpha^{(2)} & -\cos\theta^{(2)}\cos\alpha^{(2)} & -\sin\theta^{(2)} & 1 \\ \cdots & \cdots & \cdots & \cdots \\ -\cos\theta^{(N)}\sin\alpha^{(N)} & -\cos\theta^{(N)}\cos\alpha^{(N)} & -\sin\theta^{(N)} & 1 \end{bmatrix} \tag{6}$$

$\theta^{(n)}$ and $\alpha^{(n)}$ are elevation angle and azimuth angle of each satellite respectively.

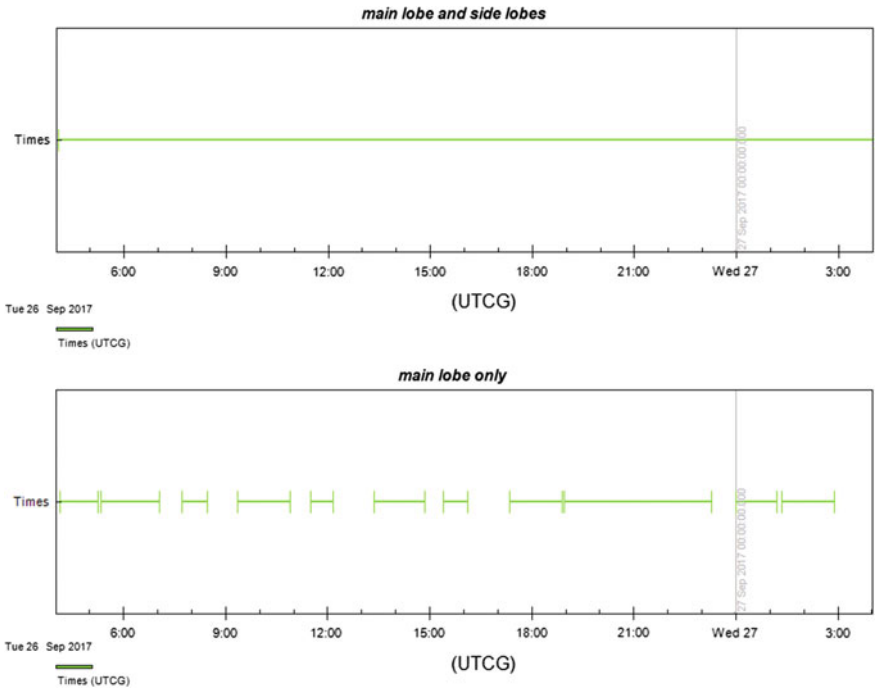


Fig. 8 Main lobe and side lobes visibility varies with time

$$\begin{aligned}\operatorname{cov}\left(\begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \varepsilon_z\\ \varepsilon_{\delta tu}\end{bmatrix}\right) &= E\left(\begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \varepsilon_z\\ \varepsilon_{\delta tu}\end{bmatrix}\begin{bmatrix}\varepsilon_x & \varepsilon_y & \varepsilon_z & \varepsilon_{\delta tu}\end{bmatrix}\right)\\ &= E\left(\left(G^TG\right)^{-1}G^T\varepsilon_\rho\left(\left(G^TG\right)^{-1}G^T\varepsilon_\rho\right)^T\right)\\ &= \left(G^TG\right)^{-1}\sigma_{URE}^2 = H\sigma_{URE}^2\end{aligned} \tag{7}$$

The error covariance matrix is on the left of the upper equation. The elements on the diagonal $\sigma_x^2, \sigma_y^2, \sigma_z^2, \sigma_{\delta tu}^2$ are the components of the error. Suppose h_{ii} represents the diagonal components, the following relation exists:

$$\begin{bmatrix}\sigma_x^2 & & & \\ & \sigma_y^2 & & \\ & & \sigma_z^2 & \\ & & & \sigma_{\delta tu}^2\end{bmatrix} = \begin{bmatrix}h_{11} & & & \\ & h_{22} & & \\ & & h_{33} & \\ & & & h_{44}\end{bmatrix}\sigma_{URE}^2 \tag{8}$$

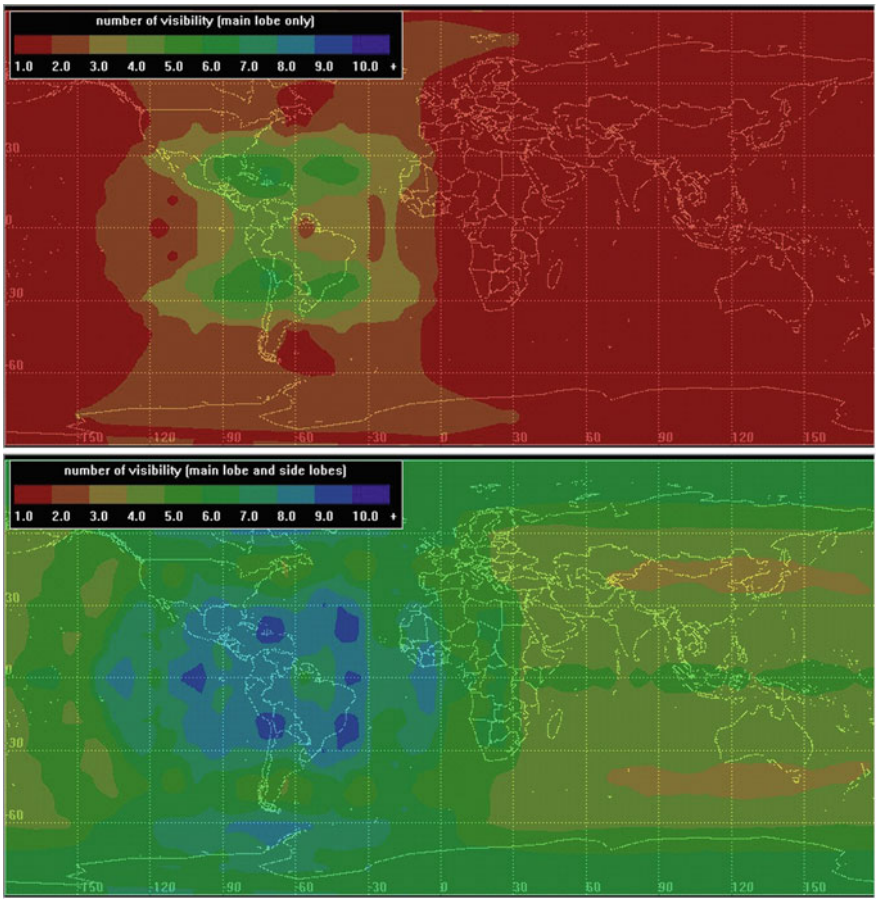


Fig. 9 Main lobe and side lobes visibility varies with time

Expressions of PDOP and position error are as follow:

$$\begin{aligned} PDOP &= \sqrt{h_{11} + h_{22} + h_{33}} \\ \sigma_p &= PDOP \cdot \sigma_{URE} \end{aligned} \tag{9}$$

Through the Eq. (9), the value of PDOP in space can be calculated. From Fig. 10, it is found that the PDOP values of the main lobe signals are larger than that of main lobe and side lobe signals. The fluctuation of main lobe signal is also larger and unstable. However, the PDOP value is decreased a lot by adding side-lobes in, which indicates the positioning error diminished, and the positioning accuracy is improved. Further verify the importance of sidelobe signals.

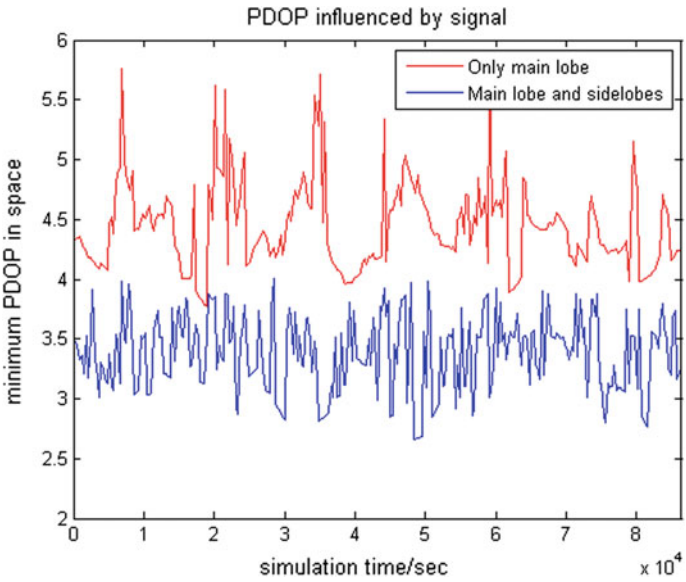


Fig. 10 Main lobe and side lobes influence on PDOP

3.2 *Receiver Gain Design*

In order to study the design of receiver antenna, this paper simulated and analyzed the direction of received power of antenna on geosynchronous orbit. The average received power is also calculated (Fig. 11).

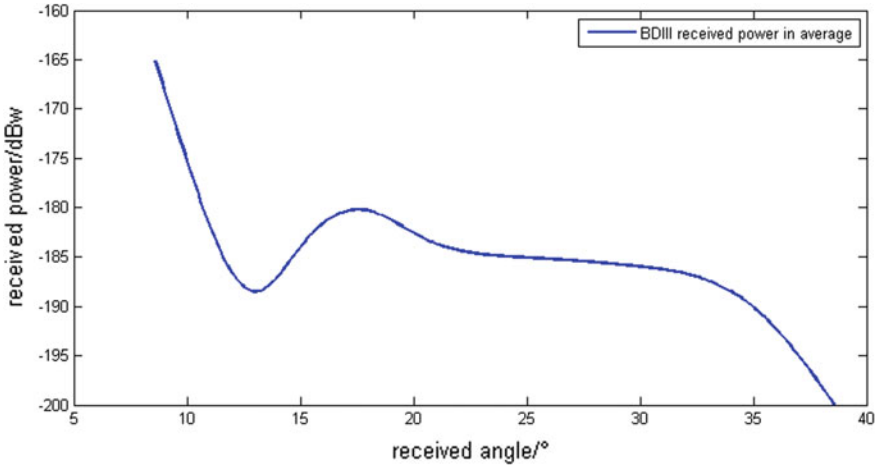


Fig. 11 Receiver power angle

Due to the block of the earth geometry, there is no signal received under 8.49°. Because BDS-3 is composed of MEO, GEO and IGSO, the orbit is different. As a result, the received power can be different in the same direction. The average power is calculated as below.

According to the direction characteristics, the following points should be taken into account in the design of receiver antenna:

(1) The antenna is affected by the earth within 8.49°, so there is no need to design antenna gain, or we can design the spaceborne receiver antenna more flexible in this period;
(2) The signals received by the receiver are mainly concentrated in the range of 10°–40°. In the medium and high orbit, the receiver antenna should be designed as narrow beam;
(3) There will be a trough in the received power of the antenna at 10°–20°, and the antenna gain can be appropriately improved in the design of this period of antenna gain.

3.3 Receiver Design Combined GPS with BDS-3

In 2015, under the support of Lockheed Martin technology, GPS announced the GPS Block IIR and IIR-M satellite antenna pattern. The detailed data of the gain and phase of each satellite antenna of L1 and L2 band are given. It can be used for the task analysis and receiver design reference of GPS SSV users. Combining BDS-3 and GPS constellation, the increase of the number of satellites can improve the space service capability [11] (Table 4).

Through modeling and simulation, the GEO coverage in the case of joint GPS and BDS-3 is obtained.

It can be obtained from Table 5, the GPS SSV capability is a little inferior compared with BDS-3. This is due to the GEO/IGSO satellites of BDS-3, that improves the service performance. However combined GPS and BDS-3, the

Table 4 GPS constellation simulation parameters

	GPS
Sources of date	Broadcast ephemeris
Inclination	55°
Orbit height	20,200 km
Periodic regression	1/2 sidereal day
Number of satellites	31
Simulation start time (UTCG)	2017-09-26 04:00:00
Simulation stop time (UTCG)	2017-09-27 04:00:00
Total time	24 h

Table 5 Visibility varies with different constellation

	GPS			GPS + BDS-3		
Sensitivity/dBW	Average visible satellite	Visible time of all simulation period/%	Minimum visible satellites	Average visible satellite	Visible time of all simulation period/%	Minimum visible satellites
−170	1.17	10.24	0	2.31	36.34	0
−173	3.43	65.23	0	4.74	57.69	2
−175	3.43	65.23	0	5.85	100	4
−177	3.43	65.23	2	6.08	100	4
−180	4.80	79.90	2	7.20	100	5
−183	6.14	97.55	2	12.31	100	5

requirement of receiver sensitivity design is more loose and flexible. Combined system can supply service with a low sensitivity of −175 dBW only.

For the spaceborne receiver in geosynchronous orbit, when the receiver sensitivity reaches −175 dBW, GPS/ BDS-3 combined system can ensure the service all the time. But even if the sensitivity increased to −183 dBW, GPS still cannot offer this kind of service.

4 Conclusion

In this paper, the software STK is used as the core of link computing. By using the analysis function of MATLAB, the design of BDS-3 spaceborne receiver in the space service volume is studied. And the following conclusions can be obtained:

(1) The application of spaceborne BDS-3 receiver is feasible, and the sensitivity of it should be above −183 dBW.
(2) The receiver antenna applying for the space service volume should be designed as a narrow beam and high gain pattern. The center beam gain is not required due to the block of the earth, and the antenna gain should be increased in the region of 10°–20°.
(3) The sidelobe signal of BDS-3 is very important under the condition of high earth orbit, so sidelobe signals must be considered in the simulation and design of the receiver.
(4) GPS is unable to meet the demand of SSV entirely even at a high sensitivity. However, utilizing BDS-3/GPS combined system can reduce the sensitivity requirement of the spaceborne receiver. And it can cover the whole space service volume in −175 dBW only.

References

1. Bauer FH, Moreau MC, Dahle-Melsaether ME et al (2006) The GPS space service volume. In: Proceedings of international technical meeting of the satellite division of the institute of navigation
2. Honglei Qin, Minmin Liang (2008) Research on GNSS high orbit satellite positioning technology. J Space Sci 28(4):316–325
3. Bauer FH, Parker JK, Welch B et al (2017) Developing a robust, interoperable GNSS space service volume (SSV) for the global space user community. ION Int Tech Meet
4. Jing S, Zhan X, Lu J et al (2015) Characterisation of GNSS space service volume. J Navig 68 (1):107–125
5. Zhenhai Li, Wenhai Jiao, Xiaorui Huang et al (2013) Comparison and analysis of signal availability of space service volume of GNSS. J Astronaut 34(12):1605–1613
6. China Satellite Navigation Office (2017) Beidou system space signal interface control file on signal B1C, B2a (test) [EB/OL]. www.beidou.gov.cn
7. Chen K (2013) Research on key technologies of low power and high sensitivity GPS receiver design. Shanghai Jiao Tong University
8. Mo J (2011) Research on the key technology of high performance GPS receiver in Weak signal environment. Xidian University
9. Changting Cai, Yiming Pi (2006) Research on progress of high sensitivity GPS technology. Global Positioning Syst 31(2):1–4
10. Chen S, Liang Z, Zhang M (2014) Using STK to design spaceborne GPS receiver. Syst Simul J 26(1)
11. Yanzhen Zhao (2016) Research on high orbit satellite positioning technology based on GPS/ Beidou integrated system. J Space Sci 36(1):77–82

A Low Complexity Algorithm for Code Doppler Compensation Using FFT-Based Parallel Acquisition Architecture

Ping Tang, Xiangming Li, Shuai Wang and Ke Wang

Abstract With the increasing competition in Global Navigation Satellite System (GNSS), the development of satellite navigation industry has become a big challenge for our country in recent years. Focusing on the development trend and effective demand of BeiDou system (BDs), a low complexity algorithm for code Doppler compensation using FFT-based parallel code searching architecture is proposed to correct Pseudo-Random Noise (PRN) code phase migration, which is caused by the code Doppler in the acquisition process of navigation signal in high dynamic and low Carrier-Power-to-Noise-Density Ratio (C/N0) environment. The new algorithm that is named as "Modified Estimation and Circular Shift of Code Phase (MECS-CP)" could cut down the influence of code Doppler on accumulation peak so as to increase the acquisition performance effectively. The MECS-CP algorithm is presented on the foundation of FFT-based full parallel acquisition algorithm. Firstly, the direction and magnitude of PRN code phase migration are estimated by the linear relationship between carrier Doppler and code Doppler in the process of correlation and coarse carrier Doppler compensation. Then, the PRN code correlation results are circularly shifted in phase dimension according to the estimation results, and then all corrected correlation outputs are kept in a three-dimensional array for later use. Finally, fine carrier Doppler compensation and coherent accumulation are carried out. This paper verifies the accuracy of code Doppler compensation, acquisition performance and computational complexity of the proposed MECS-CP algorithm by theoretical analyses and Monte-Carlo simulations. Both results show that the new algorithm could efficiently compensate the code Doppler and sharpen the accumulation peak of weak navigation signals in high dynamic environment, thus improving the peak-to-average power ratio

P. Tang (✉) · X. Li · S. Wang · K. Wang
School of Information and Electronics, Beijing Institute of Technology, Beijing, China
e-mail: tangping@bit.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_31

(PAPR) and detection rate. Numerically, it is found that the proposed algorithm can achieve a detection rate of 0.9 and a false alarm rate of 10^{-5} at C/N0 = 31dBHz during a coherent accumulation time of 32 ms in simulation.

Keywords High dynamic · Low C/N0 · FFT-based full parallel acquisition algorithm · Code Doppler compensation

1 Introduction

With the development of satellite navigation technology, the GNSS, which has been widely used in various fields, shows its great military, economic and social benefits. BDs is an independent construction and operation satellite navigation system in China. Improving the application performance of BDs in complex scenes is an urgent practical demand for Chinese satellite navigation system [1]. On account of the increasing sophisticated technology of navigation signal processing and the growing application of GNSS in the environment, where signal quality suffers serious attenuation, such as high-rise city streets, viaducts and mountain areas etc., high sensitivity receiver has become a research hotspot in recent years. Long accumulated time is necessary for high sensitivity receiver to acquire weak signals, while the code phase migration between the local PRN code and received signal caused by code Doppler, leads to the shift, broadening and attenuation of accumulation peak [2]. The impact of code Doppler would be more obvious when the Doppler is large [3]. If the code phase migration is more than 4/3 chips, the longer accumulation time is, the lower gain and worse acquisition performance will be [4]. Therefore, code Doppler compensation is great important for GNSS receivers in low C/N0 and high dynamic environment.

The code Doppler can be compensated by searching the code Doppler directly to eliminate the phase migration, but the calculation is complicated for introducing another dimension into the two-dimensional search of navigation signal [4]. Another intuitive method is to change the local PRN code frequency while changing the carrier Doppler search step using the proportional relationship between the carrier Doppler and code Doppler [5]. However, the change in local PRN code frequency brings about variation on the code length. Consequently, the filter coefficients of Digital Match Filter (DMF) and the operation points of the FFT-based circular correlation need refreshing constantly, which makes the hardware implementation of parallel code acquisition algorithm become extremely difficult [6, 7]. The code Doppler can also be compensated by correcting the migration of PRN code phase, such as spectral peak tracking, Hough transform method and so on, but these algorithms are only suitable in high C/N0 situation [8].

The MECS-CP algorithm proposed in this paper could efficiently compensate the code Doppler and sharpen the accumulation peak of weak navigation signals in high dynamic environment, thus improving the acquisition performance. Moreover, the computational complexity of the new algorithm is so low that it is easy to

implement. Firstly, on the foundation of the system model, the MECS-CP algorithm is described in detail. Then, the accuracy of code Doppler compensation, acquisition performance and computational complexity of the new algorithm are analysed theoretically, and verified by Monte Carlo simulations. Finally, a conclusion is made.

2 System Model

Figure 1 is the block diagram of the FFT-based full parallel acquisition algorithm. $s(iT_s)$ denotes the I/Q baseband data of sampled signal through orthogonal down-conversation and low-pass filter, and the expression is,

$$s(iT_s) = Ac(iT_s - \tau + \xi iT_s)e^{j(2\pi f_d iT_s + \theta_0)} + n_0(iT_s) \quad (1)$$

where $A > 0$ is the amplitude of the signal and $c(\cdot)$ is a PRN code, $T_s = \frac{1}{f_s}$ is the sampling period. We further assume $c(\cdot)$ has L_c chips. Given the chip duration of $c(\cdot)$ denoted by T_p, the length of the PRN code is $T_c = T_p L_c$. Furthermore, τ represents the delay of the received signal with respect to the local replica of the PRN code, $\xi = \frac{f_d}{f_r}$ is the normalized code Doppler, f_r is the carrier frequency, f_d is

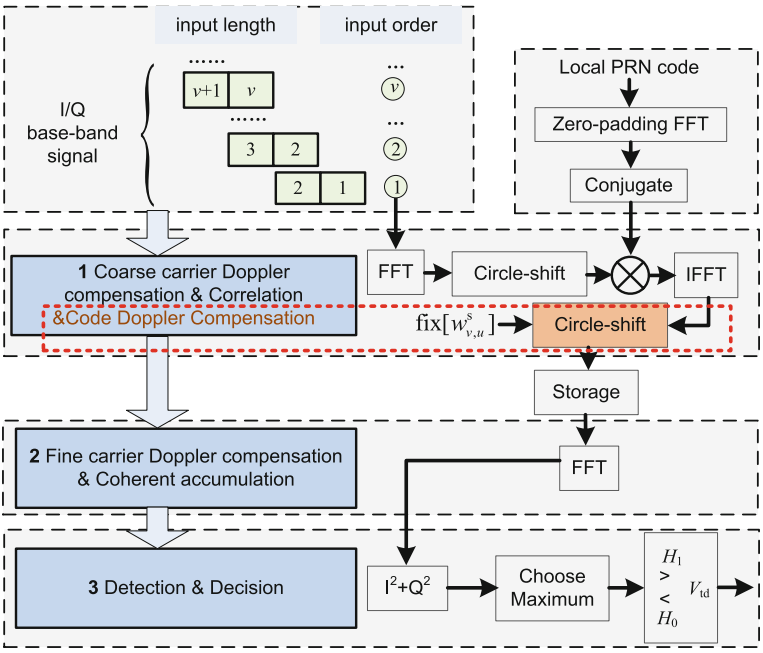


Fig. 1 Block diagram of the FFT-based full parallel acquisition algorithm [9]

the carrier Doppler frequency offset, and $\theta_0 \in [0, 2\pi]$ is the random initial phase. Finally, $n_0(iT_s)$ stands for a discrete, complexed-valued, circularly-symmetric additive white Gaussian noise (AWGN) process with zero mean and a variance of σ_n^2. To explain clearly, we assume the sampling interval of the received signal is half of the chip duration, namely $T_s = 0.5T_p$, which is sufficient for the subsequent tracking loop to take over. So there are $L = 2L_c$ samples in one PRN code period.

3 Description and Implementation of New Algorithm

3.1 Description of New Algorithm

The FFT-based full parallel acquisition algorithm contains three stages, as illustrated in Fig. 1, namely Stage (1) Correlation and coarse carrier Doppler compensation; Stage (2) Fine carrier Doppler compensation and coherent accumulation, and Stage (3) Signal detection by threshold test [9].

In stage (1), the correlation is done in frequency domain by FFT, and the coarse carrier Doppler compensation is accomplished by circularly shifting the input signal's Fourier coefficients forwards and backwards. In order to facilitate the computation of FFT, the input length N_{FFT} of each correlation operation is typically an integer power of 2. On the other hand, the input data is essential to last for at least two complete code periods to improve the accuracy of coarse carrier Doppler compensation, thus reducing the sinc loss is by 3 dB. In addition, since the local PRN code replica includes one complete period of the PRN code, only the first L entries of the IFFT output are valid correlation results, which need storing in a three-dimensional array for later use. The mathematical expression of the valid correlation results is

$$x(l, v, u) = AR(\rho)\,\mathrm{sinc}(\pi\Delta f_d T_c)e^{j\phi_{l,v,u}} + n_1(l, v, u) \tag{2}$$

Note that each $x(l, v, u)$ is the correlation output by integration over a time span of T_c; $l = 0, 1, \ldots, L-1$; $v = 1, 2, \ldots, V$; $u = 1, 2, \ldots, U$; V represents the times of coherent accumulation; U denotes the number of circular shift in coarse carrier Doppler compensation; and in the above formula, ρ is the fractional part of $\frac{\tau}{T_p}$ and $|\rho| < 1$; $|\Delta f_d| < \frac{1}{4T_c}$ stands for the residual carrier Doppler offset, $\phi_{l,v,u}$ denotes the phase of the complex-valued correlation output, and $n_1(l, v, u)$ is the noise term in the correlation output.

Given the chip rate of $c(\cdot)$ denoted by f_c, the relationship between carrier Doppler and code Doppler f_{cd} is $f_{cd} = \frac{f_c}{f_r} f_d$. As the correlation time is T_c, the PRN code phase migration caused by code Doppler is $w^c = T_c f_{cd}$ chips, and the corresponding migration of sampling points is $w^s = \frac{T_p}{T_s} T_c f_{cd}$. For the vth correlation results, the PRN code phase migration is $w_v^c = vT_c f_{cd}$ chips and $w_v^s = \frac{T_p}{T_s} vT_c f_{cd}$ sampling points. In the FFT-based parallel acquisition algorithm, the valid results of parallel correlation show

the same phenomenon of phase migration, for the reason that the input data of each correlation operation last for two continuous periods of PRN code and this ensure the implementation of circular correlation. On the basis of above analysis, the proposed MECS-CP algorithm does not perform frequency compensation or phase correction directly on the local PRN code replica or input signal, but corrects the position of correlation peak by shifting parallel correlation results $x(0{:}L\text{-}1, v, u)$circularly according to the estimation of PRN code phase migration, thus eliminating the influence of the code Doppler on the acquisition performance.

In the process of coarse carrier Doppler compensation, the number of circular shift U is determined by the range of carrier Doppler and the accuracy of coarse compensation. As described in reference [9], the sign of $\left(u-\frac{U+1}{2}\right)$ indicates the shift direction and its magnitude stands for the number of shift bits. The physical meaning of circular shift is that the carrier Doppler frequency compensation for the signal is $\left(u-\frac{U+1}{2}\right)\frac{1}{2T_c}$. Hence, the code Doppler compensation is $\hat{f}_{cd}=\frac{f_c}{f_r}\left(u-\frac{U+1}{2}\right)\frac{1}{2T_c}$ accordingly, and then the phase required for correction for v-th correlation results is $\hat{w}^c_{v,u}=vT_c\hat{f}_{cd}$ chips, that is the estimation of code phase migration is $w^c_{v,u}=-\hat{w}^c_{v,u}$ chips. In the hardware implementation of MECS-CP algorithm, all the signal processing objects are sampled discrete data and the circular shift is fixed-point operation. On one hand, the value of phase correction for (v, u)th correlation results at sampling rate is $\hat{w}^s_{v,u}=\frac{vT_pf_c}{2T_sf_r}\left(u-\frac{U+1}{2}\right)$, and the direction of circular shift is also determined by the sign of $\left(u-\frac{U+1}{2}\right)$. On the other hand, rounding is essential for $\hat{w}^s_{v,u}$ to implement the operation of circular shift.

There are two categories of rounding methods in common use: rounding to the nearest integer and rounding to the integer in a certain direction. If the method of rounding to the nearest integer is applied, the correlation peaks through phase correction will jump back and forth in two adjacent phases along with the accumulation time, and then two adjacent peaks whose amplitudes are almost equal appear in the accumulation results, as shown in Fig. 2. The two accumulation peaks almost share the signal power equally, so that the PAPR is reduced about 3 dB.

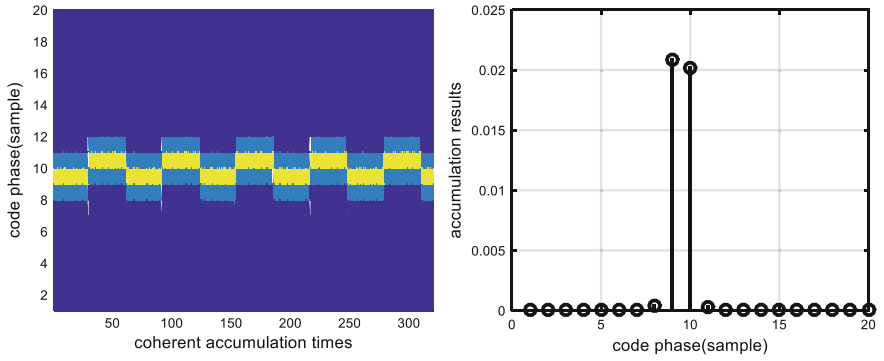


Fig. 2 Trajectory of correlation peak after phase correction and normalized coherent accumulation result without noise (rounding to the nearest integer)

Therefore, the method of rounding to the nearest integer is not suitable to modify the value of phase correction, and then the rounding method is chosen from the category of rounding to integer in a certain direction: rounding toward positive infinity, rounding toward negative infinity and rounding toward zero. Considering the symmetry in positive and negative direction, the last one, namely rounding toward zero, is adopted in the MECS-CP algorithm, and represented by the symbol of fix[·].

3.2 Implementation of New Algorithm

In summary, the implementation steps of the proposed algorithm of MECS-CP are as follows,

(1) For the (v, u)th parallel correlation results $x(0{:}L\text{-}1, v, u)$, the magnitude and sign of phase required for correction are calculated according to formulas $\hat{f}_{\mathrm{cd}} = \frac{f_{\mathrm{c}}}{f_{\mathrm{r}}}\left(u - \frac{U+1}{2}\right)\frac{1}{2T_{\mathrm{c}}}$ and $\hat{w}^{\mathrm{s}}_{v,u} = \frac{T_{\mathrm{p}}}{T_{\mathrm{s}}} v T_{\mathrm{c}} \hat{f}_{\mathrm{cd}}$. And then $\hat{w}^{\mathrm{s}}_{v,u}$ is rounded toward zero resulting in integer, that is $\mathrm{fix}[\hat{w}^{\mathrm{s}}_{v,u}]$. In accordance with the aforementioned operations, $V \times U$ modified values of phase correction are calculated successively, and all pre-stored in the memory at the corresponding address (v, u), as shown in Fig. 3.

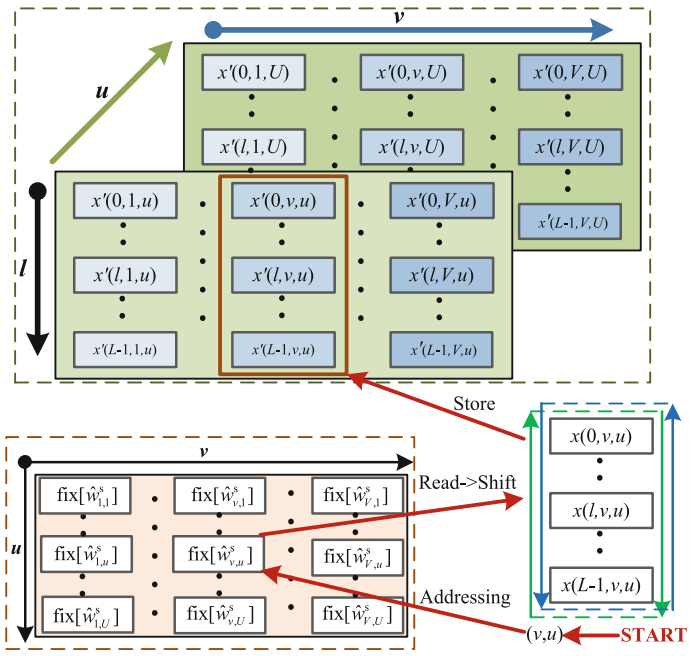


Fig. 3 Flow chart of MECS-CP algorithm

(2) In the process of correlation and coarse carrier Doppler compensation, the modified value of phase correction fix$[\hat{w}^{s}_{v,u}]$ is taken out from the mapped address space when the (v, u)th correlation results $x(0{:}L\text{-}1, v, u)$ are exported. The correlation ouput $x(0{:}L\text{-}1, v, u)$ are circularly shifted according to the value of fix$[\hat{w}^{s}_{v,u}]$, and the results $x'(0:L\text{-}1, v, u)$ are saved in the three-dimensional array preparing for the next stage of accumulation.

4 Performance Analysis and Simulation Results

4.1 Compensation Performance

The accuracy of code Doppler compensation is closely related to the accuracy of coarse carrier Doppler compensation for the proposed the MECS-CP algorithm. The accuracy of coarse compensation of FFT-based full parallel acquisition algorithm is $\frac{1}{2T_c}$, so then the accuracy of code Doppler compensation is $\frac{f_c}{2T_c f_r}$.

The phase correction accuracy is also an effective index for the algorithms, which compensate the code Doppler by correcting the phase migration, because this index influences the shape of accumulation peak and the acquisition performance in direct. The phase correction accuracy of the proposed MECS-CP algorithm is generally determined by the ratio of the sampling rate to the chip rate of PRN code, which is $\frac{T_s}{T_p}$ chips. It has been assumed earlier in the article about $T_s = 0.5T_p$, so the accuracy of phase migration is 0.5 chips.

Next, MATLAB simulations are done to analyze and evaluate the performance of the MECS-CP algorithm which employs parallel code searching architecture. The input signal is shown as Eq. (1), and main parameters are the carrier frequency $f_r = 1268.52$ MHz, the chip rate of PRN code $f_c = 10.23$ MHz [10], the length of PRN code $L_c = 1023$, the correlation time $T_c = 0.1$ ms, the carrier Doppler $f_d = 10$ kHz, and the sampling rate $f_s = 20.46$ MHz.

Hence, the code Doppler is $f_{cd} \approx 80$ Hz, and the phase migration during one span of T_c is $w^c = T_c f_{cd} \approx 0.008$ chips. If the whole accumulation time is 32 ms, the phase migration of the 320th correlation results is $w^s_{32} = 2 \times 320 w^c \approx 5$ samples, namely 2.5 chips. It is obvious that the accumulation time of 32 ms is far greater than the optimal accumulation time [4]. If no measures are executed to compensate the code Doppler or to correct the phase migration, multiple adjacent peaks will appear in the accumulation results. It can be observed visually from Fig. 4 that the position of correlation peak migrated along with time, and the trajectory shows a downward trend. Without phase correction, the main-lobe broadening and peak reduction occur in the result of coherent accumulation.

Whereas, the Fig. 5 illustrates the trajectory of correlation peaks and accumulation peak when the proposed algorithm of MECS-CP is adopted. Since the new algorithm corrects the phase migration of PRN code correlation results in direct,

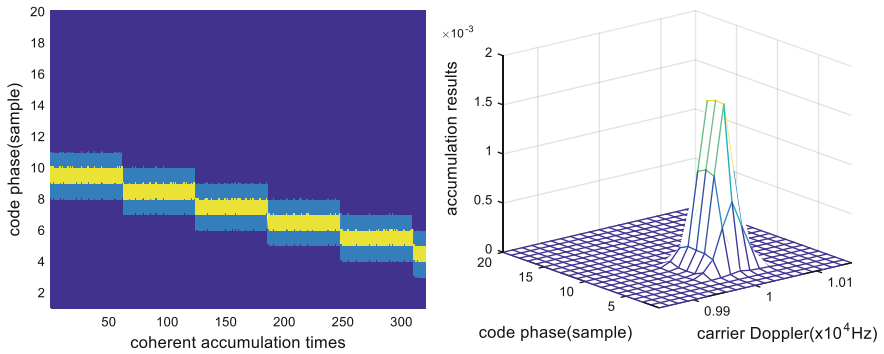


Fig. 4 Trajectory of correlation peak and normalized coherent accumulation result without phase correction and noise

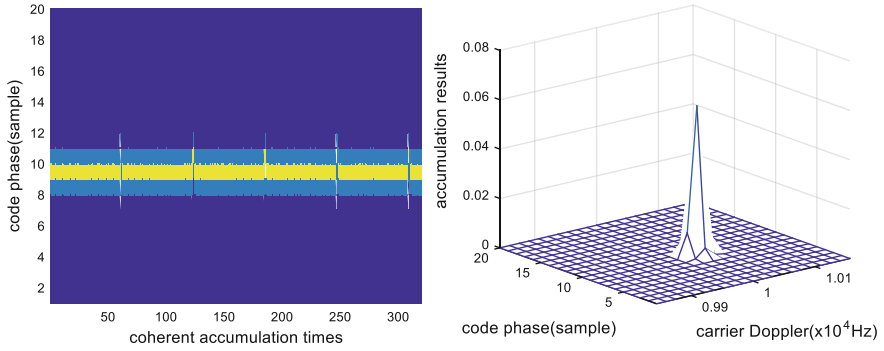


Fig. 5 Trajectory of correlation peak after phase correction and normalized coherent accumulation result when MECS-CP algorithm is adopted and noise is absent

which eliminates the offset of the correlation peak position, the trajectory of correlation peaks is approximately parallel to the X-axis. Moreover, it is noticeable that the accumulation peak is sharper, and is far higher than that of Fig. 4 due to the fact that the signal power is concentrated at the peak.

In order to further verify the performance of MECS-CP algorithm, Monte Carlo simulations of the PAPR and detection rate P_d under different C/N0 are conducted. It is required that the false alarm rate $P_{fa} = 10^{-5}$ and the estimation error of PRN code phase is less than 1 chip. The simulation results are shown in Fig. 6.

When the coherent accumulation time is 16 ms, the phase migration is about 1.3 chips. Then, the signal can still be acquired by increasing the C/N0 instead of compensating the code Doppler, and the performance is about 3 dB worse than that of the case with phase correction as shown in the Fig. 6. For the case with phase correction, the PAPR when the coherent accumulation time is 32 ms is about 3 dB larger than that when the coherent time is 16 ms, and the detection rate is about 3 dB better as well. Furthermore, the simulation curves through phase correction

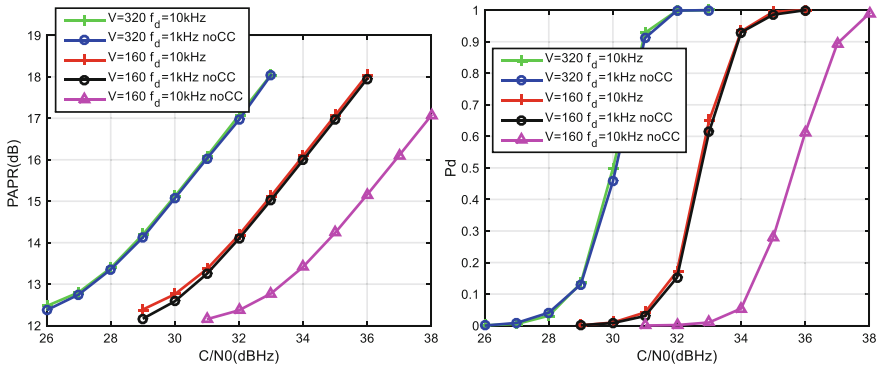


Fig. 6 Relationship between PAPR, P_d and C/N0, where "noCC" denotes there are no methods of code Doppler compensation adopted in acquisition

when the carrier Doppler is 10 kHz are almost coincided with the curves without code Doppler compensation when the carrier Doppler is 1 kHz, respectively. The above results illustrate that the MECS-CP algorithm is able to correct the phase migration caused by code Doppler effectively, so that the weak signal can be accumulated during a long time without introducing additional loss. It can be seen from the Fig. 6 that P_d of the proposed algorithm can reaches 0.9 when $P_{fa} = 10^{-5}$ and C/N0 = 31 dBHz during a coherent accumulation time of 32 ms in simulation.

4.2 Computational Complexity

The MECS-CP algorithm presented in this paper is based on code parallel acquisition architecture in terms of the previous description and analysis. The phase correction is realized primarily by circularly shifting the PRN code correlation results, so as to correct the phase migration resulting from the code Doppler. The estimation and modification of the phase migration are accomplished on computer in advance, and pre-stored in memory. In hardware implementation, the modified value of phase correction $\mathrm{fix}[\hat{w}^s_{v,u}]$ is taken out from the mapped address space in the light of (v, u). The operation of circular shift is combined with the stored procedure of correlation results, thus consuming no logic resources. In a word, the main consumption of MECS-CP algorithm is memory resources, and the size of storage space is VUN_{bit}, in which N_{bit} is the bit width of $\mathrm{fix}[\hat{w}^s_{v,u}]$.

5 Conclusions

This paper presents a low complexity algorithm, namely MECS-CP, to solve the problem of PRN code phase migration caused by code Doppler in the acquisition process for navigation receiver under the condition of large dynamic and low C/N0.

The new algorithm is carried out on the parallel code searching architecture. The phase required for correction is calculated according to the formulas $\hat{w}^{s}_{v,u} = \frac{vT_p f_c}{2T_s f_r}\left(u - \frac{U+1}{2}\right)$, and $\hat{w}^{s}_{v,u}$ is modified by rounding toward zero. Most importantly, the position of correlation peak is corrected by circularly shifting the parallel correlation results. The accuracy of code Doppler compensation of MECS-CP algorithm is $\frac{f_c}{2T_c f_r}$, and the phase correction accuracy is $\frac{T_s}{T_p}$ chips. The simulation results demonstrate that, the new algorithm could efficiently correct the phase migration of parallel correlation results and sharpen the accumulation peak, thus improving the PAPR and detection rate. In particular, it is found that the FFT-based full parallel acquisition algorithm with code Doppler compensation can achieve a detection rate of 0.9 and a false alarm rate of 10^{-5} at C/N0 = 31dBHz during a coherent accumulation time of 32 ms in simulation. In addition, only memory resources are consumed in hardware implementation, MECS-CP algorithm is proved to be a simple but efficient code Doppler compensation method.

Acknowledgements This work is financially supported by NSFC (The Joint Foundation of NSFC & Fundamental Research for General Purpose Technologies) under grant No. U1636125.

References

1. Chengqi R (2012) Development of the BeiDou navigation satellite system. In: Global navigation satellite systems. Report of the Joint Workshop of the National Academy of Engineering and the Chinese Academy of Engineering. Washington, DC
2. Ziedan NI (2006) GNSS receivers for weak signals. Artech House Publishers
3. Li H, Zhou H, Lu M et al (2010) Code-Doppler-compensation based direct acquisition method for weak GNSS long PN-code. Sci Sin Phys Mech Astron 5:012
4. Luo H, Wang Y, Luo Y et al (2013) Analysis and compensation for the effect of code Doppler frequency in PMF-FFT algorithm. Trans Beijing Inst Technol 33(11):1176–1182
5. Ma Y, Zhangyi Y, Zhang Z et al (2009) Modified method of high dynamic & high sensitivity GPS signal acquisition. Syst Eng Electron 31(2):265–269
6. Wang Y, Tian J, Leclère J et al (2014) An efficient time-frequency algorithm for the weak signal acquisition of modernized GNSS signals. ION GNSS 57(6):715–722
7. Huang X, Gong H, Zhu X et al (2015) Code Doppler compensation method for GNSS signal acquisition. J Cent South Univ Sci Technol 46(6):2134–2141
8. Zhang L, Zhang T, Li H et al (2013) Research on the effect of code doppler on acquisition performance and the compensation methods. In: China Satellite Navigation Conference (CSNC) 2013 Proceedings. Springer, Berlin, Heidelberg, pp 663–673
9. Tang P, Wang S, Li X et al (2017) A low-complexity algorithm for fast acquisition of weak DSSS signal in high dynamic environment. GPS Solutions 21(4):1427–1441
10. Chen W, Li X (2015) Accuracy assessment for carrier phase and code measurements of BDS B1/B3 frequency observation. Aerosp Shanghai 32(6):62–66

Non-binary LDPC APP-Based Decoding for BDS Navigation Data

Jiaju Song

Abstract Non-binary low density parity check (LDPC) codes over GF(64) are applied to channel coding scheme for navigation data in Beidou system (BDS) open service signal B1C and B2a. This paper reviews non-binary LDPC Extended Min-Sum (EMS) soft decision decoding algorithms, and several equivalent implementations of the decoder due to different definitions of log-likelihood ratio. Considering the characteristics of the ultra-sparse regular code used by BDS, the structure of the decoder can be simplified. Then a posteriori probability (APP) based decoding method for non-binary LDPC code is proposed, in which a posteriori probability is used not only for decision output, but also as the message input in next iteration. So complete storage of the messages between the variable and check nodes is not necessary, less memory is required, and variable node update processing is simplified. Finally, the performance of several decoders is evaluated.

Keywords Non-binary low density parity check decoders · APP decoding Iterative decoding · Belief propagation

1 Introduction

Beidou satellite navigation system (BDS) open service signal B1C and B2a use Non-binary Low-density Parity-check (NB-LDPC) codes in channel coding scheme [4]. Compared with binary LDPC, NB-LDPC has better performance, better anti-burst error ability and lower error floor, especially for medium or short code length. But it also brings engineering issues such as algorithm complexity and memory usage. Up to now, three different NB-LDPC code are deployed in BDS navigation data channel, with code length N = 200, 88 and 96. All three NB-LDPC

J. Song (✉)
Guangzhou Haige Communications Group Inc, Guangzhou, China
e-mail: nmjnsjj@mail.nwpu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_32

codes are regular codes, have same column weight $d_c = 4$ and row weight $d_v = 2$, constructed over Galois field GF(64) with primitive polynomial $p(x) = 1 + x + x^6$.

NB-LDPC is a generalization and extension of binary LDPC over Galois fields. There are two types of decoding algorithms: bit flipping (BF) and belief propagation (BP). It results in a large number of different decoding algorithms, by applying different implementations, improvements and simplifications on BF and BP. As a simplification of the BP method called Sum-Product algorithm (SPA), the log-domain min-norm algorithms family, including the Min-Sum (MS) algorithm and the Min-Max algorithm, with its excellent performance, has been widely used in binary LDPC decoding [10].

However, both SPA and log-domain min-norm algorithms family have encountered time and space complexity issues during extension to non-binary situation. To reduce complexity, fast Fourier transform (FFT) is used to process convolution in SPA [10], and Declercq proposed the Extended Min-Sum (EMS) [5]. In the EMS algorithm, probability messages over GF(q) are truncated to size n_m, where $n_m << q$. By ignoring $q - n_m$ obviously impossible elements in a message, NB-LDPC can eventually be deployed in a real-time communication system.

A posteriori probability (APP) decoding is derived from a posteriori probability iteration of one-step majority-logic decoding. The prior probabilities of all possible values of the codeword elements are obtained from the soft-decision sequence, and the corresponding posterior probabilities are calculated by decoder. Then, the set of posterior probabilities is used as new a prior probability, and the decoding process is repeated.

During APP decoding iteration, a posterior probability has correlation with the initial value, which causes performance degradation [8]. In order to avoid this correlation, the BP algorithm introduces a large amount of memory for storing probability data. In literatures about binary LDPC decoding algorithm, the min-norm algorithm family, especially the Min-Sum algorithm, which are constructed on the APP and the BP decoding, are often referred to as "APP-based" and "BP-based" algorithm [3, 6].

For NB-LDPC, we can combine the idea of truncating message in EMS with the APP iteration. In the second section, the NB-LDPC logarithmic domain Min-Sum algorithm is reviewed, and then the message truncation is applied to implement the EMS algorithm. In the third section, we use a posterior probabilities column accumulation to replace the check node data processing, to derive the NB-LDPC APP-based EMS decoding algorithm. The performances of MS, EMS and APP algorithms are simulated in the fourth section. Besides, the Galois field operations, which used by NB-LDPC decoder, are given as an appendix.

2 Log-Domain Min-Sum Algorithm

2.1 Iteration Process

The log-domain Min-Sum algorithm uses logarithms of probabilities as a likelihood measurement in order to simplify the calculation and avoid the error floor caused by numerical precision. The decoding process essentially is the propagation of likelihood information on the edge of the Tanner graph defined by the parity check matrix. The algorithm steps are as follows [2, 5, 10]:

Initialization

Initialize a prior likelihood ratio of the n-th codeword taken symbol a as $\gamma_n(\mathbf{a}) = \{\gamma_n(\mathrm{a}_1), \ldots \gamma_n(\mathrm{a}_q)\}$. There are q (q = 64) possible values for each codeword symbol, which form the vector $\mathbf{a} = \{\mathrm{a}_1, \ldots \mathrm{a}_q\}$.

Select the most likely symbol as the first decision:

$$z_n = \arg\max_{\mathrm{a}}[\gamma_n(\mathbf{a})] \tag{1}$$

Messages form variable nodes (VN) to check nodes (CN) **u** are initialized simultaneously:

$$u_{n,m}(\mathbf{a}) = \gamma_n(\mathbf{a}) \tag{2}$$

Iteration

(1) Row Processing

During row processing, messages **v** passed from CN to VN are calculated row-by-row form messages **u**. It can be completed in three steps:

(1-1) Galois field elements in **u** are rearranged by multiplying each element with the corresponding item in the parity-check matrix. The Galois field element vector **a** is rearranged to vector $\tilde{\mathbf{a}}$:

$$u_{n,m}(\tilde{\mathbf{a}}) = u_{n,m}(\mathbf{a}h_{n,m}) = \{u_{n,m}(\mathrm{a}_1\mathrm{h}_{n,m}), \ldots, u_{n,m}(\mathrm{a}_q\mathrm{h}_{n,m})\} \tag{3}$$

The multiplication here is defined over the Galois field.

(1-2) Calculate **v**:

$$v_{n,m}(\tilde{\mathbf{b}}) = \min_{\mathrm{b}\in\mathrm{C}} \sum_{\substack{t \in \mathrm{R}_m \\ t \neq n}} u_{t,m}(\tilde{\mathbf{a}}) \tag{4}$$

where R_m represents the VNs connected to the m-th CN, that is, the set of column numbers of the non-zero elements in the m-th row in the check matrix. C is called a configuration set, which is a set of sequences of all finite field elements satisfying the condition $\sum_{\substack{t \in R_m \\ t \neq n}} \tilde{a}_t = \tilde{b}$. This step is implemented by forward-backward (FB) algorithm and the elementary check node (ECN) processing unit [1, 4].

(1–3) Galois field elements in messages **v** are also rearranged by dividing with the same item in (1-1).

$$v_{n,m}(\mathbf{a}) = v_{n,m}(\tilde{\mathbf{a}} h_{n,m}^{-1}) \tag{5}$$

(2) Column Processing

Calculate new messages **u** column by column:

$$u_{n,m}(\mathbf{a}) = \gamma_n(\mathbf{a}) + \sum_{\substack{t \in S_n \\ t \neq m}} v_{n,t}(\mathbf{a}) \tag{6}$$

where S_n represents the CNs connected to the n-th VN, also the set of row numbers of the non-zero elements in the n-th column in the check matrix. Considering BDS LDPC check matrices have a fixed row weight d_v = 2, rewrite is as:

$$u_{n,m'}(\mathbf{a}) = \gamma_n(\mathbf{a}) + v_{n,m}(\mathbf{a}) \tag{7}$$

where m′ is the other non-zero element's row number in the n-th column. Therefore, we can store new $u_{n,m'}(\mathbf{a})$ into the address of $u_{n,m}(\mathbf{a})$, which not used anymore, then swap m and m′ between twice iteration, to reduce memory by half.

(3) A Posterior Probability Calculation and Decision

While performing variable node processing, a posterior probability of the n-th codeword is calculated:

$$\hat{\gamma}_n(\mathbf{a}) = \gamma_n(\mathbf{a}) + \sum_{t \in S_n} v_{n,t}(\mathbf{a}) \tag{8}$$

And codeword decision:

$$z_n = \arg\max_{\mathrm{a}} [\hat{\gamma}_n(\mathbf{a})] \tag{9}$$

So far a decoding iteration is completed. Repeat the iteration until success or exceeds the maximum number of iterations.

2.2 Equivalent Decoders

There are different definitions of a priori likelihood information in several equivalent implementations of the Min-Sum algorithm, and their VN processing equations Eq. (6) are also different [7, 9].

In binary channel, the n-th codeword x_n taken the value of a, The corresponding binary bits compose vector $\boldsymbol{\alpha}$, and the soft decision data output by the receiver form a vector $\mathbf{y}$. Suppose the components of $\mathbf{y}$ are statistically independent, binary channel modulation function is $\mathrm{B}(x) = (-1)^x$, and the variance of additive white Gaussian noise (AWGN) in the channel is σ^2. Initialization and variable node processing steps of two equivalent implementations are as follows:

2.2.1 Standard Min-Sum

Define the likelihood information as the negative logarithm of probability:

$$\begin{aligned}\gamma_n(\mathrm{a}) &= -\ln[\Pr(x_n = a|\mathbf{y})] = -\ln\left[\prod_{t=1}^{p}\Pr(\alpha_t|y_t)\right] \\ &= -m\ln\frac{1}{\sqrt{2\pi}\sigma} + \frac{1}{2\sigma^2}\sum_{t=1}^{p}[\mathrm{B}(\alpha_t) - y_t]^2\end{aligned} \tag{10}$$

Omitted the DC components in Eq. (10) where do not affect the cumulative and minimum operations, the likelihood information is initialized as:

$$\gamma_n(\mathrm{a}) = \frac{1}{2\sigma^2}\sum_{t=1}^{p}[\mathrm{B}(\alpha_t) - y_t]^2 \tag{11}$$

2.2.2 Log Likelihood Ratio Min-Sum

Specify a Galois field element â as the basis for likelihood information, define likelihood information as logarithmic likelihood ratio (LLR):

$$\begin{aligned}\gamma_n(\mathrm{a}) &= -\ln\frac{\Pr(x_n = a|\mathbf{y})}{\Pr(x_n = \hat{a}|\mathbf{y})} = \frac{1}{\sigma^2}\sum_{t=1}^{p} y_t[B(\alpha_t) - B(\hat{\alpha}_t)] \\ &= \frac{2}{\sigma^2}\sum_{t=1}^{p}|y_t|(\alpha_t \oplus \hat{\alpha}_t)\end{aligned} \tag{12}$$

where $\oplus$ represents binary XOR, â can be any element in the Galois field. Usually 0 or the most likely element in this codeword is chosen as â. Correspondingly,

the normalization operation should also be introduced in the VN processing, replace Eq. (6) with following two equations:

$$u'_{n,m}(\mathbf{a}) = \gamma_n(\mathbf{a}) + \sum_{\substack{t \in S_n \\ t \neq m}} v_{n,t}(\mathbf{a}) \tag{13}$$

$$u_{n,m}(\mathbf{a}) = u'_{n,m}(\mathbf{a}) - u'_{n,m}(\hat{\mathrm{a}}) \tag{14}$$

If the â is the most likely element, as the direct decision of channel data, then Eq. (14) is replaced by:

$$u_{n,m}(\mathbf{a}) = u_{n,m}(\mathbf{a}) - \min_{a \in \{\mathbf{a}\}} \left[u'_{n,m}(a) \right] \tag{15}$$

At this time, all likelihoods information is always nonnegative and the LLR of the most probable element is fixed to zero.

2.3 *Extended Min-Sum (EMS)*

The huge amount of data in row processing steps restricts NB-LDPC applications. For each non-zero element in parity-check matrix, the decoder row processing unit needs to perform Galois field multiplication and division each of q (d_c − 1) times, calculates q^{d_c-1} summation results, and selects the minimum values for q configuration sets. Here, d_c is the weight of the row where the element is located, and q is for GF(q).

In Extended Min-Sum (EMS) algorithm, the amount of computation is reduced form q to n_m by sorting and truncating the likelihood information. If a truncated element needs to be referenced in the decoding process, a sum of a constant offset and the maximum likelihood in the element will be taken as the truncated element's likelihood information, indicating that its probability is lower than all the reserved elements.

3 APP-Based EMS Algorithm

Comparing Eqs. (6) and (8), the difference between them is only whether the accumulation process contains the likelihood ratio message of the current edge itself.

Equation (6) calculates the VN to CN message, which defines the possibility of current variable match with each Galois field element, under the conditions of other variables in the parity check equation are specified, and the check equation is

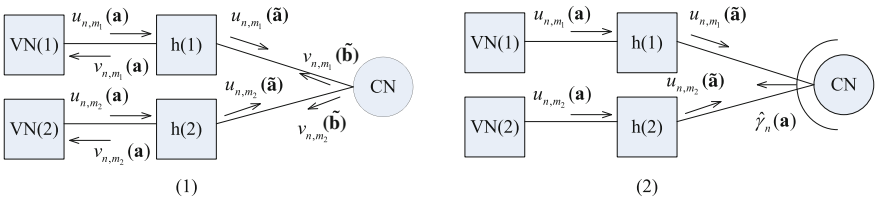


Fig. 1 Min-Sum (**1**) and APP (**2**) decoder check node processing structures

satisfied. Equation (8) obtains the most probable value of the variable after this iteration, which includes not only the information calculated in Eq. (6), but also the variable node prior information.

If we substitute the result of Eq. (8) for the Eq. (6), we get the APP-based EMS algorithm. At this point, Eq. (13) is replaced by:

$$u'_{n,m}(\mathbf{a}) = \hat{\gamma}_n(\mathbf{a}) = \gamma_n(\mathbf{a}) + \sum_{t \in S_n} v_{n,t}(\mathbf{a}) \tag{16}$$

Column operations are reduced to a column-by-column accumulation, which can be decomposed and assigned in the row processing, saves a lot of memory and time. But the approximation also introduces additional performance degradation (Fig. 1).

In EMS, VN to CN and CN to VN data each stores n_z sets of truncated LLR messages, where n_z is the number of non-zero elements in the parity check matrix. In the APP-based EMS algorithm, the data needs to be stored is APP likelihood ratio, and its accumulating cache. The total number of truncated LLR messages need to be stored is reduced from $2n_z$ to $2N$.

Following figures show the structure of the MS/EMS decoder and APP decoder. They have exactly the same initial likelihood ratio memory (Initial LLRs) and row processing module, but the CN to VN data memory and column processing module of the EMS decoder are replaced by a posterior likelihood ratio accumulator in the APP decoder. The channel data enters the decoder from the initial LLRs memory. During iteration, the data flows counterclockwise between the modules, until the decoding succeeds or reaches the maximum number of iterations (Fig. 2).

Scalar multiplication and offset can compensate the LLR overestimate, which caused by the Min-Sum algorithm itself, and improve the decoding performance [10]. These techniques can also be applied to EMS and APP-based EMS algorithm.

4 Performance Simulation

The three NB-LDPC codes used by Beidou system are simulated by Monte-Carlo method, and the decoding performance curves under several algorithms are respectively obtained.

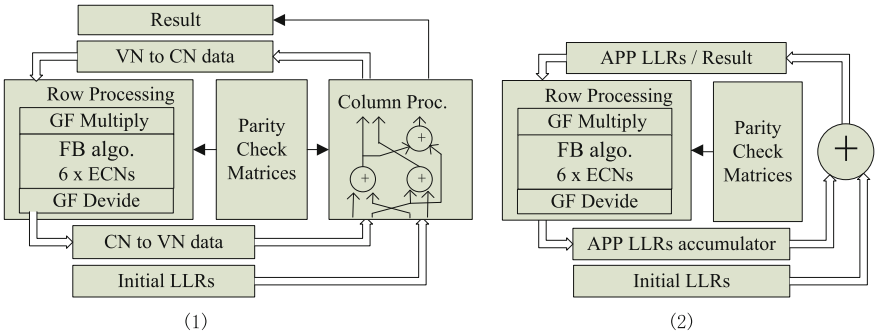


Fig. 2 Min-sum (**1**) and APP (**2**) LDPC decoder structures

The channel model used in simulation is additive white Gaussian noise (AWGN). The soft decision data is quantized to 4 bits, the length of information truncation $n_m = 10$ and the maximum number of iterations is limited to 100 (Fig. 3).

Concisely, only the curves of LDPC used by B1C subframe II ($N = 200$) are shown in figure, and the other two codes each has $\sim$0.5 dB performance degradation due to the shorter codeword length N.

Min-Sum (MS) and Offset Min-Sum (Offset-MS) have the best performance, Extended Min-Sum (EMS) have acceptable complexity and little performance loss. The two curves of the APP algorithm are the decoding performance of decoder parameter $n_{LLR} = 10$ and 64, when $n_m = 10$. Where n_m is the number of GF domain

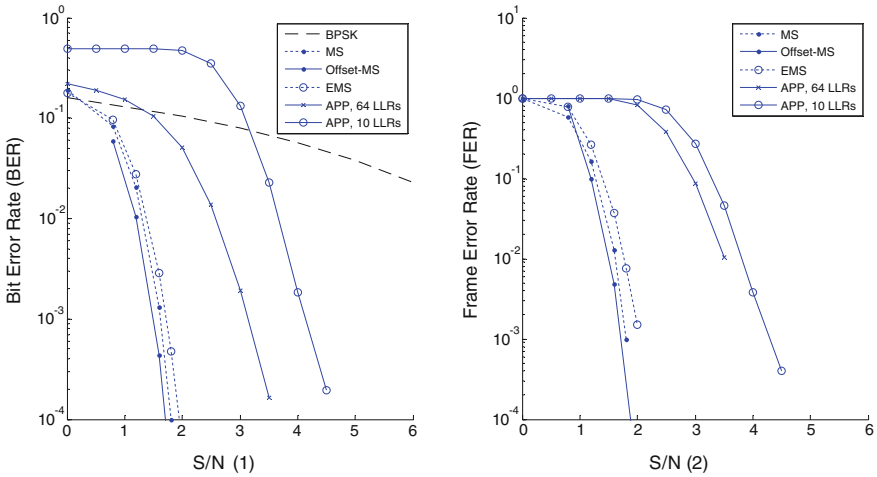


Fig. 3 Decoding performance, including the bit error rate (BER) (**1**) and frame error rate (FER) (**2**), to the signal noise ratio. The BER performance of uncoded BPSK modulation is also plotted as a reference, while the uncoded FER is always close to unity

elements participating in row processing operations, and n_{LLR} is the number of GF domain elements storing in a posteriori likelihood ratio accumulator.

5 Conclusion

Using SPA, or offset MS algorithms, the BDS NB-LDPC codes can obtain the performance approaching the Shannon limit. The complexity and performance needs to be balanced in a receiver, and EMS algorithm provides a flexible solution: choosing a suitable data truncation length based on resource and performance requirements.

APP-based EMS algorithm simplifies the column processing, accomplishes column-by- column accumulation at the same time as row processing, and eliminates VN to CN and CN to VN data memory. It requires less memory, and is easier to implement. However, it has decoding performance degradation about 2 dB. So it is suitable for consumer receivers, soft decoders and other applications with strict resource limitation.

Appendix: Operations Over Galois Field

In the decoding algorithm, using unsigned p bits represent the p binary coefficients of the polynomial elements in GF(2^p). The addition and subtraction is equivalent to the bitwise XOR of the polynomial coefficients. And there are two alternative implementations of multiplication and division: Logarithmic and shift register method.

Logarithmic method requires four steps, with fixed time-consuming:

- Obtain the logarithms of two operands by a look-up table;
- Add or subtract two logarithms over Integral Ring;
- Modulo $2^p - 1$;
- Complete the exponential operation by another look-up table, and get the result of multiplication or division.

The shift register method does not require look-up tables, but needs more and unfixed time. BDS using Galois field GF(64) based on the primitive polynomial p $(x) = 1 + x + x^6$, the multiplier/divider consists of two 6-bit shift registers and the primitive polynomial decides its feedback, as shown in following figure.

In a multiplication operation, the factor a is loaded into the shift register A and the number 1 is loaded into the other shift register B. Then the two LSBs are XORed with the MSB in each shift cycle. Until the value in the shift register B equals the other factor b, the result in the shift register A is the product ab.

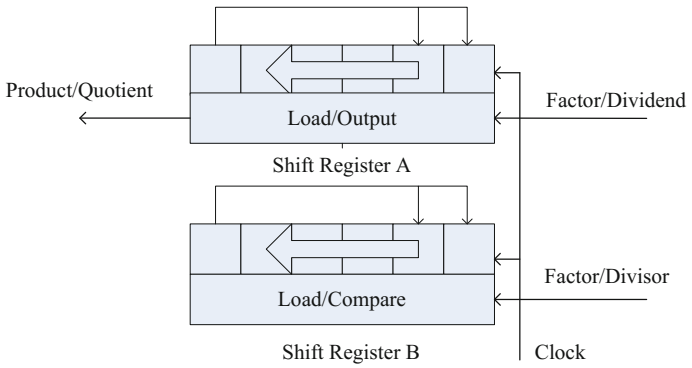


Fig. 4 shows the shift register structure of the multiplier/divider over GF(64) with a primitive polynomial $p(x) = 1 + x + x^6$

In a division operation, the dividend a is loaded into the shift register A while the divisor b is loaded into the shift register B. Perform the same shift operation as in multiplication, until the value in the shift register B equals 1. Then the result in shift register A is the quotient a/b (Fig. 4).

References

1. Abassi O, Conde-Canencia L, Ghouwayel AA, Boutillon E (2016) A novel architecture for elementary check node processing in non-binary LDPC decoders. IEEE Trans Circuits Syst II Express Briefs 64(2):136–140
2. Boutillon E, Conde-Canencia L, Ghouwayel AA (2013) Design of a GF(64)-LDPC decoder based on the EMS algorithm. IEEE Trans Circuits Syst. Part I: Regular Papers. 60 (10):2644–2656
3. Chen J, Fossorier MPC (2001) Decoding low-density parity check codes with normalized APP-based algorithm. IEEE Glob Telecommun Conf 2:1026–1030
4. China Satellite Navigation Office (2017) BeiDou Navigation Satellite System signal in space interface control document: open service signal B1C and B2a (Test Ed.)
5. Declercq D, Fossorier M (2007) Decoding algorithms for nonbinary LDPC codes over GF(q). IEEE Trans Commun 55(4):633–643
6. Fossorier MPC (1999) Reduced complexity iterative decoding of low-density parity check codes based on belief propagation. IEEE Trans Commun 47(5):673–680
7. Ghouwayel AA, Boutillon E (2011) A systolic LLR generation architecture for non-binary LDPC decoders. IEEE Commun Lett 15(8):851–853
8. Lin S, Costello DJ (2004) Error control coding, 2nd edn. Prentice-Hall, Inc.
9. Savin V (2008) Min-max decoding for non-binary LDPC codes. IEEE Int Symp Inf Theory 960–964:2008
10. Zhang X (2015) VLSI architectures for modern error-correcting codes. CRC Press

An Improved RAIM Algorithm and Hatch-Type Filter Smoothing Strategy

Si-long Luo, Li Wang, Rui Tu, Ya-bing Zhang and Wei-qi Zhang

Abstract An improved RAIM algorithm is proposed against the inaccurate of small outlier detection and identification, which based on the standard deviation of the unit weight in the positioning solution and in the form of an iterative selection of satellites. At the same time, the RAIM algorithm is applied to the positioning data preprocessing stage while the cycle slips only detect without repair. It solves the problem of the failure of the Hatch-type smoothing during the initial period, because of satellite fault or cycle slips. The results show that the new data processing method combined with Hatch-type filtering and RAIM algorithm can not only solve the divergence problem of Hatch-type filtering, but also can detect and identify 20 m small outlier satellites in different situations, which improve the integrity and positioning accuracy.

Keywords SPP · Hatch-type filter · RAIM · Small fault · Missing detection

S. Luo (✉)
School of Geological Engineering and Geomatics,
Chang'an University, Xi'an 710054, Shaanxi, China
e-mail: luosilong_xian@126.com

L. Wang
National Time Service Center, Chinese Academy of Sciences,
Xi'an 710600, Shaanxi, China
e-mail: wangli@chd.edu.cn

R. Tu
State Key Laboratory of Geographic Information Engineering,
Xi'an 710054, Shaanxi, China

Y. Zhang · W. Zhang
National Administration of Surveying, Mapping and Geo-Information
Engineering Research Center of Geographic National Conditions Monitoring,
Xi'an 710054, Shaanxi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_33

1 Introduction

Single Point Position (SPP) accuracy based on the double-frequency ionosphere-free model is generally about 10 m [1] and is mainly used for non-precision positioning and navigation. The precision of non-differential precision single-point positioning using high-precision carrier phase observation is higher, but its convergence time is longer and needs to deal with carrier ambiguity and cycle slips with more complex process. Carrier phase smoothing pseudo-range is a method which weighing the high-precision carrier phase observations and pseudo-ranges to improve the accuracy of single-point positioning [2–5]. The commonly used smoothing methods are Hatch filtering [3], elevation-angle-based Hatch [4], weighted smoothness of observation [5] and so on. This kind of filtering needs the accurate initial value of filtering to ensure the reliability of the subsequent pseudo-range of the satellite. Therefore, it is necessary to detect and identify the satellites in the epoch.

Receiver Autonomous Integrity Monitoring (RAIM) refers to the monitoring of the integrity of the localization solution using only the redundant observations of user receivers, and a large number of studies have been conducted [6–9]. However, the traditional RAIM algorithm is insensitive to minor faults of about 20 m or smaller, which makes it hard to detect and identify completely, resulting in missing detection and decrease the positioning accuracy.

To solve the defect that Hatch Filter failure because satellite fault in initialization phase or cycle slips in satellite, as well as the problem that RAIM algorithm is not sensitive to the small fault, this paper proposes a data processing method which combines Hatch filtering and RAIM algorithm to simulate the original data two kinds of fault situations, and use this method for experimental analysis. The results show that SPP after the combination of Hatch filtering and RAIM algorithm, which not only can detect and identify the fault of 20 m fault satellites, but also can properly correct the pseudo-range deviation or rejection of the faulty satellites, greatly reducing the missed detection rate to improve the integrity and positioning accuracy.

2 Algorithm Principle

2.1 Hatch-Type Filter

Except the ionospheric delay, multipath effect and observed noise term, the other types of error terms are the same in GPS pseudo-range observation equation and carrier phase observation equation. Thus the two types of observation equations without considering the same error can be expressed as [1]:

$$P^j = R^j + c\left(t_R - t_s^j\right) + D_{ion}^j + D_{mp}^j + \varepsilon_p \tag{1}$$

$$\Phi^j = \lambda\left(\varphi^j + N^j\right) = R^j + c\left(t_R - t_s^j\right) + D_{ion}^j + D_{mp}^j + \varepsilon_\varphi \tag{2}$$

In the above formula, P^j is the pseudorange observation of the jth satellite; R^j is the geometric distance from the satellite to the receiver; t_R, t_S is the receiver clock error and the satellite clock error; D_{ion}^j is the ionospheric delay; $D_{mp}^j, D_{m\varphi}^j$ for the corresponding pseudorange and phase multipath delay; Φ^j for the phase pseudorange; λ^j for the carrier wavelength; φ^j for the corresponding phase observation; N^j is integer ambiguity; ε is observed noise;

Hatch filtering weighted-smoothing formula [3] can be obtained by making the difference between (1) and (2) in epoch t_k and $t_{k\text{-}1}$.

$$\overline{P_{LC}^j}(1) = P_{LC}^j(1) \tag{3}$$

$$\overline{P_{LC}^j}(i) = \frac{1}{i}P_{LC}^j(i) + \left(1 - \frac{1}{i}\right)\left[(i-1) + \Delta\Phi_{LC}^j(i, i-1)\right]$$

Hatch filtering is simple and easy. When i = 1, pseudoranges have the heaviest weight, and then reduce the weights of the pseudoranges in smoothing time to make full use of high-precision carrier phase observations. However, it is also found that the Hatch filter relies heavily on the correctness of the smoothed initial epoch satellite pseudorange. If the initial epoch satellite is fault, that is, the value of P_{LC}^j is abnormal, the satellite will transmit the fault offset in subsequent epochs, so that the smoothed value deviates from the actual value even diverges. Based on this, this paper combines RAIM algorithm before ephemeris smoothing to detect and identify faults.

2.2 RAIM Algorithm

Adopting the RAIM algorithm based on the least squares method [9], the GPS pseudorange observation equation is given as:

$$\mathrm{y} = \mathrm{Gx} + \varepsilon \tag{4}$$

where y is the pseudorange observation value of n available satellites, G is the observation matrix of dimension $n \times 4$, formed by the geometry vectors of n available satellites and a column of clock error vectors; x is the vector of navigation solutions $[dx\, dy\, dz\, dt]^{\mathrm{T}}$; ε is the n-dimensional observation of the pseudo-range noise vector, if there is deviation, then represent as ε + b.

The unit weight's standard deviation of the pseudorange residual vector obtained by least-squares principle is:

$$\sigma = \sqrt{V^T PV/(n-4)} = \sqrt{SSE/(n-4)} \tag{5}$$

Wherein, V is the pseudorange residual vector, P is the observation of the weight matrix. Assuming that the pseudo-range residual limit is σ_T under normal conditions, $\sigma > \sigma_T$ indicates that a satellite fault can be detected and the user can be alerted. Afterwards, the fault satellite can be identified using the Baldar data detection method [12], which constructs the statistics based on the least square residual vector:

$$d_i = \frac{|v_i|}{\sigma_0 \sqrt{Q_{vii}}} \tag{6}$$

According to the statistical theory, d_i obeys the normal distribution whose mean value is zero when there is no fault. The threshold T_d is calculated according to the parameters of the approach phase, and the statistic d_i of each satellite is compared with T_d respectively. If $d_i > T_d$, it indicates that the ith satellite is faulty. In addition, the RAIM algorithm must be given its integrity assurance. We uses Horizontal Protection Level (HPL) value as RAIM usability standards. Before the RAIM algorithm is performed, the Horizontal Alert Limit (HAL) is calculated from the non-precision phase parameters of the aircraft, and the HPL is compared with the HAL. If HPL $< HAL$, the RAIM algorithm is available, otherwise unavailable. The fault detection and identification are based on the selection of stars based on the change of error in the unit weights.

3 Hatch and RAIM Combined Strategy

The SPP data processing method with Hatch filtering and RAIM algorithm combination shown in Fig. 1.

As can be seen from Fig. 1, the data processing with RAIM and Hatch combination has the following characteristics:

(1) *Pseudo-range smoothing is performed after SPP and detecting the fault first.*

Considering the importance of Hatch filtering in fault detection and recognition at the beginning of smoothing, the RAIM algorithm and the star selection iterative method based on the change of error in the unit weights are given priority to detect and identify the epochs. Then the epochs are smoothed by Hatch filtering Ensure the smoothness of the correctness.

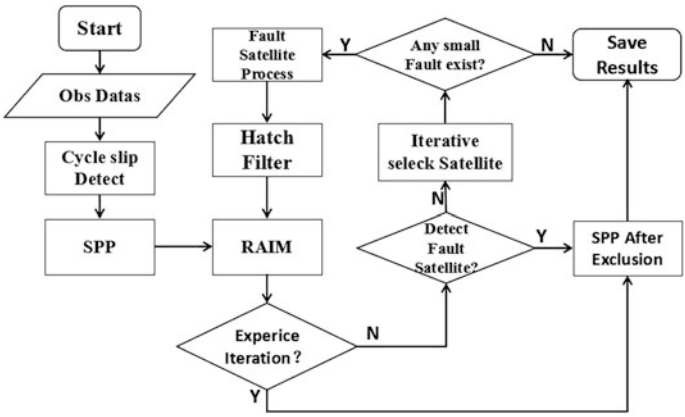


Fig. 1 Hatch-type filter and RAIM algorithm combination processing flow

(2) *Detection of small fault using satellite selection iterations based on changes in the unit weights.*

The satellite selection iterative method based on the change of error in unit weight as a supplement to RAIM. If the fault value is larger, the faulty satellite can be successfully identified by the RAIM algorithm; otherwise, based on the variation of error in the epoch unit weight, the jth visible satellite is sequentially eliminated from the n available satellites, and the remaining n−1 visible satellite to locate and solve the value of σ_j so as to obtain n $\Delta\sigma_j$ value. This paper draws to take $\Delta\sigma_T =$ 5m as the recognition threshold. If $\Delta\sigma_j > \Delta\sigma_T$, it indicates that the j-th satellite has poor positioning performance and is regarded as a fault identifier.

(3) *Fault satellite processing.*

For the faulty satellite detected and identified, if the satellite is normally involved in Hatch filtering in the previous epoch, then the smoothing of the epoch is continued using the smoothing pseudo-range value of the previous epoch; if the satellite does not exist in the previous epoch or the previous epoch is a fault satellite, then be removed. In addition, the satellite is also removed in the event of a cycle slip. For the current case of redundant observation of GPS, it is feasible to eliminate some satellites with poor positioning performance to improve the integrity and positioning accuracy.

4 Experimental Analysis

4.1 Experimental Data and Schemes

GPS observation data of the first day of 2017 of the IGS BJFS station with a sampling interval of 30 s are used, and all the epochs between 3:00 and 11:00 are selected for the example analysis.

Two different satellite fault scenarios were simulated. *Case 1*: Satellite simulation of random failures, that is, an interval of 10 epochs is added with a deviation of 20 m on the pseudo-range observations of the 3rd satellite of each observation epoch. Since the faulty satellite is uncertain and non-existent, so the situation is random and sudden. *Case 2*: Simulation of medium and long term sustained failure of satellites, that is, a specific numbered satellite (PRN14) is selected as a fault satellite, and 20 m deviation is added to its pseudo range observation, the situation is specific and continuous.

Figure 2 shows the availability of the RAIM algorithm. During the experimental period, the number of available satellites for the RAIM algorithm is between 6 and 12, and its availability is 100% (HPL < HAL). Since the RAIM usability standard HPL is only related to non-precision approach phase parameters and satellites geometry locations, the HPL values are equal in two cases.

The following solutions are designed for the two kinds of fault cases, two aspects of positioning accuracy and RAIM performance are analyzed.

- Direct Pseudo Range Point Positioning (SPP)
- RAIM Algorithm Processing (RAIM)
- Hatch filter processing (Hatch)
- RAIM algorithm processing after Hatch filtering processing (H_R)
- Hatch/RAIM combined data processing (HR).

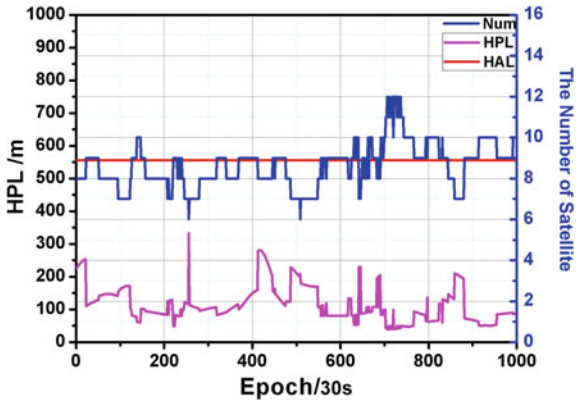


Fig. 2 RAIM algorithm availability graph

4.2 Experiment 1: Random Satellite Burst Failure Situation

Figure 3 shows the epoch sigma sequence processed by different schemes for 20 m deviation on the pseudo-range. It can be seen from the figure, under the pseudo-range offset of 20 m, there will be a large number of sigma < sigmaT of the epoch when only use the RAIM algorithm, resulting in missing detection. If only the Hacth filtering is used to smooth the pseudo-range, It only detects the cycle slip and does not repair it, so there are still some unsmooth satellite positioning solvers between 300–350 and 620–750 epoch, resulting in a large sigma value; if the RAIM algorithm is performed after the Hatch filtering, the effect is obviously better than the first two, in fact, it can be seen as a combination of two independent repair processes, but there are still some epoch that has not improved; When the combination of Hatch/RAIM is used, the defect of scheme four is made up. and accuracy is improved by eliminating satellites with poor positioning performance. It should be pointed out that the height angles of PRN12, PRN13 and PRN14 are not enough to reach the elevation angle of 10° at the pseudoranges of 140–221, 350–510 and 650–690 epochs and thus are not reflected in sigma sequence of the figure.

Figure 4 shows the residual sequences and RMS plots of E, N, and U axes under different schemes with the bjfs station's high-precision coordinates as true values. It can be seen from the figure that the residuals of E, N and U after Hatch/RAIM combination processing are the smallest, but their RMS values are not as good as the simple RAIM algorithm and RAIM algorithm processing after Hatch filtering processing in the U direction. The reason is that the false distance repair value of the fault satellite is less than its pseudo range deviation value, so the positioning effect is still not as good as directly removing the faulty satellite.

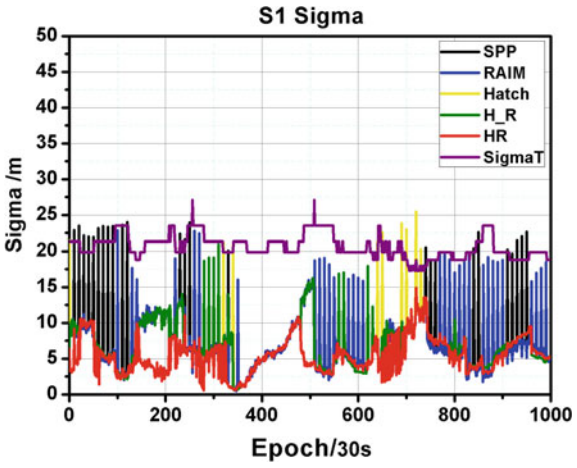


Fig. 3 Sigma values for different scenarios under fault **situation** 1

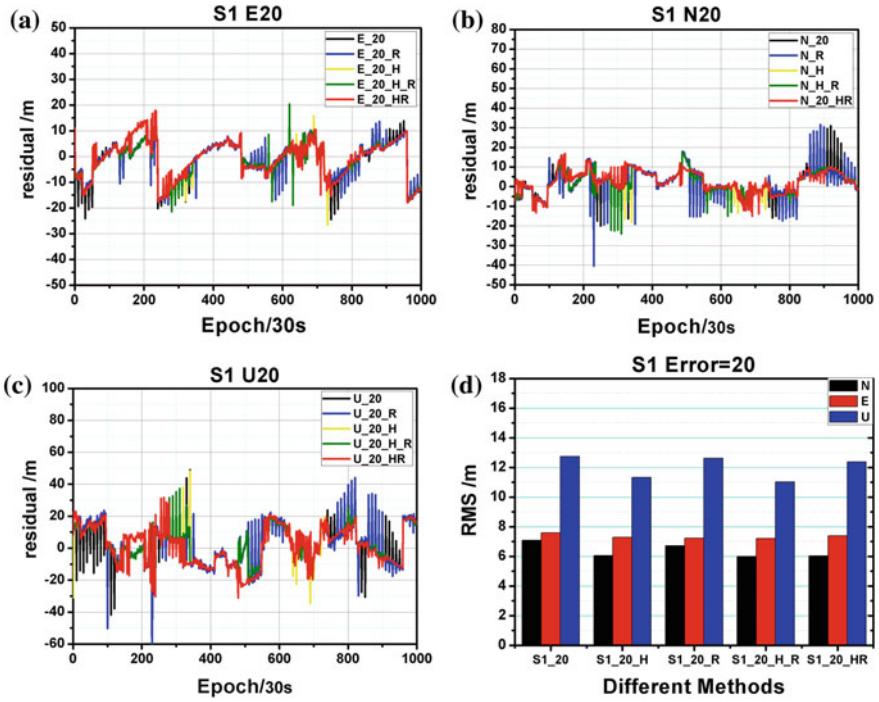


Fig. 4 E/N/U direction residual and various solutions RMS accuracy under fault situation 1

4.3 Experiment 2: Specific Satellites Continued Fault Situation

Figure 5 shows epoch sigma value sequence after adding 20 m pseudo-range deviations to PRN14 under the fault of case 2. It can be seen from the figure, only the case 3, which is Hatch filtering processing, clearly reflects the defects of distortion in the epoch of satellite fault under the initial Hatch smoothing,, so that RAIM algorithm processing after Hatch filtering processing results are similar to the RAIM algorithm; And when using Hatch/RIAM combination of programs to deal with, glitches play a better detection, thus the Sigma value is improved. Therefore, with enough redundant observations to be located, it is entirely possible to eliminate glitch satellites or poorly positioned satellites to improve accuracy and integrity.

Figure 6 shows the residual sequences and RMS plots in E, N and U directions after processing of different schemes. It can be seen from the figure that for the fault case 2, the residual E, N and U residuals obtained after the satellite is removed are smaller than other schemes because the faulty satellite identified by the Hatch/RAIM detection can not perform pseudo-range smoothing repair. At the same time, it can be seen from the RMS plots of each scheme that the positioning scheme combined with Hatch/RAIM also has higher positioning accuracy than other schemes.

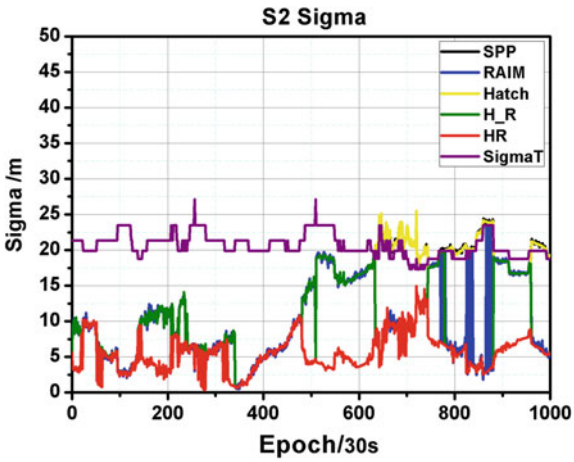


Fig. 5 Sigma values for different scenarios in case of fault **situation** 2

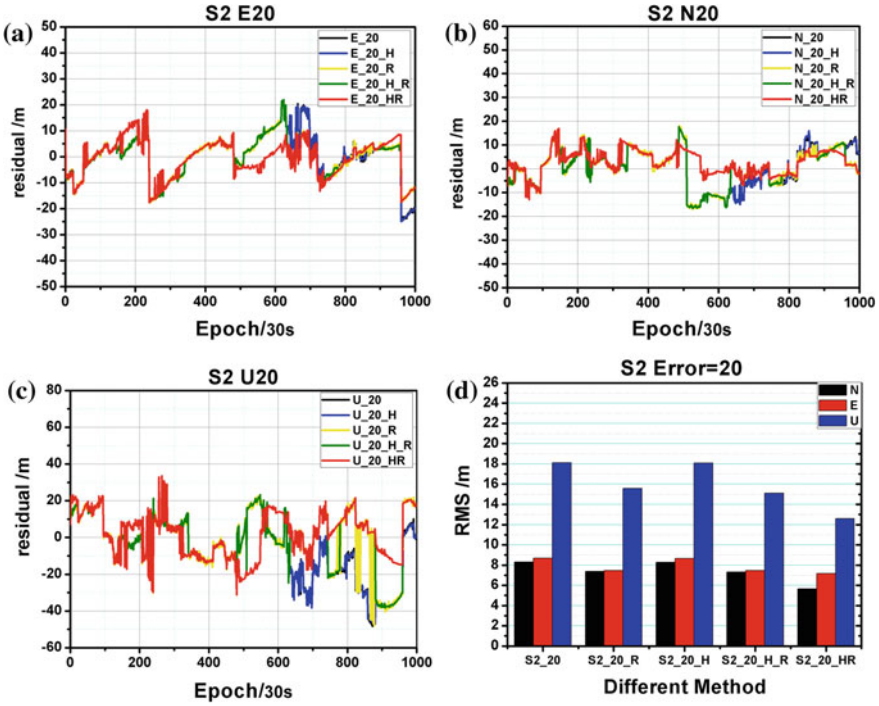


Fig. 6 E/N/U direction residual and various solutions RMS accuracy under fault situation 2

Table 1 RAIM performance results under different schemes

Schemes	Detection number	Identification number	Detection rate (*100%)	Identification rate (*100%)
R_0_R	0	0	0	0
R_0_H_R	0	0	0	0
R_0_HR	194	194	194/0	194/0
R1_20_R	34	34	34/75	33/34
R1_20_H_R	9	9	9/75	9/9
R1_20_HR	218	218	218/75	75/218
R2_20_R	241	241	241/490	241/241
R2_20_H_R	252	251	252/490	251/252
R2_20_HR	658	658	658/490	490/658

Table 1 shows the RAIM results of experimental data processed by different schemes in two fault situations. It can be seen from the table, after the Hatch/RAIM fusion process for the original data, 194 poorly positioned satellites are still removed from the epoch.

For fault *Scheme 1*, 34 fault epochs can be detected by using RAIM algorithm, and the fault detection rate is 45.3%. After Hatch smoothing, the RAIM algorithm is used to reduce the fault detection rate to 12% by restoring the pseudo-range bias of some satellites. After the combination of Hatch/RAIM, the detection rate can reach 290.7%, which shows that this algorithm also performs the star removal for the epoch that has not artificially added.

For fault *Scheme 2*, only the fault detection rate of RAIM algorithm is 49.2%, and a considerable number of faulty satellites fail to identify. However, the RAIM algorithm fault detection rate is 51.4% after performing Hatch filtering smoothing for each epoch, indicating that for a specific in the case of satellite failure, Hatch filtering not only does not play the role of pseudo-range smoothing, but also increases the pseudo-range discrepancy of the faulty satellite in the later epoch for the smooth. When the combination of Hatch/RAIM is adopted, the Hatch filtering can successfully detect and identify 658 failed satellites, detection rate of 134%.

5 Conclusion

The proposed Hacth/RAIM positioning data processing method can effectively solve the initial defects of Hatch filtering under different fault conditions. In view of the shortcoming that the miniaturization of RAIM algorithm is hard to identify, this paper adopts the star selection iterative method based on mean square error of unit weight to perform small Fault detection as a supplement, you can fully identify the deviation of 20 m, reduce the risk of missed inspection to improve the positioning accuracy and integrity.

With the rapid development of GNSS system, GNSS Multi-System positioning is more practical and the probability of multiple faults will also increase. In this paper, we only study the single-fault location of GPS. The next step will focus on the data processing and satellite optimization of multi-satellite faults. Improving the existing algorithms at the same time so that the algorithm can adapt to real-time location and integrity monitoring in a variety of situations

Acknowledgements This research was partially supported by The National 973 Program of China, No. 2014CB744700; the Program of the National Natural Science Foundation of China (41731066, 41774025, 41604001, 41504006, 41304033), the Grand Projects of the Beidou-2 System (GFZX0301040308) and the Special Fund for Basic Scientific Research of Central Colleges (310826172006, 310826172202, 310826173101, Chang'An University). Heartfelt thanks to the editors and anonymous referee experts provide valuable opinions and suggestions for this article!

References

1. Li ZH, Huang JS (2010) GPS measurements and data processing. Wuhan University Press
2. Zhang CJ, Yang L, Chen J (2009) Study on the algorithm for improving the accuracy of GPS carrier phase smoothing pseudorange. J Geodes Geodyn 29(4):106–110
3. Hatch R (1982) The synergism of GPS code and carrier measurements 1:1213–1231
4. Fan SJ, Guo JM, Peng XY (2005) Study on accuracy of GPS dual-frequency phase smoothing pseudo-range and its single point positioning. Surv Mapp Eng 14(4):39–42
5. Xu JY, Li JL, He HB et al (2014) Beidou code differential positioning based on weighted measurements carrier phase smoothing. J Geodes Geodyn 34(4):123–126
6. PLA Information Engineering University (2015) Research and application of new RAIM algorithm. Information Engineering University of Chinese People's Liberation Army
7. Hewitson S, Wang J (2006) GNSS receiver autonomous integrity monitoring (RAIM) performance analysis. GPS Solut 10(3):155–170
8. Brown RG, Mcburney PW (1988) Self-contained GPS integrity check using maximum solution separation. Navigation 35(1):41–53
9. Parkinson BW, Axelrad P (1988) Autonomous GPS integrity monitoring using the pseudorange residual. Navigation 35(2):255–274

The Performance Assessment on Lunar Navigation for GNSS

Bo Qu, Tao Yan, Xingyuan Han, Yanguang Wang, Longlong Li and Yansong Meng

Abstract With the gradual implementation of China's lunar exploration project, the ground monitoring and control network also has to complete a large number of lunar monitoring and control tasks. In order to alleviate this pressure, lunar spacecraft needs to employ other methods to finish its positioning and timing tasks. Global navigation satellite system is designed to provide positioning and timing service for terrestrial users. Because of the improvement on GNSS receiver, it has been used on LEO satellites and spacecraft. However, the signal power of navigation satellites in GNSS is weak, and GDOP of visible satellites is bad, which is a big challenge for the design of GNSS receiver on lunar orbit. In this paper, the lunar navigation performance of GPS and BDS III is evaluated. We analyzed the propagation attenuation loss of GPS and BDS III signals on the moon, and gave the propagation attenuation of these signals at different frequencies. In addition, we investigated the power variation characteristic, the number variation, visible time, and GDOP of navigation satellites in GPS and BDS, which can improve the design of GNSS receiver for lunar mission.

Keywords GNSS · Lunar · Receiver · Visible satellite · GDOP

1 Introduction

Recent years, the demand for navigation and positioning on the moon increases gradually with the further implementation of lunar space exploration activities in China. In order to improve the autonomy, reliability and safety of lunar space vehicle, it is necessary to investigate new navigation and positioning methods as backup means [1].

Global navigation satellite system (GNSS) was initially designed for terrestrial users to provide location and timing services. However, with the study on spatial

B. Qu (✉) · T. Yan · X. Han · Y. Wang · L. Li · Y. Meng
China Academy of Space Technology(Xi'an), Xi'an, China
e-mail: qb840528@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_34

characteristics of global positioning system (GPS) and the performance improvement of GNSS receiver, more and more spacecraft and LEO satellites are equipped with GNSS receivers to realize autonomous navigation and positioning.

However, there is a big challenge for GNSS receiver to use in lunar space mission. Because the altitude of lunar spacecraft is higher than that of GNSS satellites whose antenna points to the earth, lunar spacecraft can only receive GNSS signals from the other side of the earth. The main lobe signal from the antenna on the GNSS satellite is blocked by the earth, which decreases the beam angle of the signal from a visible satellite. The number of visible GNSS satellites is smaller, which degrades the geometric dilution of precision (GDOP). In order to improve the GDOP and increase the number of visible satellites, the GNSS receiver has to receive the side lobe signal of the antenna on navigation satellites [2]. However, the side lobe signal power of GPS satellites is 20 dB less than that of main lobe signal [3]. The receiver requires the higher sensitivity to acquire and track the side lobe signal. In addition, the distance between the spacecraft on lunar orbit and GNSS satellites is far greater than the distance between the ground receiver and GNSS satellites, so the power attenuation of the navigation signal received in lunar mission is larger than that of the navigation signal received by terrestrial users, which results in the requirement on the high sensitivity of GNSS receiver.

In recent years, there have been some literature on the above problems. Reference [2] designed a high orbit GPS receiver, and simulated its performance in geostationary earth orbit (GEO) and lunar space mission. The performance of GPS antenna sidelobe signal on a high orbit satellite was analyzed and discussed in [3]. GPS navigation performance on the Earth-Moon transfer track is analyzed in [4]. Reference [5] designed a high sensitivity receiver for GPS L1CA signals for lunar missions. However, the above literature mainly analyzed the GPS navigation performance for GEO or lunar mission, and didn't analyze the lunar navigation performance of Beidou navigation satellite system (BDS).

The navigation and positioning performance of GPS and BDS in lunar mission has been analyzed in this paper. The variation characteristic of the navigation satellite power is simulated. Besides, we have investigated the number of visible satellites and the variation characteristics of GDOP under different receiver sensitivities, which supports the development and design of lunar GNSS receiver.

The paper is organized as follows: In the Sect. 2, GPS and BDS are briefly introduced. Section 3 describes the principle of GNSS signal reception in lunar missions and the signal propagation attenuation analysis. Section 4 shows the assumptions and conditions of the simulation, and analyzes the power variation, the number of visible satellites and the GDOP variation of GPS and BDS. Finally, conclusions are drawn in Sect. 5.

2 Introduction on GPS and BDS

GPS is a global navigation satellite system developed by the United States. Its core constellation consists of 24 medium earth orbit (MEO) satellites in order to ensure that the number of visible satellites is more than 4 at anywhere and any time. At present, GPS has about 31 satellites whose orbit height is about 20,000 km. These satellites are distributed on 6 orbital planes, and there are about 4–6 satellites on each orbital plane. The angle between the orbital plane and Earth equator is 55°, and the difference between the ascending node longitude and the adjacent orbital planes is 60°. The orbit period of GPS satellites is about 11 h and 58 min [6, 7].

GPS signals mainly include L1C/A signals, L2C signals, and L5 signals. All of GPS satellites are transmitting L1C/A signals, but only part of GPS satellites are transmitting L2C signals and L5 signals. The minimum power level of L1C/A signal is −158.5dBW, and that of L2C signal is −160 dBW [6, 8].

BDS is the Beidou satellite navigation system developed by China. Its construction and development are divided into three stages. The first stage is called as Beidou-I, the second stage is Beidou-II, and the third stage is Beidou-III. Beidou-III constellation consists of three geostationary earth orbit (GEO) satellites, three inclined geosynchronous orbit (IGSO) satellites, 24 MEO satellites and several backup satellites. The orbit height of GEO satellites is 35,786 km, and the longitude of three GEO satellites is 80°, 110.5° and 140° respectively. The orbit height of IGSO satellites is also 35,786 km, and the angle between the IGSO orbital plane and Earth equator is 55°. The orbit height of MEO satellites is 21,528 km, and the angle between the orbital plane of MEO satellites and Earth equator is also 55° [9]. The MEO constellation of BDS contains three orbital planes, and the phase of MEO satellites is selected from the Walker24/3/1 constellation. The IGSO satellites are distributed in three orbital planes, whose difference of the right ascension of ascending node (RAAN) is 120° [10].

At present, BDS signals mainly consists of B1I, B2I, B1C and B2a signals. The B1C and B2a signals are transmitted by MEO and IGSO satellites in BDS-III. If the elevation of the BDS satellites is about 5° and the BDS receiver with the 0dB right

Table 1 Minimum receiving power level of BDS signals

Signal	Satellite Type	Minimum power level (dBW)
B1I	GEO, MEO, IGSO	−163
B2I	GEO, MEO, IGSO	−163
B1C	MEO	−158.5
	IGSO	−160.3
B2a	MEO	−155.5
	IGSO	−157.3

circularly polarized antenna, the minimum power level of the navigation signal from the BDS satellites is shown in Table 1[9, 11].

It can be seen from Table 1 that the minimum power level of B1I and B2I signals is −163 dBW, which is 4.5 and 2.7 dB smaller than that of B1C signal transmitted by MEO and IGSO satellites, respectively. The minimum power level of B1I and B2I signals is 5.7 and 7.5 dB smaller than that of B2a signal transmitted by MEO and IGSO satellites, respectively. The B1C and B2a signals transmitted by IGSO satellites are 1.8 dB weaker than those transmitted by MEO satellites. The moon is far away from the earth which causes large propagation attenuation. B1C and B2a signals have higher power than B1I and B2I signals, which can reduce the sensitivity requirement on lunar receiver for signal acquisition and tracking.

Comparing the constellations of GPS and BDS, we can see that MEO satellites of GPS are distributed in 6 orbital planes, while MEO satellites of BDS is distributed in 3 orbital planes. In addition, BDS has 3 IGSO, which causes that BDS and GPS has different coverage characteristics and affects the visibility of these two navigation systems on the moon. From the point of view in signal power, the minimum power level of L1C/A signal is −158.5 dBW and higher than that of B1I and B2I signals, but L1C/A signal has the same minimum power level as B1C signal from MEO satellites and its minimum power level is 3 dB weaker than that of B2a signal. Therefore, in the situation of the same propagation loss, the power of B2a signals on the moon is higher than that of L1C/A signals and B1C signals.

3 Principle and Link Analysis of Signal Reception on Lunar Orbit

In this section, we introduce the principle of GNSS signal reception on lunar orbit and analyze the propagation loss of different navigation signals.

3.1 Principle of GNSS Signal Reception on Lunar Orbit

The distance between the lunar vehicle and the earth is much larger than the height of the GNSS satellites. The lunar vehicle needs to receive the GNSS satellite signals from the other side of the earth, as shown in Fig. 1. The main lobe signal of GNSS satellites is blocked by the earth, and the beam angle from visible satellites is small, which leads to the shorter visible time, the smaller number of visible satellites, and the worse GDOP in the moon space mission.

In order to increase the number of visible navigation satellites, it is necessary for the lunar receiver to receive the side lobe signal from the antenna on GNSS satellites. We can get the relative gain table of the side lobe gain relative to the gain of 15° antenna angle of GPS Block IIR satellites in Ref. [3]. When the

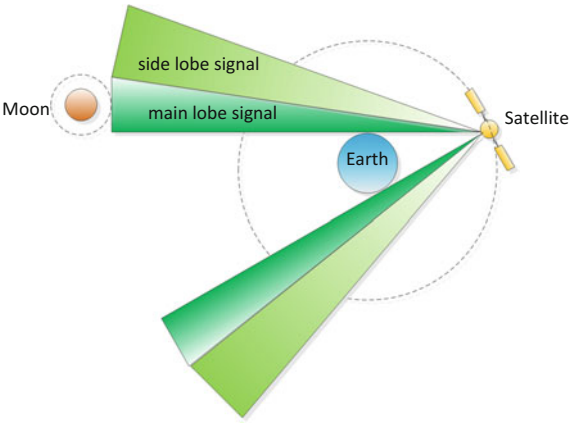


Fig. 1 Diagram of lunar GNSS signal reception

antenna angle is greater than 69°, the antenna gain relative to that of the 15° antenna angle is smaller than −20.61 dB. When the sensitivity of the lunar receiver is high enough to receive the signal from the 69° antenna angle, the beam angle of navigation satellites can be increased, which effectively increase the visible time of navigation satellites. However, the high sensitivity increases the difficulty in the design of a lunar receiver.

3.2 GNSS Link Analysis on Lunar Orbit

The distance between the lunar spacecraft and the earth is about 380,000 km, and the orbit height of GNSS satellites is about 20,000 km, so the signal transmission distance is assumed to be 400,000 km. We can calculate the free space propagation loss of the GPS and BDS signals at the lunar distance.

From the Table 2, we can see that the propagation loss of L1CA, B1I, and B1C signals is more than 208 dB. When the signal frequency is reduced to around 1200 MHz, the free space propagation loss of L2C and B2I signals is reduced to 206 dB. The propagation loss of L5 and B2a signals is further reduced to 205.89 dB. Because the signal received by lunar receiver is very weak, the decrease of the space propagation loss can decrease the sensitivity of lunar receiver, which can reduce the difficulty in the design of a lunar receiver.

Table 2 Propagation loss of GPS and BDS signals

System	Signal	Frequency (MHz)	Propagation loss (dB)
GPS	L1C/A	1575.42	−208.43
	L2C	1227.6	−206.26
	L5	1176.45	−205.89
BDS	B1I	1561.098	−208.35
	B2I	1207.140	−206.12
	B1C	1575.42	−208.43
	B2a	1176.45	−205.89

4 Simulation Results

4.1 Simulation Conditions and Assumptions

In order to compare the signal power and coverage characteristics of GPS and BDS constellations, L1C/A and B1C signals, which have the same minimum power level and frequency, are assumed to be received in simulation. Due to the lack of BDS antenna characteristic data, Both GPS and BDS adopt GPS BLOCK IIR antenna. Besides, the GNSS receiver antenna gain is set to be 11 dB.

We need the positions of GPS and BDS navigation satellites and spacecraft at different times to calculate the power, visible time, number of GPS and BDS visible satellites. In simulation, satellite tool kit (STK) is used to simulate GPS and BDS constellations to obtain the above information.

STK software has GPS constellation, which can be used directly. However, BDS constellation needs to be set up. The B1C signal is only transmitted by IGSO and MEO satellites, so BDS constellation didn't contain GEO satellites in the simulation [8]. The types and quantities of GPS and BDS satellites used in the simulation are shown in Table 3.

BDS satellites have not yet been fully deployed, and we can't obtain the specific orbit of all 24 BDS MEO satellites. According to the Ref. [10], BDS MEO constellation uses Walker24/3/1 in simulation. There are five IGSO satellites in Ref. [11], but the constellation of BDS III contains only three IGSO satellites [9]. We remove one satellite from the orbital planes having two IGSO satellites and another one from the orbital planes having three IGSO satellites so that the number of IGSO satellites decreases to three and meets the constellation design of BDS III.

Table 3 Type and quantity of navigation satellites in simulation

System	Satellite type	Quantity
GPS	MEO	31
BDS	IGSO	3
	MEO	24

4.2 Simulation Analysis

This section focuses on the power variation of GPS and BDS satellites, the visibility of navigation satellites, the number of visible satellites and the GDOP under different receiver sensitivities. Figures 2 and 3 show the power variations of GPS and BDS satellites on lunar orbit. We can see from Figs. 2 and 3 that the power of GPS L1C/A signals and BDS B1C signals at most of the time is in the region between −182 dBW and −190 dBW. The sensitivity of GNSS receiver needs to reach −190 dBW in order to receive most of the navigation satellite signals. Considering that the power of B2a signal is 3 dB higher than that of L1C/A signals and B1C signals and its path propagation loss is 2.54 dB lower than that of L1C/A signals and B1C signals, the sensitivity of GNSS receiver receiving B2a signals is 5.54 dB lower than that of GNSS receiver receiving L1CA signals and B1C signals.

Figures 4 and 5 show the number variation of visible satellites in GPS and BDS at different reception sensitivities. The simulation time is 24 h in these figures. When the number of visible satellites isn't less than 4 satellites, GNSS receiver can be located. It can be seen from Fig. 4 that the number of visible satellites in either GPS or BDS can't be more than four at any time when the receiver sensitivity is −185 dBW, which means that GNSS receiver can't get its position at any time. In order to improve the visibility of navigation satellites, the joint reception of GPS and BDS signals can be used to increase the number of visible satellites. It can be seen that the number of visible satellites can be no less than 4 at any time under the −185 dBW receiver sensitivity, after employing the joint reception.

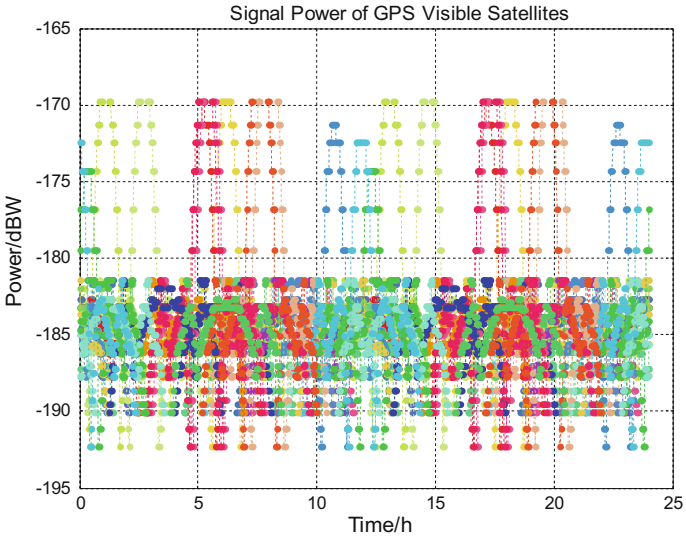


Fig. 2 Power of GPS satellites (Antenna gain included)

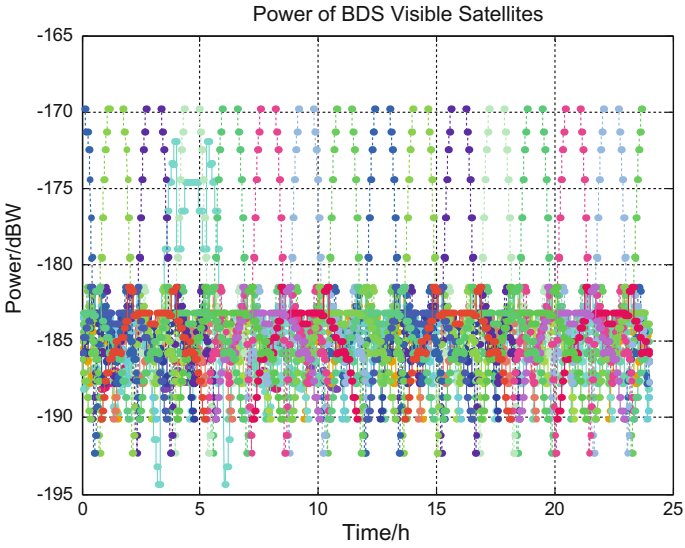


Fig. 3 Power of BDS satellites (Antenna gain included)

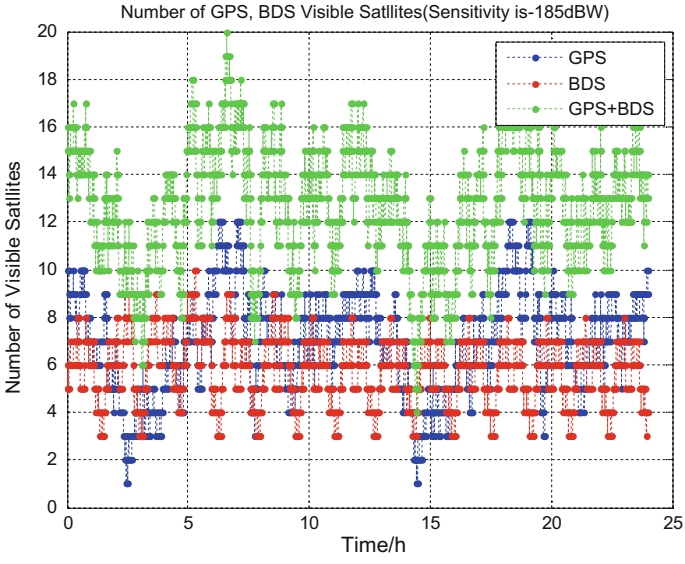


Fig. 4 Number of visible satellites of GPS, BDS, and GPS + BDS (receiver sensitivity is −185 dBW)

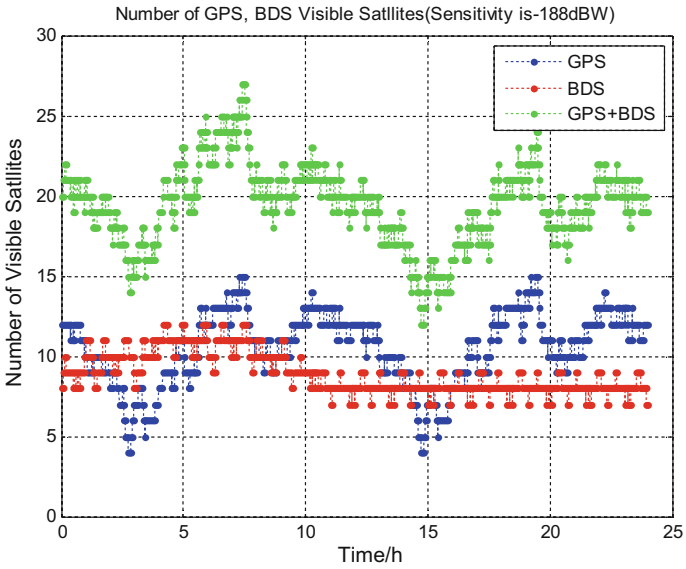


Fig. 5 Number of visible satellites of GPS, BDS, and GPS + BDS (receiver sensitivity is −188 dBW)

As shown in Fig. 5, either GPS or BDS can guarantee that no less than 4 satellites can be visible at any time, when the receiver sensitivity is increased to −188 dBW. The number of BDS visible satellites in the region between 0 and 10 h is more than that in the region between 10 and 24 h. This is because that the signal of IGSO satellites in China can cover the moon area, which increases the number of BDS visible satellites. When the signals from these IGSO satellites are unable to cover the moon area, the number of BDS visible satellites decreases. Although one of GPS and BDS can guarantee 4 satellites visible at any time, the number of redundant satellites is still small. When receiving GPS and BDS jointly, the number of visible satellites can be greatly increased.

GDOP is an important indicator to measure the positioning precision of GNSS. Figures 6 and 7 show GDOP variations under different receiver sensitivities. As shown in Fig. 6, when the receiver sensitivity is −185 dBW, the GDOP value of GPS and BDS is between 500 and 4000. At some time, the GDOP value of GPS and BDS can reach more than 8000. This is because that only 4 satellites in GPS or BDS is visible at this time and the distance between two of these satellites is very close. When receiving GPS and BDS jointly, the GDOP decreases to about 1000 due to the increase of visible satellites.

As shown in Fig. 7, when the receiver sensitivity is −188 dBW, the GDOP value of GPS and BDS is reduced to the region between 500 and 2000. When receiving GPS and BDS jointly, the GDOP can be reduced to about 500. Reducing the GDOP can effectively improve the positioning accuracy of the GNSS receiver.

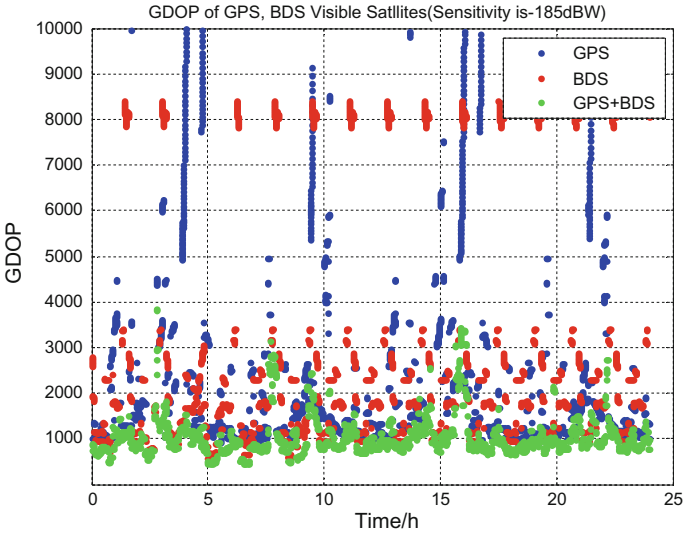


Fig. 6 GDOP of GPS, BDS, and GPS + BDS (receiver sensitivity is −185 dBW)

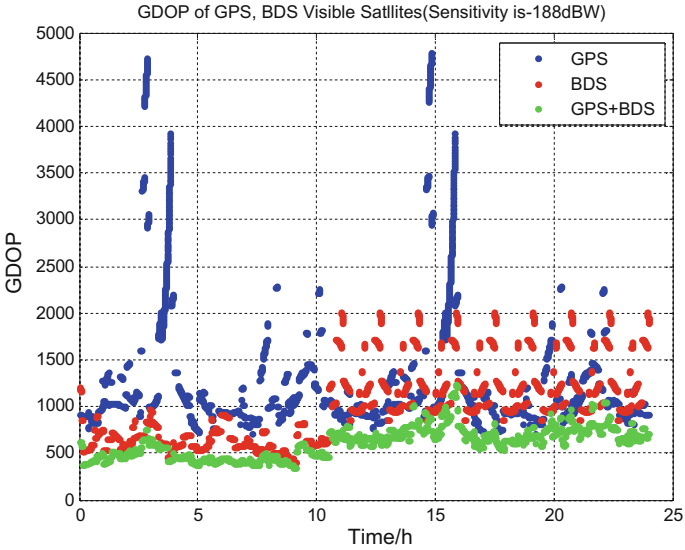


Fig. 7 GDOP of GPS, BDS, and GPS + BDS (receiver sensitivity is −188 dBW)

5 Conclusions

This paper analyzes the navigation performance of GPS and BDS on the lunar distance. In this paper, we analyze the path propagation loss of GPS and BDS signals on the moon distance. With the navigation signal frequency decreases from L1 and B1 frequency to L5 and B2 frequency, the signal propagation loss is reduced by 2.54 dB, so the sensitivity of GNSS receiver receiving L5 or B2 frequency signal can be 2.54 dB lower than that receiving L1 or B1 frequency. In addition, the transmitting power of B2a signal is 3 dB higher than that of B1C signal, which can further reduce the receiver sensitivity for B2a signal. This can reduce the complexity of GNSS receiver design. According to the antenna characteristics and the power of L1CA signal and B1C signal, we analyze the power change, the number variation, and the GDOP value of visible satellites in GPS and BDS under different receiver sensitivities. It is shown that the joint reception of GPS and BDS can effectively increase the number of visible and reduce the GDOP, which provides the basis for the design of GNSS lunar receiver.

Acknowledgements This work is supported by National Natural Science Foundation of China (Grant 91438107; Grant 61627817).

References

1. Miller J (2011) Enabling a fully interoperable GNSS space service volume. In: The 6th International Committee on GNSS(ICG), Tokyo, Japan
2. Winternitz LM, Bamford WA, Heckler GW (2009) A GPS receiver for high-altitude satellite navigation. IEEE J Sel Top Signal Process 3(4):541–556
3. Martzen P, Highsmith D (2015) GPS antenna characterization experiment(ACE): receiver design and initial results
4. Capuano V, Blunt P, Botteron C, Tian J, Wang Y, Leclere J, Farine P (2014) GNSS to reach moon. In: The 65th International Astronautical Congress, Toronto, Canada
5. Capuano V, Blunt P, Botteron C, Tian J, Leclere J, Wang F, Basile F, Farine P (2016) Standalone GPS L1 C/A. Sensors 16:347
6. ICD-GPS-200F Navstar GPS Space Segment/User Segment Interfaces (21 Sept 2011)
7. Xie G (2014) Principles of GPS and receiver design. Publishing House of Electronics Industry, Beijing
8. Kou Y (2012) Understanding GPS: principles and applications, 2nd edn. Publishing House of Electronics Industry, Beijing
9. China Satellite Navigation Office (2017) BeiDou navigation satellite system signal in space interface control document for B1C and B2a (Test Version)
10. China Satellite Navigation Office (2013) Specification for open service performance of Beidou satellite navigation system (Version 1.0)
11. China Satellite Navigation Office (2016) BeiDou Navigation satellite system signal in space interface control document (Version 2.1)

The Development of Real-Time Vector Receiver on Hardware Platform and the Assessment of Anti-spoofing Capability

Xinran Zhang, Hong Li, Chun Yang, Li He and Mingquan Lu

Abstract Anti-spoofing technology is currently a hot spot in the field of navigation. In comparison with the traditional GNSS receiver, vector receiver takes the advantage of the coupling of satellite signals in a combining way regarding the state of receiver. Limited to attack conditions and implementation complexity, spoofing signals are generally not coupled with authentic satellite signals. This means that vector receiver is likely to have anti-spoofing capability. To validate this conception, we develop a real-time GNSS vector receiver on FPGA platform and test its response to spoofing signals. The experimental results show that the vector receiver is insensitive to the spoofing attacks when the number of satellites is four. While if the number is five, some loops of the vector receiver will lose lock on the condition that spoofing signals delays are greater than $2 \sim 15$ μs. It demonstrates preliminarily that the vector receiver can detect spoofing attack to a certain extent. Moreover, we also investigate the influence of satellite geometrical distribution, spoofing signal delay and other factors on the anti-spoofing capability of vector receiver.

Keywords Vector receiver · Real-time · Spoofing attack · Anti-spoofing

1 Introduction

Global Navigation Satellite System (GNSS) has been widely used around the world. Nevertheless, navigation system is vulnerable to spoof due to low signal power as remote transmission, open structure of civil signals, and slow updating of navigation message. Spoofing is a kind of deliberate interference, in which the

X. Zhang · H. Li (✉) · L. He · M. Lu
Department of Electronic Engineering, Tsinghua University,
Beijing 100084, China
e-mail: lihongee@tsinghua.edu.cn

C. Yang
China Academy of Engineering Physics, Mianyang 621000, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_35

spoofer provides the receiver with erroneous position, velocity and time solutions by sending spoofing signals. This type of attack is more than subtle and the receiver is usually unaware of spoofing which might lead to extremely severe consequences [1]. Therefore, to prevent spoofing interference and improve user security has become an urgent topic in the field of navigation.

The traditional GNSS receiver adopts scalar receiving mode that each channel tracks a satellite signal respectively, then the receiver performs a PVT (Position-Velocity-Time) calculating by combining the observations from satellite signals to get a receiver state, including position, velocity and clock correction. But in fact, satellite signals have some couplings with each other instead of being completely independent. Vector receiving technique utilizes the coupling to track all satellite signals in a combining regarding the receiver state, which can significantly enhance tracking performance [2].

The concept of vector tracking was first proposed in Ref. [2]. It points out that the vector tracking algorithm can improve tracking performance by combining satellite signals together. Afterwards, a vector tracking structure based on Kalman filtering is proposed [3]. In which traditional scalar loop is replaced by Vector Delay Locked Loop (VDLL) and Vector Frequency Locked Loop (VFLL), and the simulation result shows that the vector tracking algorithm can reduce signal recapturing time and improve tracking performance. Receivers based on vector tracking structure also have advantages in anti-interference ability that the receiver can quickly recover the signal tracking after the interference is suppressed [4, 5]. In general, the vector receiver is superior to the traditional scalar receiver in terms of tracking performance and anti-interference ability. The development of vector receiver is also a hot spot at present.

Considering that the authentic satellite signals are always coupled with each other. While limited to attack conditions and implementation complexity, spoofing signals are generally not coupled with authentic satellite signals. This means that vector receiver is likely to have anti-spoofing capability. To validate this conception, we develop a real-time GNSS vector receiver on FPGA platform and test its response to spoofing signals. Moreover, we also investigate the influence of satellite geometry, spoofing signal delay and other factors on the anti-spoofing capability of vector receiver. The experimental results demonstrate that the vector receiver is insensitive to the spoofing attacks when the number of satellites is four. While if the number is five, some loops of the vector receiver will lose lock on the condition that spoofing signals delays are greater than 2–15 μs. For different satellite geometrical distributions, the minimum delay that leads to loss of lock is different. And the greater the amount of delay, the more likely the vector receiver is to lose lock. It demonstrates preliminarily that the vector receiver can detect spoofing attack to a certain extent. The research results can be used as reference for further GNSS anti-spoofing technology research and the development of GNSS customer terminal with anti-spoofing capability.

The rest part is arranged as follows. The second section introduces the principle and structure of vector receiver. The third section discusses the anti-spoofing

capability of vector receiver. The fourth section presents the anti-spoofing experiments and analyses of the results. The last section concludes the paper.

2 Vector Receiver

2.1 Satellite Signals Coupling

In the traditional GNSS scalar receiver, each channel separately tracks each satellite signal to obtain the carrier Doppler, code phase and other observations, as well as the satellite ephemeris. Afterward the receiver combines the observations of at least four satellite signals to get position, velocity and clock correction. In the process of tracking carrier Doppler and code phases, all the channels operate independently (Fig. 1).

In practice, however, the carrier Doppler and code phase of each satellite signal are directly relevant to the position, velocity, and clock correction of the receiver as shown in Eqs. (1) and (2) [6]:

$$f_n = -\frac{(x_u - x_n)(vx_u - vx_n) + (y_u - y_n)(vy_u - vy_n) + (z_u - z_n)(vz_u - vz_n)}{\lambda\sqrt{(x_u - x_n)^2 + (y_u - y_n)^2 + (z_u - z_n)^2}} \quad (1)$$

$$C_n = \frac{f_c\left(\sqrt{(x_u - x_n)^2 + (y_u - y_n)^2 + (z_u - z_n)^2} + \tau\right)}{c} \quad (2)$$

Among them, $(x_u, y_u, z_u, vx_u, vy_u, vz_u)$ and $(x_n, y_n, z_n, vx_n, vy_n, vz_n)$ represent the three-dimensional position and velocity of receiver and satellite of the channel n respectively. And τ denotes the receiver's clock correction. The carrier Doppler f_n

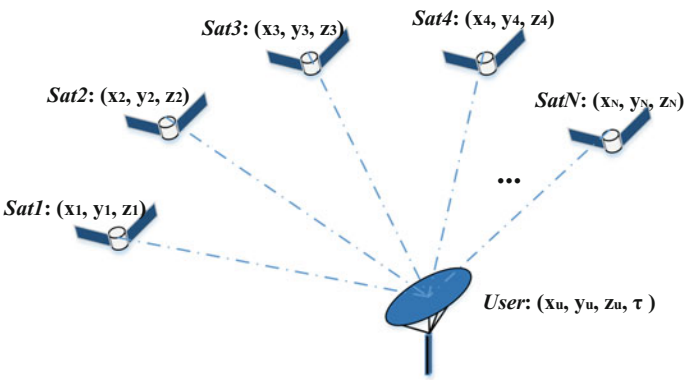


Fig. 1 GNSS positioning principle

can be derived from the relative velocity between receiver and satellite. We define λ to be the wavelength of the satellite signal. The code phase C_n is relevant to the pseudorange from receiver to each satellite, where f_c is the code rate and c is the velocity of light. That is, the satellite signals received are relevant to the receiver state. In the known ephemeris conditions, provided that the receiver's position, velocity, clock correction and other information are completely exact, carrier Doppler and code phase calculated by Eqs. (1) and (2) are exact too. Conversely, if the receiver's state is incorrect, the calculated carrier Doppler and code phase would very likely be erroneous.

2.2 *Vector Tracking Algorithm*

Since vector tracking algorithm requires a receiver state to start, the vector receiver we designed has a process of transition from scalar to vector mode [7]. The vector receiver adopts scalar mode at startup until a PVT solution is achieved. Once the receiver's position, velocity, clock correction and other information are obtained, the receiver switches to vector mode. The structure of vector mode receiver is shown in Fig. 2 [8].

In vector mode, the outputs of discriminator no longer update tracking loops directly, but adjust the carrier Doppler and code phase firstly:

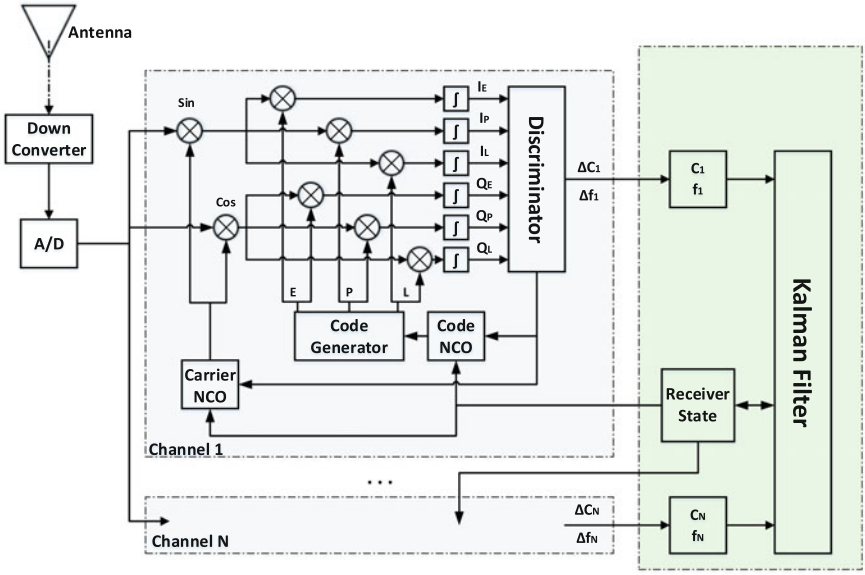


Fig. 2 The structure of vector mode receiver

$$f_n \Rightarrow f_n + \Delta f_n \tag{3}$$

$$C_n \Rightarrow C_n + \Delta C_n \tag{4}$$

Then the adjusted carrier Doppler f_n and code phase C_n are used as the observations of Kalman filter to update receiver state. Finally, carrier Doppler and code phase calculated by the corrected receiver state update tracking loops with discriminator outputs [9].

2.3 Kalman Filter

Kalman filter is the principle part of the vector tracking algorithm. The main idea is that whenever the system obtains a set of observations, the state of the system is corrected so that the system state estimation reaches the minimum mean square error [6]. In the Kalman filtering of the vector tracking algorithm, system state X is the state of the receiver, including three-dimensional position and velocity, clock correction and clock drift. The system observation Y contains code phases and carrier Doppler measurements of all receiving satellite signals adjusted by the discriminators [8, 9].

$$X = [\ x_u \quad vx_u \quad y_u \quad vy_u \quad z_u \quad vz_u \quad \tau_u \quad b_u\ ]^{\mathrm{T}} \tag{5}$$

$$Y = [\ C_1 \quad C_2 \quad \cdots \quad C_N \quad f_1 \quad f_2 \quad \cdots \quad f_N\ ]^{\mathrm{T}} \tag{6}$$

Assume that the state vector of the receiver is obtained at time k:

$$X_k = [\ x_{uk} \quad vx_{uk} \quad y_{uk} \quad vy_{uk} \quad z_{uk} \quad vz_{uk} \quad \tau_{uk} \quad b_{uk}\ ]^{\mathrm{T}} \tag{7}$$

Firstly, we estimate the receiver state at time $k+1$:

$$\hat{X}_k = AX_k \tag{8}$$

In Eq. (8), A is the state transition matrix and the transfer cycle is 1 ms (taking the GPS L1 C/A signal as an example). Subsequently, code phases and carrier Doppler of all receiving satellite signals are estimated through the receiver state according to Eqs. (1) and (2).

$$h(\hat{X}_k) = [\ \hat{C}_{1k} \quad \hat{C}_{2k} \quad \cdots \quad \hat{C}_{Nk} \quad \hat{f}_{1k} \quad \hat{f}_{2k} \quad \cdots \quad \hat{f}_{Nk}\ ]^{\mathrm{T}} \tag{9}$$

Finally, receiver state is corrected by the difference between the observed value Y_k and estimated value $h(\hat{X}_k)$ of code phases and carrier Doppler:

$$X_{k+1} = \hat{X}_k + K_k \times \left(Y_k - h(\hat{X}_k)\right) \tag{10}$$

And K_k is the gain matrix of Kalman filter [9].

After Kalman filtering is accomplished and corrected receiver state X_{k+1} is obtained, the difference between the estimated value $h(X_{k+1})$ and observed value Y_k of code phases and carrier Doppler is combined with the outputs of discriminators $[\ \Delta C_1 \quad \Delta C_2 \quad \cdots \quad \Delta C_N \quad \Delta f_1 \quad \Delta f_2 \quad \cdots \quad \Delta f_N\ ]^{\mathrm{T}}$ to update tracking loops. It is the entire process of vector tracking algorithm, which strengthens the tracking performance by combining satellite signals of all channels with receiver state.

3 Anti-spoofing Capability Analysis

Spoofing attack, no matter self-generated spoofing, or forwarding spoofing, aims to distort the PVT solution of the receiver.

As shown in Fig. 3, $U_{rt} : (x_{ut}, y_{ut}, z_{ut}, vx_{ut}, vy_{ut}, vz_{ut}, \tau_t)$ indicates the real three-dimensional position, velocity and clock correction of receiver. When one or more satellite signals are replaced by spoofing signals, the PVT solution $U_{rs} : (x_{us}, y_{us}, z_{us}, vx_{us}, vy_{us}, vz_{us}, \tau_s)$ deviates from the actual value. According to the GNSS positioning principle:

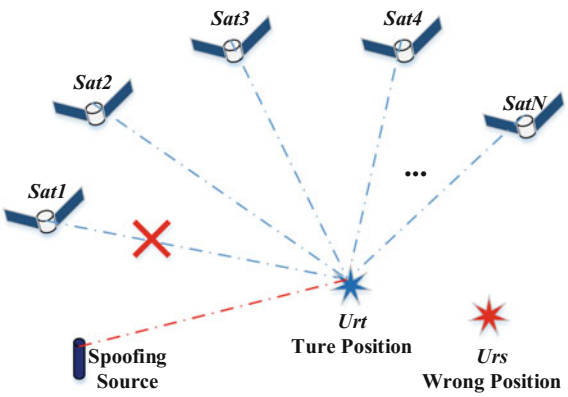


Fig. 3 Principle of spoofing attack

$$\begin{cases} \sqrt{(x_u - x_1)^2 + (y_u - y_1)^2 + (z_u - z_1)^2} + \tau = \rho_1 \\ \sqrt{(x_u - x_2)^2 + (y_u - y_2)^2 + (z_u - z_2)^2} + \tau = \rho_2 \\ \vdots \\ \sqrt{(x_u - x_N)^2 + (y_u - y_N)^2 + (z_u - z_N)^2} + \tau = \rho_N \end{cases} \tag{11}$$

The position and clock correction deviation from the true value in the presence of spoofing signals can be formulated as follows [10, 11]:

$$\begin{bmatrix} \varepsilon x \\ \varepsilon y \\ \varepsilon z \\ \varepsilon \tau \end{bmatrix} = \begin{bmatrix} x_{us} - x_{ut} \\ y_{us} - y_{ut} \\ z_{us} - z_{ut} \\ \tau_s - \tau_t \end{bmatrix} \approx \left(G^{\mathrm{T}} G\right)^{-1} G^{\mathrm{T}} \begin{bmatrix} \Delta\rho_1 \\ \Delta\rho_2 \\ \vdots \\ \Delta\rho_N \end{bmatrix} \tag{12}$$

In Eq. (12), matrix G is called geometric matrix that relates merely to the distribution of satellites relative to the receiver. Pseudorange deviation $[\Delta\rho_1 \quad \Delta\rho_2 \quad \cdots \quad \Delta\rho_N]^{\mathrm{T}}$ can be expressed as $\Delta\rho_n = c \times \Delta t_n$, which is the speed of light c multiplies spoofing delay Δt_n. When the signal is true satellite signal, $\Delta\rho_n = 0$. While if the signal is false, $\Delta\rho_n \neq 0$. Analysis reveals that the deviations of receiver position and clock correction are related to the matrix G and the spoofing delay.

Postulating that the Doppler shift of the spoofing signals relative to the true signals is small that the velocity deviation due to the Doppler shift can be neglected. (In fact, The Doppler shift cannot be neglected in reality environment. The neglect here is to facilitate the derivation. It happens only with perfect spoofing, and considering Doppler shift should make the spoofing detection more sensitive and effective.) Thereafter the deviation of the velocity in the presence of the spoofing signals can be expressed as:

$$\begin{aligned} \begin{bmatrix} \varepsilon vx \\ \varepsilon vy \\ \varepsilon vz \\ \varepsilon b \end{bmatrix} &= \frac{1}{\Delta t} \begin{bmatrix} x_{usB} - x_{usA} \\ y_{usB} - y_{usA} \\ z_{usB} - z_{usA} \\ \tau_{sB} - \tau_{sA} \end{bmatrix} - \frac{1}{\Delta t} \begin{bmatrix} x_{utB} - x_{utA} \\ y_{utB} - y_{utA} \\ z_{utB} - z_{utA} \\ \tau_{tB} - \tau_{tA} \end{bmatrix} \\ &= \frac{1}{\Delta t} \left\{ \begin{bmatrix} x_{usB} - x_{utB} \\ y_{usB} - y_{utB} \\ z_{usB} - z_{utB} \\ \tau_{sB} - \tau_{tB} \end{bmatrix} - \begin{bmatrix} x_{usA} - x_{utA} \\ y_{usA} - y_{utA} \\ z_{usA} - z_{utA} \\ \tau_{sA} - \tau_{tA} \end{bmatrix} \right\} \\ \approx \frac{1}{\Delta t} &\left\{ \left(G_B^{\mathrm{T}} G_B\right)^{-1} G_B^{\mathrm{T}} \begin{bmatrix} \Delta\rho_{B1} \\ \Delta\rho_{B2} \\ \vdots \\ \Delta\rho_{BN} \end{bmatrix} - \left(G_A^{\mathrm{T}} G_A\right)^{-1} G_A^{\mathrm{T}} \begin{bmatrix} \Delta\rho_{A1} \\ \Delta\rho_{A2} \\ \vdots \\ \Delta\rho_{AN} \end{bmatrix} \right\} \end{aligned} \tag{13}$$

The velocity is expressed as the rate of position shift. The true position of the receiver is from $U_{rtA} : (x_{utA}, y_{utA}, z_{utA}, \tau_{tA})$ to $U_{rtB} : (x_{utB}, y_{utB}, z_{utB}, \tau_{tB})$ in time Δt. While under the condition of spoofing signals, the position result is from $U_{rsA} : (x_{usA}, y_{usA}, z_{usA}, \tau_{sA})$ to $U_{rsB} : (x_{usB}, y_{usB}, z_{usB}, \tau_{sB})$. Assuming that the shift of pseudorange deviation is small enough, the deviation of velocity is mainly determined by the matrix G and its rate of shift, as well as the deviation of pseudorange.

From the structure of vector receiver, we can find that if the receiver state is incorrect, then the estimated carrier Doppler and code phases are likely different to the measurements from the received signals. When the deviation between the estimated value and the measured value reaches a certain threshold, the loss of lock will occur. This abnormal loss of lock in the loop is due to the disturbed coupling of the tracked signals and this disturbance comes from spoofing signals. This means that vector receiver is capable of detecting the presence of spoofing signal.

From all the analysis above, it is clear that the deviations of receiver position, velocity and clock correction in the presence of spoofing signals are relevant to the matrix G and its rate of shift, as well as the spoofing signals delays (deviation of pseudorange). While the deviations of receiver position, velocity and clock correction solution will give rise to the deviations of carrier Doppler and code phase estimations. Therefore, the anti-spoofing capability of vector receiver is also relevant to these factors.

4 Experimental Test and Result Analysis

To validate the conception that vector receiver possesses anti-spoofing capability, we develop a real-time GNSS vector receiver on FPGA platform and test it with GPS L1 C/A code signals.

4.1 Testing Platform

The testing platform is shown in Fig. 4.

The testing platform contains signal source simulator, transmitting antenna, receiving antenna, vector receiver and data displayer. The signal source simulator can simulate satellite RF signals at different times and scenarios, and can simulate the process of forwarding spoofing by adding delays to signals. The simulated signals are emitted from the transmitting antenna, received by the receiving antenna and transmitted to the downconverter. Then the vector receiver processes the down converted signals and the calculation results are displayed on the data displayer.

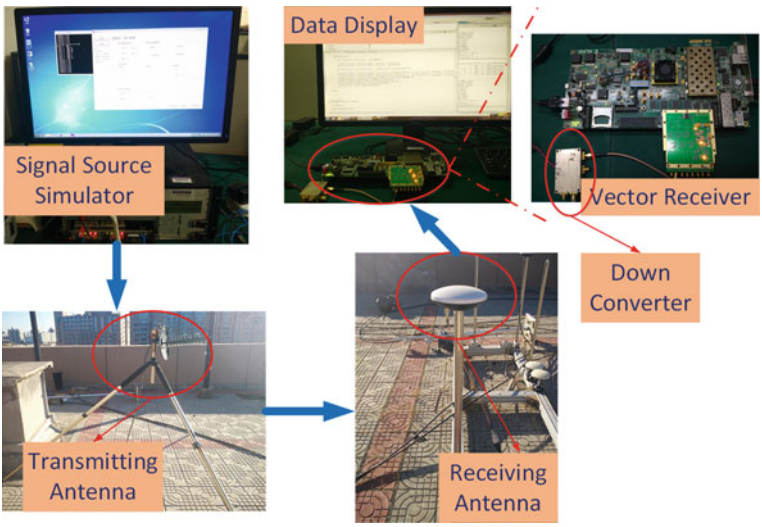


Fig. 4 GNSS vector receiver testing platform

4.2 Testing Results and Analysis

Taking into the account that the limited computation capacity of the hardware-based test platform, we mainly test the response of receiver under different satellite geometrical distributions and spoofing delays for four or five satellites. The signal source simulator uses GPS ephemeris on November 16, 2017, and the signal power at the down converter is around −120dBm.

4.2.1 Four Satellites

The vector receiver is insensitive to the spoofing signals when the number of satellites is four. After the receiver switches to vector mode, it takes time for the loops to lose lock. The loss of lock times in different satellite combinations and different spoofing delays are tested respectively for No. 11 and No. 28 satellite. The testing results are shown in Table 1.

The results show that when four satellites, the vector receiver can respond to spoofing signals with delays roughly greater than 1 ms. After a period of time some of the loops will lose lock. And the greater the amount of delay, the faster the vector receiver loses lock.

Four satellites situation is special to the vector receiver. When all four satellite signals are true, the code phases $[C_{1t} \quad C_{2t} \quad C_{3t} \quad C_{4t}]^{\mathrm{T}}$ estimated by the actual receiver position and clock correction $U_{rt} : (x_{ut}, y_{ut}, z_{ut}, \tau_t)$, are equal to the code

Table 1 Loss of lock time under different delays for four satellites

Delay / *Satellite Number*	*10ms*	*5ms*	*1ms*	*500us*
(11,3,1,17)	2min	3min	11min	Long enough
(11,3,6,28)	1min	1min	8min	Long enough
(11,1,6,22)	1min	1min	8min	Long enough
(28,1,3,17)	1min	1min	7min	Long enough
(28,6,3,11)	1min	2min	6min	Long enough
(28,11,17,1)	0.5min	5min	10min	Long enough

phases $\left[\frac{f_c \cdot \rho_1}{c} \quad \frac{f_c \cdot \rho_2}{c} \quad \frac{f_c \cdot \rho_3}{c} \quad \frac{f_c \cdot \rho_4}{c}\right]^{\mathrm{T}}$ measured from the signal [according to the Eq. (11)]:

$$C_{nt} = \frac{f_c\left(\sqrt{(x_{ut}-x_n)^2+(y_{ut}-y_n)^2+(z_{ut}-z_n)^2}+\tau_t\right)}{c} = \frac{f_c \times \rho_n}{c} \tag{14}$$

Considering the existence of spoofing signals. Take the case that only one satellite signal is false as an example. It is assumed that the deviation of pseudo-range due to spoofing delay is $\Delta\rho_1$. As shown in equation set (15):

$$\begin{cases} \sqrt{(x_{us}-x_1)^2+(y_{us}-y_1)^2+(z_{us}-z_1)^2}+\tau_t = \rho_1+\Delta\rho_1 \\ \sqrt{(x_{us}-x_2)^2+(y_{us}-y_2)^2+(z_{us}-z_2)^2}+\tau_t = \rho_2 \\ \sqrt{(x_{us}-x_3)^2+(y_{us}-y_3)^2+(z_{us}-z_3)^2}+\tau_t = \rho_3 \\ \sqrt{(x_{us}-x_4)^2+(y_{us}-y_4)^2+(z_{us}-z_4)^2}+\tau_t = \rho_4 \end{cases} \tag{15}$$

There are discrepancies between the solutions $U_{rs}:(x_{us},y_{us},z_{us},\tau_s)$ of Eq. (15) and the true values $U_{rt}:(x_{ut},y_{ut},z_{ut},\tau_t)$. The code phase C_{ns} estimated by $U_{rs}:(x_{us},y_{us},z_{us},\tau_s)$ can be expressed as:

$$C_{ns} = \frac{f_c\left(\sqrt{(x_{us}-x_n)^2+(y_{us}-y_n)^2+(z_{us}-z_n)^2}+\tau_s\right)}{c} \tag{16}$$

While the measured values of the code phases obtained from the signal are:

$$\left[\frac{f_c(\rho_1+\Delta\rho_1)}{c} \quad \frac{f_c\cdot\rho_2}{c} \quad \frac{f_c\cdot\rho_3}{c} \quad \frac{f_c\cdot\rho_4}{c}\right]^{\mathrm{T}} \tag{17}$$

It can be seen from the Eq. (15):

$$\left[C_{1s} \quad C_{2s} \quad C_{3s} \quad C_{4s}\right]^T = \left[\frac{f_c(\rho_1+\Delta\rho_1)}{c} \quad \frac{f_c\cdot\rho_2}{c} \quad \frac{f_c\cdot\rho_3}{c} \quad \frac{f_c\cdot\rho_4}{c}\right]^{\mathrm{T}} \tag{18}$$

For four satellites, there are four unknowns (three-dimensional position and clock correction of the receiver) and four equations of the GNSS positioning system. The calculation results satisfy each equation. Therefore, even if the calculated receiver state deviates from the true value due to the spoofing signals delays, the subsequent code phases estimated by the erroneous receiver state and the measurement results obtained from the signals are still basically equal. Therefore when four satellites, the deviations of code phases are not enough to cause the loss of lock in code loop.

Spoofing delays will not only lead to the deviation of receiver's position and clock correction, but also add a velocity that is caused by the shift of matrix G and can keep stable over a period time. And this velocity deviation is proportional to the delays. Therefore, when the spoofing delays reach a certain value, the vector receiver loops may lose lock.

4.2.2 Five Satellites

Spoofing delay gives rise to the deviations of receiver state. Unlike the case of four satellites, in a five satellites scenario, there must be code phase discrepancy between the estimations from erroneous receiver state and the measurements from received signals. The receiver loops will lose lock when the deviation reaches a certain threshold. Therefore, when the number of satellites is more than four, the sensitivity of vector receiver to spoofing signals increases greatly.

Testing on the condition of five satellites. Situation 1: one satellite signal is false and the other four satellite signals are true. Add spoofing delay to No. 3, No. 17 and No. 28 satellite separately in different satellite combinations. Situation 2: two satellite signals are false and other three satellite signals are true. Add spoofing delay to No. 3 with No. 17 satellites, and No. 8 with No. 11 satellites separately in different satellite combinations. And T donates the vector receiver keeping working normally while S donates losing lock after the receiver witches into vector mode. All test results are presented in Tables 2, 3, 4, 5 and 6.

Table 2 shows the working status of vector receiver in different satellite geometrical distributions and spoofing delays for the No. 3 satellite. Normal working (T) of the vector receiver indicates that no spoofing has been detected, while the case of loss of lock (S) indicates that the vector receiver has response to spoofing signals. Detecting spoofing critical delay is the minimum amount of delay that can

Table 2 Response to the spoofing attack of the satellite 3

delay(us) / *Satellite Number*	*2*	*3*	*4*	*6*	*8*	*10*	*12*	*15*	*30*	*Critical delay*
(3,11,7,8,1)	T	T	T	T	T	S	S	S	S	**10us**
(3,7,28,19,11)	T	S	S	S	S	S	S	S	S	**3us**
(3,8,28,7,19)	T	T	S	S	S	S	S	S	S	**4us**
(3,17,8,1,11)	T	T	T	T	T	T	S	S	S	**12us**
(3,6,22,17,19)	T	S	S	S	S	S	S	S	S	**3us**
(3,6,22,1,28)	T	T	S	S	S	S	S	S	S	**4us**

Table 3 Response to the spoofing attack of the satellite 17

delay(us) / *Satellite Number*	*2*	*3*	*4*	*6*	*8*	*10*	*12*	*15*	*30*	*Critical delay*
(17,11,7,8,1)	T	T	T	T	T	T	S	S	S	**12us**
(17,28,19,7,3)	T	T	S	S	S	S	S	S	S	**4us**
(17,19,1,8,3)	T	S	S	S	S	S	S	S	S	**3us**
(17,8,11,7,28)	T	T	T	T	T	S	S	S	S	**10us**
(17,6,22,19,1)	T	T	T	T	T	S	S	S	S	**10us**
(17,6,22,28,3)	T	T	T	S	S	S	S	S	S	**6us**

Table 4 Response to the spoofing attack of the satellite 28

delay(us) / *Satellite Number*	*2*	*3*	*4*	*6*	*8*	*10*	*12*	*15*	*30*	*Critical delay*
(28,11,7,8,1)	T	T	T	S	S	S	S	S	S	**6us**
(28,3,7,19,11)	S	S	S	S	S	S	S	S	S	**2us**
(28,3,17,8,19)	T	S	S	S	S	S	S	S	S	**3us**
(28,17,8,1,11)	T	S	S	S	S	S	S	S	S	**3us**
(28,6,22,7,19)	T	T	T	S	S	S	S	S	S	**6us**
(28,6,22,1,3)	T	T	T	S	S	S	S	S	S	**6us**

Table 5 Response to the spoofing attack of the satellite 3 and 17

delay(us) / *Satellite Number*	*2*	*3*	*4*	*6*	*8*	*10*	*12*	*15*	*30*	*Critical delay*
(3,17,1,11,8)	T	T	S	S	S	S	S	S	S	**4us**
(3,17,6,11,8)	S	S	S	S	S	S	S	S	S	**2us**
(3,17,1,8,28)	T	S	S	S	S	S	S	S	S	**3us**

Table 6 Response to the spoofing attack of the satellite 8 and 11

delay(us) / *Satellite Number*	*2*	*3*	*4*	*6*	*8*	*10*	*12*	*15*	*30*	*Critical delay*
(8,11,3,17,1)	T	S	S	S	S	S	S	S	S	**3us**
(8,11,6,28,1)	T	S	S	S	S	S	S	S	S	**3us**
(8,11,17,6,28)	T	T	T	S	S	S	S	S	S	**6us**

lead a vector receiver to lose lock. Tables 3 and 4 are the testing results for No. 17 and No. 28 satellite. Tables 5 and 6 are the testing results of two spoofing satellite signals.

The experimental results show that when the number of satellites is five, vector receiver can detect the spoofing attack that spoofing delay is greater than 2–15 μs. For different satellite geometrical distributions, the minimum delay that can be detected is different. And for the same satellite geometrical distribution, the greater the amount of delay, the more likely the vector receiver is to lose lock. When the number of satellites is more than five, due to the computational burden, the current platform cannot work normally. The follow-up effort will solve the platform's problems. Then we can test the anti-spoofing capability of vector receiver in the case of more satellites to further improve the test system.

But on the whole, through theoretical analysis and experiments, it demonstrates preliminarily that the vector receiver indeed has anti-spoofing capability.

5 Conclusion

In this paper, we study the structure of vector receiver and analysis that the carrier Doppler and code phase estimated by the erroneous receiver state as spoofing signals are likely different to the measurements from signals. When the deviation

between them reaches a certain value, the tracking loops of vector receiver will lose lock. Furthermore, we demonstrate that due to its nature, the vector receiver cannot work normally in the presence of spoofing signals. Thus, we propose the conception that vector receiver has anti-spoofing capability. Through developing a real-time vector receiver on FPGA platform, we carried out series of experiments to demonstrate this conception. Moreover, we also investigate the influence of satellite geometry, spoofing signal delay and other factors on the anti-spoofing capability of vector receiver. For different satellite geometrical distributions, the minimum spoofing delays triggering loss of lock are different. As for the same satellite geometrical distribution, the greater the amount of delay, the more likely the vector receiver is to lose lock. The research results can be used as reference for further GNSS anti-spoofing technology research and the development of GNSS customer terminal with anti-spoofing capability.

Acknowledgements This study is funded by the National Natural Science Foundation of China (61571255).

References

1. Swaszek PF, Hartnett RJ (2013) Spoof detection using multiple COTS receivers in safety critical applications. In: Proceedings of international technical meeting of the satellite division of the institute of navigation
2. Spilker JJ (1996) Fundamentals of signal tracking theory. In: Global positioning system: theory and applications, vol. I. Progress in astronautics and aeronautics, vol 163, AIAA, Washington, DC, St. Petersburg, Russia
3. Lashley M, Bevly DM (2007) Analysis of discriminator based vector tracking algorithms. In: Proceedings of ION NTM 2007, pp 570–576, Institute of Navigation, San Diego, CA, 22–24 Jan 2007
4. Li F, Wu R, Wang W (2016) The anti-jamming performance analysis for vector tracking loop. In: 7th China satellite navigation conference, CSNC 2016, 18–20 May 2016
5. Sun Z, Wang X, Feng S et al (2016) Design of an adaptive GPS vector tracking loop with the detection and isolation of contaminated channels. GPS Solut 1–13
6. Xie G (2012) Principles of GPS and receiver design. Publishing House of Electronics Industry, Beijing, pp 69–155
7. Wu C, Xie S, Zhang J et al (2017) The real-time realization and performance analysis of BDS vector receiver. In: China satellite navigation conference (CSNC) 2017 proceedings, vol I
8. Vu BN, Andrle M (2014) Implementation of vector tracking loop algorithm in modern GPS receiver. In: Digital avionics systems conference IEEE, 2014, pp 3E1(1–9)
9. Chen X, Wang X, Xu Y (2014) Performance enhancement for a GPS vector-tracking loop utilizing an adaptive iterated extended Kalman filter. Sensors 14(12):23630–23649
10. Wang Q, Li H (2013) Residual vector analysis method (RVAM) for evaluating the performance of GNSS Part of channels' replay attacks. In: Proceedings of IEEE China Summit & international conference on signal and information processing, 2013, pp 561–565
11. Bai W, Li H, Yang Y et al (2016) Motion state monitoring based GNSS spoofing detection method for repeater spoofing attack. In: Proceeding of the 2016 international technical meetings of the institute of navigation, pp 732–738

Applying Probabilistic Graphical Models in a GNSS Software Receiver

Xin Zhang, Lingxiao Zheng and Xingqun Zhan

Abstract A probabilistic graphical modelling approach to the parameter and state estimation problems in a GNSS receiver is proposed. This method turns out to be universal and comprehensive since as a framework it can accommodate both a traditional software receiver with scalar tracking loop (STL) and one with a vector tracking loop (VTL). It can even be extended to a multi-sensor integration setting and serves as a candidate software architecture for All Source Positioning and Navigation (ASPN).

Keywords PGM · GNSS · INS · Integration · Sensor

1 Introduction

Probabilistic graphical model (PGM) is a graph-theoretic framework for representation, inference, and learning [1–7]. This paper reviews those PGM-based methods that are suitable for different stages of signal processing in a contemporary scaler or vector GNSS software receiver.

Traditionally, signal processing in a Global Navigation Satellite System (GNSS) software receiver focuses on (i) understanding explicit relationship between consecutive parts of the processing stages including acquisition, tracking, and navigation solution; (ii) closed-form, non-probabilistic solutions to each of these stages subject to certain simplifications or assumptions. Here 'non-probabilistic' refers to inexplicit inclusion of any conditional probabilities in equations.

This work tries to consider the parameter and state estimation problems encountered in a GNSS receiver from a different viewpoint. A framework based on PGMs will be explored wherever possible. Previous researches have been dedicated to one or several stages of the signal processing chain, either in a pure GNSS

X. Zhang (✉) · L. Zheng · X. Zhan
School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai, China
e-mail: xin.zhang@sjtu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_36

receiver or multi-sensor integration setting. The current attempt is expected to facilitate improvements of information utilization within a receiver. The reason why this approach should possess a decent performance is the data-centric trends in modern computing. Several possible candidates of PGM are readily available from modern machine learning. In particular, PGM combines the power of machine learning with its unique strength in modeling the independencies (or dependencies) between random variables.

The paper is structured as follows. The theoretic and application background of this research is **first** introduced in this section. This is followed by the **second** part where the framework of PGM is summarized and those results that both reside in this framework and are important for the parameter and state estimation problems in a contemporary GNSS receiver is given in proper orders. In particular, results related to template model and independence preservation and evolution will be presented, which are the keys to facilitating the current tracking problem. In the **third** section, PGM application in signal tracking will be examined. This is followed by examples of PGM application in in a vector GNSS receiver in the **fourth** part and an example of PGM application in multisensor integration of GNSS with other sensors in the **fifth** part. The final part provides some discussions on its direct or indirect applications and concludes the paper.

2 Probabilistic Graphical Models (PGM) Framework

Estimation problems in navigation and in particular, GNSS and INS/GNSS integration can be modelled and solved using the framework of PGM. This section summarizes the basic set of rules that will be useful in later development of estimator equations.

2.1 Representation

Throughout the paper, we will mainly deal with Dynamic Bayesian Networks (DBN). Specifically, the estimation problems, as we shall show later, belong to a class of temporal modelling and inference problem. A typical unrolled DBN (or 'ground network' [8]) is shown in Fig. 1. Here 'unrolled' means temporal

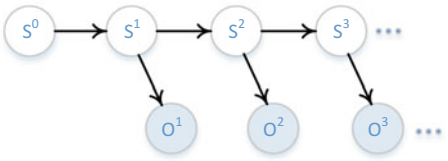


Fig. 1 A typical DBN for a parameter/state estimation problem

repetition of a template. This network can also be considered as Hidden Markov Model (HMM). $S^{(t)}$, t = 1, 2, … is the hidden state (vector).

A list of symbols used in later development of inference paradigm is given here for easy reference. This list also includes some of the symbols present in Fig. 1.

t	Current epoch.
$S^{(t)}$	state vector at time t.
$O^{(t)}$	observation vector at time t.
$O^{(1:t)}$	observation (vector) from time 1 (time at which the first observation is obtained) through t (including time t).
$\sigma^{(t)}(S^{(t)})$, $P(S^{(t)}\|o^{(1:t)})$	the belief state at time t of current state $S^{(t)}$ given observations up to the current time, i.e. time t. In $P(\cdot)$, symbol S is uppercase because this is a random variable while symbol o is lowercase because this is considered observed here.
$\sigma^{(\cdot t+1)}(S^{(t+1)})$, $P(S^{(t+1)}\|o^{(1:t)})$	the belief state at time t + 1 of current state $S^{(t+1)}$ given observations up to the previous time, i.e. time t.

To distinguish between these two categories of nodes easily in Fig. 1, the always observed node '$o^{(t)}$' (lower case 'o') is grayed.

2.2 Inference

Using the DBN model shown in Fig. 1, inference can be accomplished in four steps, which in turn could be summarized as two equations.

Step 1. Unroll the template model (either DBN or plate model) to obtain the ground (unrolled) temporal network.

A typical 2-TBN is shown in Fig. 2, where a robot localization problem is used as an example. The states at each time epoch include robot's acceleration, velocity, and position, in addition to modeled sensor failure and weather conditions. The inter-relationships between these variables within one time slice and between two consecutive time slices are shown at the right side of Fig. 2. The left side of Fig. 2 shows the necessary initial condition.

Figure 3 shows the unrolled 2-TBN from Fig. 2 and is named ground or unrolled network. In this example, only variables at time t = 0, 1, 2 are shown.

Now assume we are in the middle of the iterative estimation process. Steps 2 through 4 describes how to propagate the current state to the next.

Step 2. Input belief state obtained from last step, i.e. $\sigma^{(t)}(S^{(t)})$.

Step 3. Belief state propagation: use $\sigma^{(t)}(S^{(t)})$ to output $\sigma^{(\cdot t+1)}(S^{(t+1)})$.

Step 4. Output belief state: use $\sigma^{(\cdot t+1)}(S^{(t+1)})$ to output $\sigma^{(t+1)}(S^{(t+1)})$.

Steps 2 through 4 can be summarized as two equations [8].

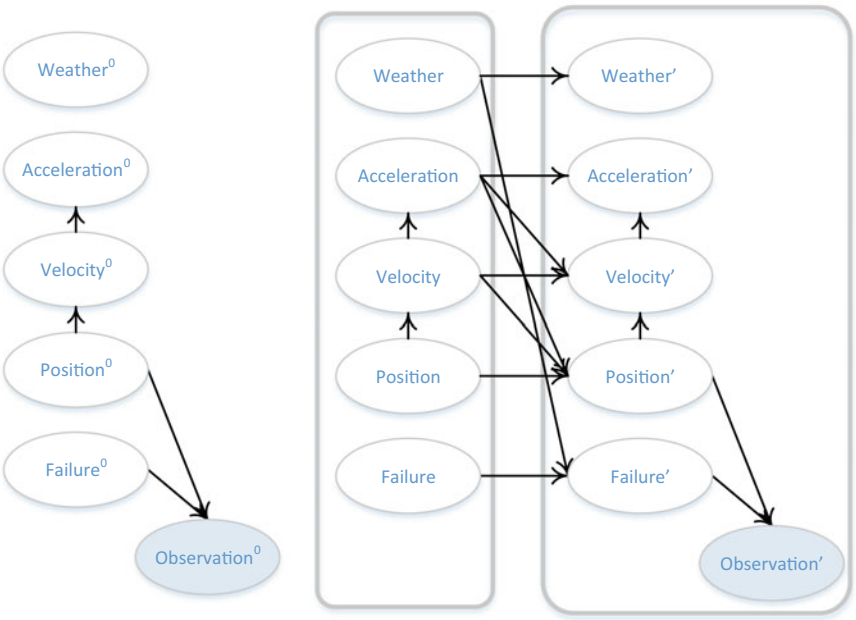


Fig. 2 A 2-TBN example

$$\begin{aligned}\sigma^{(\cdot t+1)}\left(X^{(t+1)}\right) &= P\left(X^{(t+1)}|o^{(1:t)}\right)\\ &= \sum_{X^{(t)}} P(X^{(t+1)}|o^{(1:t)}, X^{(t)})P(X^{(t)}|o^{(1:t)})\\ &= \sum_{X^{(t)}} P(X^{(t+1)}|X^{(t)})\sigma^{(t)}(X^{(t)}).\end{aligned} \tag{1}$$

which is belief state forward propagation and

$$\begin{aligned}\sigma^{(t+1)}(X^{(t+1)}) &= P(X^{(t+1)}|o^{(1:t)}, o^{(t+1)})\\ &= \frac{P(o^{(t+1)}|X^{(t+1)}, o^{(1:t)})P(X^{(t+1)}|o^{(1:t)})}{P(o^{(t+1)}|o^{(1:t)})}\\ &= \frac{P(o^{(t+1)}|X^{(t+1)})\sigma^{(\cdot t+1)}(X^{(t+1)})}{P(o^{(t+1)}|o^{(1:t)})}.\end{aligned} \tag{2}$$

which is conditioning or re-normalization

As shown in later sections. These two equations can be rewritten in many different forms among which the most easy-to-understand one is the Kalman filter [10], and it will be shown that the prediction and update of a Kalman filter correspond exactly to (1) and (2), respectively.

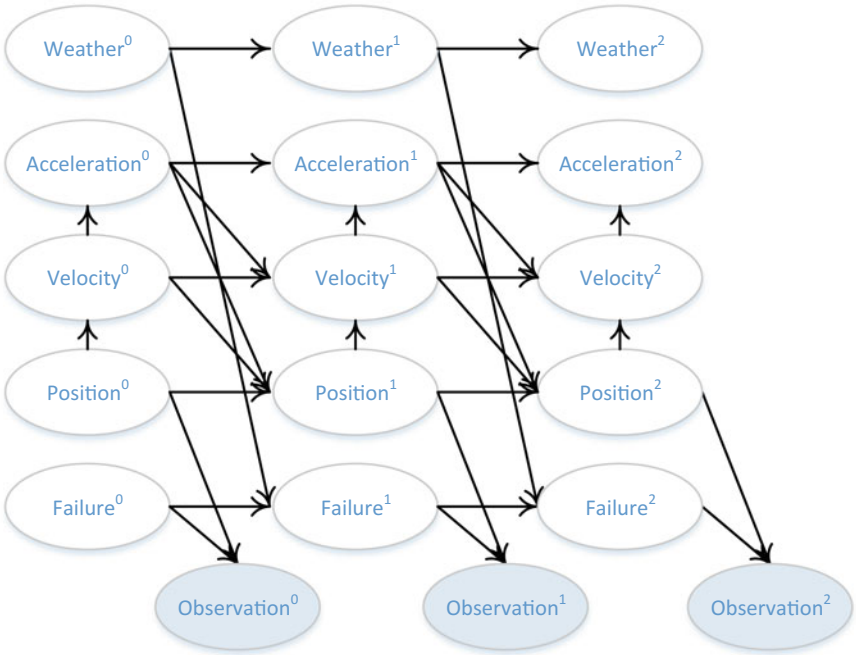


Fig. 3 Unrolled 2-TBN from Fig. 2

3 Applying PGM to GNSS SDR

Estimation problems, whether a DLL or PLL, can be readily fitted into the framework. Two common implementations take place here. One is scale tracking loop (STL), and the other is vector tracking loop (VTL).

3.1 STL

Step 1. Representation.
In case of STL, what is estimated is a fixed variable, θ, using the received signal r. In case of PLL, θ is carrier phase, while in DLL, θ is code phase. Usually a LSQ estimator is used. In this case, the transition model that is characteristic of an HMM is nullified (the transition probability is 1, since we consider θ is a constant in a relatively small interval around the linearization point of the S-curve), and hence our model will degenerate into a Naïve Bayes model (NB), shown in Fig. 4.
Steps 2, 3, & 4. Inference & Probabilistic Independence
In the right-hand side of Fig. 4, r^1 through r^n are mutually independent given the parameter, θ. This is a typical application of the concept of 'active trail' in PGM

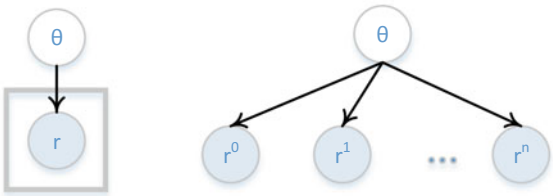


Fig. 4 The Naïve Bayes model used for STL

framework. Here the trail between every pair of r^i and r^j, $i \neq j$, is not active given θ and this means that probabilistic influence cannot flow from r^i to r^j, if θ is given. This property is called probabilistic independence. Whether or not this property is preserved will be essential in deriving different forms of PGM model and its corresponding inference algorithms.

Let's assume a pure carrier (i.e. code wipe-off is complete) and the received signal is

$$r(t) = x(t, \theta) + n(t) = A\cos(2\pi f_c t + \theta) + n(t) \tag{3}$$

where $n(t)$ is white Gaussian noise (with power spectrum of $N_0/2$ over the bandwidth). The receiver can observe $r(t)$ for a given time interval from $t = 0$ to $t = t_0$ and is trying to estimate θ. Using techniques similar to that of Gram-Schmidt orthonormalization, it is possible to find a basis of K orthonormal signals $\Psi_k(t)$, which should be guaranteed complete (the method of choosing this set however, is beyond the scope of this paper) for $x(t, \theta)$ (for all the possible values of θ) for $t \in [0, t_0]$:

$$x(t, \theta) = \sum_{k=1}^{K} x_k(\theta)\psi_k(t), \quad t \in [0, t_0] \tag{4}$$

Then the signal $x(t, \theta)$ can be represented through vector $\mathbf{x}(\theta)$, and the useful part of r(t) can be represented through vector

$$\boldsymbol{r} = \boldsymbol{x}(\theta) + \boldsymbol{n} \tag{5}$$

where $\boldsymbol{n}$ is a vector of K statistically independent Gaussian random variables with zero mean and variance equal to $N_0/2$. Given θ, the conditional probability density function for $\boldsymbol{r}$, encoded by the model as shown in the right-hand side of Fig. 4 is

$$f_{\boldsymbol{r}|\theta}(\boldsymbol{r}|\theta) = \frac{1}{(\pi N_0)^{K/2}} \exp\left(-\frac{\|\boldsymbol{r} - \boldsymbol{x}(\theta)\|^2}{N_0}\right) \tag{6}$$

The ML synchronizer outputs the estimate $\hat{\theta}$ of θ that maximizes $f_{\boldsymbol{r}|\theta}(\boldsymbol{r}|\theta)$ for the received $\boldsymbol{r}$:

$$\hat{\theta}_{ML} = \max_{\theta}^{-1}[f_{\boldsymbol{r}|\theta}(\boldsymbol{r}|\theta)] \tag{7}$$

Solving (7) we can obtain the structure of our desired tracking loop including Numerically Controlled Oscillator (NCO), loop filter and discriminator function. Solutions can be easily found in a classical PLL textbook [9].

3.2 VTL

In case of VTL, the model takes the usual form of HMM. An example unrolled network is shown in Fig. 5.

Here P^i is the position at time i, O^i the observation, and S_j the presence of satellites. The subscript of S_i represents PRN number assigned to this satellite. As an example, only GPS satellites are considered here, so S_j ranges from S_1 through S_{32}. In this particular setting, at time 1 and 2, satellites 1, 2, and 32 are visible while at time 3, only satellites 2 and 32 are visible.

4 Extended Framework for Multisensor Integration

As we have shown in the previous three examples, what PGM describes can be the probabilistic independence between state and observed variables and signal parameters. In this regard, this framework can be readily suitable for solving multi-sensor integration problems. An example of INS/GNSS Integration is shown in Fig. 6.

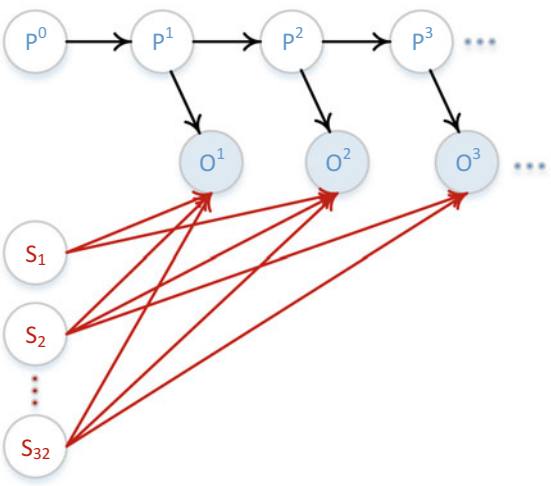


Fig. 5 Unrolled network for VTL

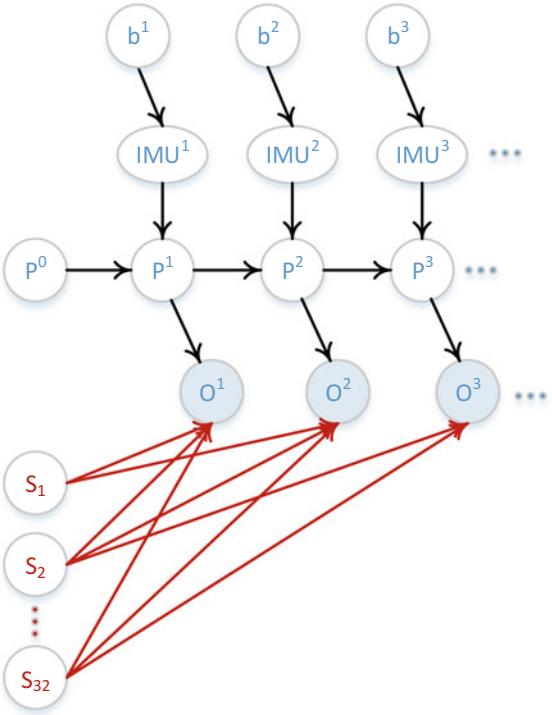


Fig. 6 Model extended for INS/GNSS integration

Here b^i (i = 1, 2, 3, …) is the bias at time epoch i. IMU^i is the IMU readings (usually rate) at the same epoch. This is a simple illustration of how PGM can integrate two sensors (i.e. INS and a GNSS receiver). Integration of more than two sensors can be easily accommodated into this framework by adding nodes representing parameter, state variables, or observables carefully in an epoch by epoch manner. An example of INS/GNSS integration using DBN and sampling methods can be found in [11].

5 Conclusions

As shown in Sects. 2 through 4, PGM can be considered as a universal framework. It is universal because

(1) It can describe an estimation problem. Moreover, the modelled problem can be solved using various existing algorithms (e.g. EKF) or other algorithms based on graph theories (for example, A*, Dijkstra's, or even Convolutional Neural Networks (CNNs)). The latter is extremely resourceful. This is the primary advantage that this framework can offer and this is especially promising in multi-sensor fusion since a whole bunch of sensors of different nature, and

inference algorithms, and machine learning algorithms can be unified in a same framework and thus solved, using existing proved algorithms.

(2) It is theoretically important since it provides an intuitive way to explain the probabilistic relationships within statistical estimation problems that once seem to be different. Whether the probabilistic independence property is preserved with the temporal evolution of the model will be essential in deriving different forms of PGM model and its corresponding inference algorithms.

In this paper, a PGM framework is suggested to describe and solve estimation problems in multisensor avigation. This approach turns out to be universal since it can accommodate estimation problems of different nature, whether it is a two-stage problem (i.e. tracking and navigation in a GNSS receiver) including both parameter and state estimation or it is a multisensor integration problem trying to fuse sensors of different natures. It can be considered as a promising candidate to the general problems of All Source Positioning and Navigation (ASPN).

Acknowledgements This work is supported by National Natural Science Foundation of China (No. 61403253) "A plug-and-play solution to All Source Positioning and Navigation problems based on factor-graph."

References

1. Sahawneh LR, Brink KM (2017) Factor graphs-based multi-robot cooperative localization: a study of shared information influence on optimization accuracy and consistency. In: Proceedings of the 2017 international technical meeting of the institute of navigation (ION ITM 2017), Monterey, California, January 2017, pp 819–838
2. Chen W, Zeng QH, Liu JY, Chen LJ, Wang HZ (2016) Research on the multi-sensor information fusion method based on factor graph. In: Proceedings of IEEE/ION PLANS 2016, Savannah, GA, April 2016, pp 502–506
3. Chen D, Gao X (2015) Robust MAV state estimation using an M-estimator augmented sensor fusion graph. In: Proceedings of the 28th international technical meeting of the satellite division of the institute of navigation (ION GNSS + 2015), Tampa, Florida, September 2015, pp 841–848
4. Chen D, Gao X (2015) Simultaneous state estimation of UAV trajectory using probabilistic graph models. In: Proceedings of the 2015 international technical meeting of the institute of navigation (ION ITM 2015), Dana Point, California, Jan 2015, pp 804–810
5. Trinh NH, Zhou X, Chiu HP, Samarasekera S, Kumar R (2013) A general approach to online sensor calibration for navigation using factor graphs. In: Proceedings of the 2013 international technical meeting of the institute of navigation (ION ITM 2013), San Diego, California, Jan 2013, pp 901–908
6. Kaess M (2008) iSAM: incremental smoothing and mapping. IEEE Trans Robot 24(6)
7. Kaess M (2011) iSAM2: incremental smoothing and mapping with fluid relinearization and incremental variable reordering. In: IEEE international conference on robotics and automation (ICRA 2011), May 2011
8. Koller D, Friedman N (2009) Probabilistic graphical models, 1st edn. The MIT Press, London

9. Myer H, Moeneclaey M, Fechtel SA (1997) Digital communication receivers, synchronization, channel estimation and signal processing. 1st edn. Wiley-Interscience
10. Kalman RE (1960) A new approach to linear filtering and prediction problems. Retrieved on 4 Oct 2017 at https://www.cs.unc.edu/∼welch/kalman/media/pdf/Kalman1960.pdf
11. Zhang X, Zhan X (2017) Graphical approach to representation and inference in multi-sensor state estimation, In: Proceedings of the 30th international technical meeting of the satellite division of the institute of navigation (ION GNSS + 2017), 25–29 Sept 2017, Portland, OR, USA. pp 2603–2611

GNSS Spoofing Mitigation Method After Despreading

Zhenglin Geng, Yangbo Huang, Huaming Chen and Feixue Wang

Abstract GNSS spoofing signals have the same signal structure as authentic signals, so it is difficult to mitigate the spoofing in time and frequency domains, and space domain processing has become one of the most powerful methods for spoofing cancelation. But most of the existing spoofing mitigation methods are implemented before signal despreading and their computational complexity is generally low, but their mitigation performance deteriorates seriously as the spoofing power decreases. To solve the problem, a spoofing mitigation method after signal despreading is proposed in this paper. Simulation results show that the proposed method is practicable and outperforms the methods before signal despreading in a large variation range of spoofing power.

Keywords GNSS · Spoofing mitigation · Despreading

1 Introduction

Spoofing mitigation is the most powerful GNSS anti-spoofing method because it can effectively eliminate the impact of spoofing and restore the availability of the receiver [1]. In 2007, McDowell proposed a spoofing detection and mitigation system based on the spatial correlation of spoofing signals [2]. The system uses the signal electronic angle to distinguish the spoofing and can eliminate the spoofing and jamming simultaneously, but hardware modification is required to apply the method for a conventional receiver. Subsequently, Daneshmand proposed a low-complexity spoofing mitigation method in 2011 [3] and improved it in 2012 [4], which utilizes the cyclical characteristic of the GPS C/A signal PRN code and the uncorrelation of noise on different antenna to estimate the spatial signature vector (SSV) of the spoofing. And then the orthogonal subspace of the spoofing is

Z. Geng · Y. Huang (✉) · H. Chen · F. Wang
College of Electronic Science, National University of Defense Technology, Changsha 410073, China
e-mail: yangbohuang@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_37

calculated and the received signals are projected into the orthogonal subspace. In literature [5], eigenvalue decomposition method is used to estimate the subspace of noise then the received signals are projected into the noise subspace.

In addition, the space-time processing approaches are used to cancel the spoofing in literature [6] and [7]. However, the above methods are seriously affected by the power of the spoofing, and the mitigation performance deteriorates significantly when the spoofing power is low. In order to ensure the spoofing cancelation performance under the condition of relatively low power of spoofing, this paper proposes a method to mitigate the spoofing after signal despreading.

2 Signal Model

In order to present a credible navigation solution to the receiver, a spoofer is required to transmit at least 4 GNSS spoofing signals simultaneously. It is generally assumed that the spoofing signals are transmitted from a single antenna due to the restriction of cost and engineering implementation, so, they arrival the receiving antenna in the same direction [8]. Considering GPS L1 C/A code, the signals received by a N_{Ant}-element antenna-array receiver can be expressed as

$$\mathbf{r}(nT_s) = \mathbf{A}\mathbf{S}^a(nT_s) + \mathbf{B}\mathbf{S}^s(nT_s) + \boldsymbol{\eta}(nT_s) \tag{1}$$

where,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{N_{Auth}} \end{bmatrix} \tag{2}$$

$$\mathrm{B} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_{N_{Spoof}} \end{bmatrix} \tag{3}$$

$$\mathbf{S}^a(nT_s) = \begin{bmatrix} \sqrt{p_1^a} F_1^a(nT_s) \\ \vdots \\ \sqrt{p_{N_{Auth}}^a} F_{N_{Auth}}^a(nT_s) \end{bmatrix} \tag{4}$$

$$\mathbf{S}^s(nT_s) = \begin{bmatrix} \sqrt{p_1^s} F_1^s(nT_s) \\ \vdots \\ \sqrt{p_{N_{Spoof}}^s} F_{N_{Spoof}}^s(nT_s) \end{bmatrix} \tag{5}$$

$$\boldsymbol{\eta}(nT_s) = \begin{bmatrix} \eta_1(nT_s) \\ \vdots \\ \eta_{N_{Ant}}(nT_s) \end{bmatrix} \tag{6}$$

where,

$$F_m^a(nT_s) = d_m^a\left(nT_s - \tau_m^a\right)c_m^a\left(nT_s - \tau_m^a\right)\mathrm{e}^{j\phi_m^a + j2\pi f_m^a nT_s} \tag{7}$$

$$F_q^s(nT_s) = d_q^s\left(nT_s - \tau_q^s\right)c_q^s\left(nT_s - \tau_q^s\right)\mathrm{e}^{j\phi_q^s + j2\pi f_q^s nT_s} \tag{8}$$

where, T_s is the time interval, N_{Auth} and N_{Spoof} are the number of authentic signals and spoofing signals respectively. $\mathbf{a}_m$ and $\mathbf{b}_q$ are the SSVs [3] of the authentic and spoofing signals respectively. The $\boldsymbol{\eta}(nT_s)$ is the complex additive white Gaussian noise vector with covariance matrix $\sigma^2\mathbf{I}$, where, the $\mathbf{I}$ is an identity matrix. p, ϕ, f and τ are the signal power, carrier phase, Doppler frequency and code delay of the received signals respectively. $d(nT_s)$ and $c(nT_s)$ represent navigation data bits and PRN code. The superscripts 'a' and 's' represent the authentic and spoofing signals, and the subscripts 'm' and 'q' represent the *m-th* authentic and *q-th* spoofing signal respectively. ϕ_m^a and ϕ_q^s are the carrier phases of the *m-th* authentic and *q-th* spoofing signal at the reference antenna.

Because the spoofing signals with different PRN numbers have the same arrival angle and the influence of signal wavelength on SSV $\mathbf{b}_q$ is negligible, the subscripts 'q' is omitted. Equation (1) can be rewritten as

$$\mathbf{r}(nT_s) = \mathbf{A}\mathbf{S}^a(nT_s) + \mathbf{b}\mathbf{l}_{1\times N_{Spoof}}\mathbf{S}^s(nT_s) + \boldsymbol{\eta}(nT_s) \tag{9}$$

where, $\mathbf{l}_{1\times N_{Spoof}}$ is an all one vector.

3 Spoofing Suppression Method

The spoofing mitigation methods before signal despreading are generally based on the spatial correlation of the spoofing, and rely on the presence of a dominant spatial power in order to extract the spoofing SSV or calculate the subspace of the spoofing. So, the decrease of the spoofing power will significantly deteriorate their performance. To improve the spoofing mitigation performance under the condition of low spoofing signal power, a method after despreading is proposed in this paper, and its block diagram is shown in Fig. 1. This method can be mainly divided into 4 steps, namely signal parameter estimation, array signal dispreading, spoofing identification and spoofing mitigation.

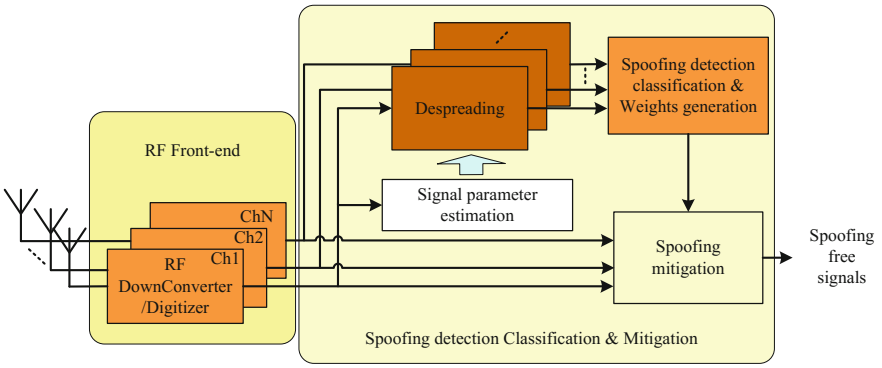


Fig. 1 Block diagram of the proposed method

3.1 Signal Parameter Estimation

In this step, a reference channel is selected among the plurality of channels firstly, and the parameters of all the signals received by the reference channel can be estimated after signals fully tracked.

Take the signal with pseudo random noise (PRN) code i for example, the parameters include the code delay $\hat{\tau}_i$, carrier Doppler frequency $\hat{f}_i$, navigation data bit $\hat{d}_i$ and carrier phase $\hat{\phi}_i$.

3.2 Array Signal Despreading

After signal parameter estimation, the authentic and spoofing signal can be despreaded according to the estimated parameters. ds_i^a and ds_i^s are constructed based on the parameters which belong to the authentic and spoofing signal. Superscripts *'a'* and *'s'* are used to indicate the authentic and spoofing signal.

$$ds_i^a(nT_s) = \hat{d}_i^a\left(nT_s - \hat{\tau}_i^a\right)c_i^a\left(nT_s - \hat{\tau}_i^a\right)\mathrm{e}^{-j\hat{\phi}_i^a - j2\pi\hat{f}_i^a nT_s} \tag{10}$$

$$ds_i^s(nT_s) = \hat{d}_i^s\left(nT_s - \hat{\tau}_i^s\right)c_i^s\left(nT_s - \hat{\tau}_i^s\right)\mathrm{e}^{-j\hat{\phi}_i^s - j2\pi\hat{f}_i^s nT_s} \tag{11}$$

The authentic and spoofing signals after despreading can be expressed as

$$\mathbf{x}_i^a(k) = \sum_{n=1}^{L}\mathbf{r}((k-1)LT_s + nT_s)\cdot ds_i^a((k-1)LT_s + nT_s) \tag{12}$$

$$\mathbf{x}_i^s(k) = \sum_{n=1}^{L} \mathbf{r}((k-1)LT_s + nT_s) \cdot ds_i^s((k-1)LT_s + nT_s) \tag{13}$$

Using $\mathbf{x}_i(k)$ to express the correlator output, and the L is the data length of correlation, if the authentic signal is captured,

$$\mathbf{x}_i(k) = \mathbf{x}_i^a(k) \approx \mathbf{a}_i \sqrt{p_i^a} L + \tilde{\boldsymbol{\eta}}_i^a(k) \tag{14}$$

and if the spoofing signal is captured,

$$\mathbf{x}_i(k) = \mathbf{x}_i^s(k) \approx \mathbf{b} \sqrt{p_i^s} L + \tilde{\boldsymbol{\eta}}_i^s(k) \tag{15}$$

where,

$$\tilde{\boldsymbol{\eta}}_i^a(k) = \sum_{n=1}^{L} \boldsymbol{\eta}_i((k-1)LT_s + nT_s) \cdot ds_i^a((k-1)LT_s + nT_s) \tag{16}$$

$$\tilde{\boldsymbol{\eta}}_i^s(k) = \sum_{n=1}^{L} \boldsymbol{\eta}_i((k-1)LT_s + nT_s) \cdot ds_i^s((k-1)LT_s + nT_s). \tag{17}$$

3.3 Spoofing Identification

The spoofing identification can be implemented from two aspects, namely signal power and signal direction of arrival (DOA) [9, 10]. On the one hand, in order to drag the tracking loop to the spoofing signal effectively, the spoofing signal needs higher power than the authentic one, so that spoofing could be identified from the signal power. On the other hand, the spoofing signals with different PRN numbers are usually transmitted by a single antenna, so they arrive at the receiver in the same direction, which can be used for spoofing classification.

(1) Power identification. Construct a test statistic $p = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_i^H(k) \mathbf{x}_i(k)$, which denotes the signal power after despreading. Set a binary hypothesis. '*Th*' is the detection threshold.

H0: $p < Th$; H1: $p > Th$
H0: Authentic signal; H1: Spoofing signal

(2) DOA identification. Construct a test statistic $\boldsymbol{\alpha}_i = \angle \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_i(k)$, where, $\angle \cdot$ denotes the phase angle calculation. Without loss of generality, the first element of $\boldsymbol{\alpha}_i$ is selected as a reference, which is denoted by $\alpha_{i,1}$, calculate the phase difference $\boldsymbol{\beta}_i$, $\boldsymbol{\beta}_i = \boldsymbol{\alpha}_i - \alpha_{i,1}\mathbf{1}$, $\mathbf{1}$ is an $N_{Ant} \times 1$ all 1 vector. When the phase difference of different signals are obtained, we calculate the difference $\boldsymbol{\gamma}_{ij}$ between tow arbitrary signals, where, $\boldsymbol{\gamma}_{ij} = \boldsymbol{\beta}_i - \boldsymbol{\beta}_j$. If at least one of the two

signals compared is an authentic signal, the signal phase difference γ_{ij} is usually nonzero, and if the two signals compared are both spoofing signals, each element of γ_{ij} approximates zero. Set test statistic $r_{ij} = \gamma_{ij}^H \gamma_{ij}$, then r_{ij} can be used for spoofing identification. Where, *'Th'* is the detection threshold.

H0: $r_{ij} > Th$; H1: $r_{ij} < Th$
H0: At least one of the two compared signals is an authentic signal;
H1: Both compared signals are spoofing signals

Through the two methods above, the spoofing could be distinguished effectively, then, it can be mitigated by the following methods.

3.4 Spoofing Mitigation

After spoofing identification, the correlator output $\mathbf{x}_i(k)$ can be divided into authentic and spoofing categories, and represented by $\mathbf{x}_i^a(k)$ and $\mathbf{x}_i^s(k)$. And in spoofing mitigation, the key point is to generate a weight to mitigate the spoofing signals by multiplying with the received signals, and 3 different approaches can be used for weights generation, which are described as follow.

(a) **Subspace Projection**

In the subsection, the SSVs of spoofing and authentic signals are estimated and the subspace projection method [4] is used for weights generation. The details are described as follow. The estimation of $\mathbf{b}$ is indicated by $\hat{\mathbf{b}}$.

$$\hat{\mathbf{b}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{x}_i^s(k) \approx \mathbf{b}\sqrt{p_i^s}L \tag{18}$$

The orthogonal subspace of $\hat{\mathbf{b}}$ is denoted by

$$\begin{aligned} \mathbf{P}_\perp =& \mathbf{I}_N - \hat{\mathbf{b}}\left(\hat{\mathbf{b}}^H\hat{\mathbf{b}}\right)^{-1}\hat{\mathbf{b}}^H \\ =& \mathbf{I}_N - \mathbf{b}\sqrt{p_i^s}L\left(\mathbf{b}^H\sqrt{p_i^s}L\mathbf{b}\sqrt{p_i^s}L\right)^{-1}\mathbf{b}^H\sqrt{p_i^s}L \\ =& \mathbf{I}_N - \mathbf{b}\left(\mathbf{b}^H\mathbf{b}\right)^{-1}\mathbf{b}^H \end{aligned} \tag{19}$$

Obviously, $\mathbf{P}_\perp$ is also the orthogonal subspace of $\mathbf{b}$, and $\mathbf{P}_\perp = \mathbf{P}_\perp^H$. Furthermore, arbitrary $N_{Ant} \times 1$ unit vector $\mathbf{h}$ satisfies the relation $\mathbf{W} = \mathbf{P}_\perp\mathbf{h} \in \mathbf{P}_\perp$. Where, $\mathbf{W}$ denotes the weight. Then, signal after weighting is denoted by $v(nT_s)$ as

$$\begin{aligned} v(nT_s) =& \mathbf{W}^H \mathbf{r}(nT_s) = \mathbf{h}^H \mathbf{P}_\perp^H \mathbf{r}(nT_s) \\ =& \sum_{m=1}^{N_{Auth}} \mathbf{h}^H \mathbf{P}_\perp^H \mathbf{a}_m \sqrt{p_m^a} F_m^a(nT_s) + \mathbf{h}^H \mathbf{P}_\perp^H \boldsymbol{\eta}(nT_s) \end{aligned} \tag{20}$$

It can be seen that the spoofing is cancelled after weighting, and when $\mathbf{h} = \mathbf{P}_\perp^H \mathbf{a}_m / \left| \mathbf{P}_\perp^H \mathbf{a}_m \right|$, the *m-th* authentic signal power is maximized. $\mathbf{a}_m$ can be estimated in the same way, and be denoted by $\hat{\mathbf{a}}_m$.

$$\hat{\mathbf{a}}_m = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_m^a(k) \approx \mathbf{a}_m \sqrt{p_m^a} L \tag{21}$$

Then,

$$\mathbf{h} = \mathbf{P}_\perp^H \hat{\mathbf{a}}_m / \left| \mathbf{P}_\perp^H \hat{\mathbf{a}}_m \right| \approx \mathbf{P}_\perp^H \mathbf{a}_m \sqrt{p_m^a} L / \left| \mathbf{P}_\perp^H \mathbf{a}_m \sqrt{p_m^a} L \right| = \mathbf{P}_\perp^H \mathbf{a}_m / \left| \mathbf{P}_\perp^H \mathbf{a}_m \right|. \tag{22}$$

(b) **Eigenvalue Decomposition**

In this subsection a method based on eigenvalue decomposition [5] is discussed. The covariance matrix $\mathbf{R}_{xx}$ can be calculated using $\mathbf{x}_i^s(k)$ and expressed in the form as Eq. (23).

$$\mathbf{R}_{xx} = E\left[\mathbf{x}_i^s(k) \mathbf{x}_i^{sH}(k) \right] = \begin{bmatrix} \mathbf{v} & \mathbf{U} \end{bmatrix} \begin{bmatrix} p_i^s L^2 + \sigma_n^2 & 0 \\ 0 & \sigma_n^2 \end{bmatrix} \begin{bmatrix} \mathbf{v}^H \\ \mathbf{U}^H \end{bmatrix} \tag{23}$$

where, $\mathbf{v}$ is the eigenvector corresponding to the maximal eigenvalue, and $\mathbf{U}$ is the other eigenvectors namely the noise subspace. The received signals can be projected into the noise subspace to mitigate the spoofing. Using $\mathbf{W}$ denotes the weight, where,$\mathbf{W} = \mathbf{U}\mathbf{h} \in \mathbf{U}$ and $\mathbf{h}$ is an arbitrary $(N_{Ant} - 1) \times 1$ unit vector. Signal after weighting can be written as

$$\begin{aligned} v(nT_s) =& \mathbf{W}^H \mathbf{r}(nT_s) = \mathbf{h}^H \mathbf{U}^H \mathbf{r}(nT_s) \\ =& \sum_{m=1}^{N_{Auth}} \mathbf{h}^H \mathbf{U}^H \mathbf{a}_m \sqrt{p_m^a} F_m^a(nT_s) + \mathbf{h}^H \mathbf{U}^H \boldsymbol{\eta}(nT_s). \end{aligned} \tag{24}$$

Obviously, the output power of *m-th* authentic signal can be maximized when $\mathbf{h} = \mathbf{U}^H \mathbf{a}_m / \left| \mathbf{U}^H \mathbf{a}_m \right|$, we can use the $\hat{\mathbf{a}}_m$ to calculate the $\mathbf{h}$.

$$\mathbf{h} = \mathbf{U}^H \hat{\mathbf{a}}_m / \left| \mathbf{U}^H \hat{\mathbf{a}}_m \right| = \mathbf{U}^H \mathbf{a}_m \sqrt{p_i^a} L / \left| \mathbf{U}^H \mathbf{a}_m \sqrt{p_i^a} L \right| = \mathbf{U}^H \mathbf{a}_m / \left| \mathbf{U}^H \mathbf{a}_m \right| \tag{25}$$

(c) **PI and MVDR Technique**

The covariance matrix of the despreaded spoofing signal $\mathbf{R}_{xx}$ can be written in another form as Eq. (26).

$$\begin{aligned}\mathbf{R}_{xx} &= E\left[\mathbf{x}_i^s(k)\mathbf{x}_i^{sH}(k)\right] \\ &= E\left[\left(\mathbf{b}\sqrt{p_i^s}L + \tilde{\boldsymbol{\eta}}_i^s(k)\right)\left(\mathbf{b}^H\sqrt{p_i^s}L + \tilde{\boldsymbol{\eta}}_i^{sH}(k)\right)\right] \\ &\approx \mathbf{b}p_i^s L^2\mathbf{b}^H + \mathbf{I}\tilde{\sigma}^2\end{aligned} \tag{26}$$

where, $\tilde{\sigma}^2$ is noise variance after despreading, weight according to the power inversion (PI) technique [11] is

$$\mathbf{W}_{PI} = \frac{\mathbf{R}_{xx}^{-1}\mathbf{b}}{\mathbf{b}^H\mathbf{R}_{xx}^{-1}\mathbf{b}} \tag{27}$$

where, $\mathbf{b} = [1, 0, \ldots, 0]^T$. If the spatial signature vector of authentic signal is known, minimum variance distortionless response (MVDR) technique can be used [12], according weight is

$$\mathbf{W}_{MVDR} = \frac{\mathbf{R}_{xx}^{-1}\mathbf{a}_m}{\mathbf{a}_m^H\mathbf{R}_{xx}^{-1}\mathbf{a}_m}. \tag{28}$$

Replacing the $\mathbf{a}_m$ with $\hat{\mathbf{a}}_m$, then

$$\mathbf{W}_{MVDR} = \frac{\mathbf{R}_{xx}^{-1}\hat{\mathbf{a}}_m}{\hat{\mathbf{a}}_m^H\mathbf{R}_{xx}^{-1}\hat{\mathbf{a}}_m} = \frac{\mathbf{R}_{xx}^{-1}\mathbf{a}_m}{\mathbf{a}_m^H\mathbf{R}_{xx}^{-1}\mathbf{a}_m}\frac{1}{\sqrt{p_i^a}L} \tag{29}$$

In the performance comparison of the three methods above, the weights are normalized and multiply with the SSV of the spoofing signal, and their normalized inner products are used to measure their spoofing suppression performance.

4 Simulation Results

In this section, the proposed method is simulated and its performance is compared with the spoofing mitigation methods before signal despreading which are described in literature [4] and [5]. A set of spoofing and authentic signals in intermediate frequency with random DOAs are generated and used to calculate the weights according to the proposed method. The detailed parameters are shown in the Tab.1. Where, f_s, f_c, f_{RF} and f_{IF} are sampling frequency, code frequency, radio frequency (RF) and intermediate frequency (IF) of the carrier, respectively. B_w represents the bandwidth of the IF filter.

Table 1 Simulation parameters

Parameter	Value
N_{Auth}	9
N_{Spoof}	9
f_s	38.192 MHz
f_c	1.023 MHz
f_{RF}	1575.42 MHz
f_{IF}	9.548 MHz
B_w	8.184 MHz
Authentic CNR	45 dB Hz
Spoofing CNR	41–50 dB Hz

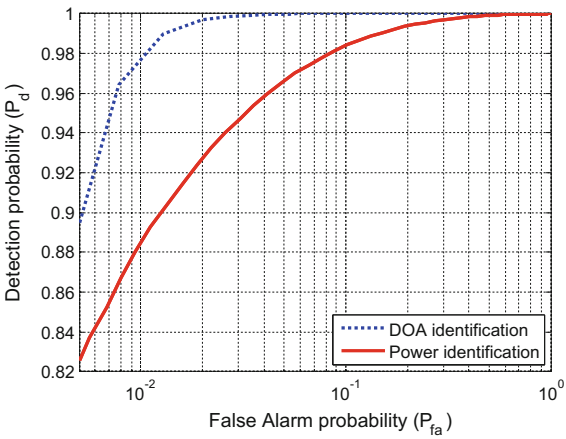


Fig. 2 ROC curves of spoofing identification

First, the spoofing identification methods described above are simulated. The receiver operating characteristic (ROC) curves are used as a measure of spoofing identification performance and obtained from 30000 Monte Carlo runs as shown in Fig. 2. In this figure, 10 ms signal are used for spoofing identification. And the power of each spoofing signals is 3 dB higher than the authentic one's and the DOAs of spoofing and authentic signals randomly distribute in elevation angle 0°–90° and azimuth angle 0°–360° in these simulations.

It is observed that the spoofing can be effectively distinguished by the power identification method and DOA identification method when the power of spoofing is 3 dB higher than the authentic one's or spoofing signals are transmitted from one antenna.

Subsequently, the normalized inner products of the weights and the SSVs of spoofing and authentic signals are used to measure the spoofing suppression performance and authentic signal power maximization performance respectively. And the simulation results are shown in Figs. 3 and 4. The spoofing mitigation performance of the proposed method is compared with the spoofing mitigation methods before signal despreading which are described in literature [4] and [5].

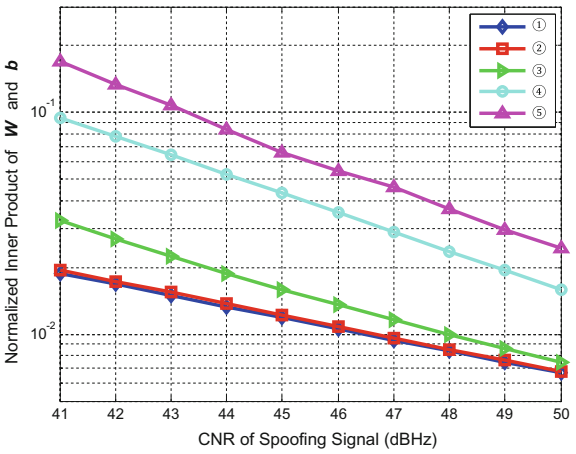
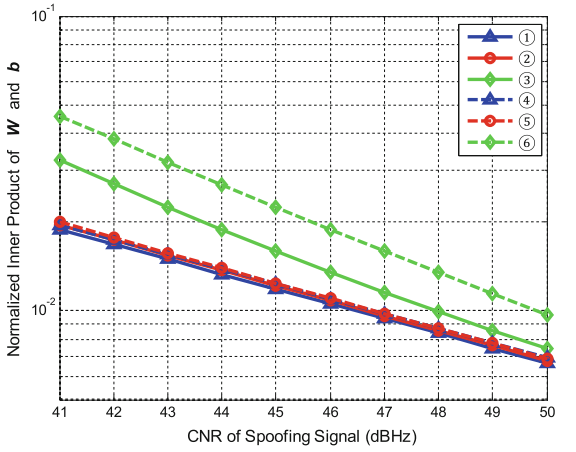
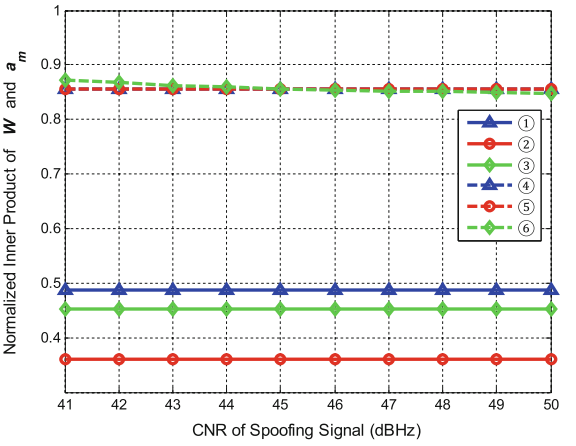


Fig. 3 Normalized inner products of weights and SSV of spoofing



(a) Normalized inner products of weights and SSV of spoofing



(b) Normalized inner products of weights and SSV of authentic signal

Fig. 4 Normalized inner products of weights and SSVs of authentic and spoofing signal before and after authentic signal power maximization

In Fig. 3, curves ①, ② and ③ denote the proposed methods with different weights generation approaches subspace projection, eigenvalue decomposition and PI technique, curves ④ and ⑤ denote the methods described in literature [4] and [5] respectively.

In Fig. 4a, the curves ①, ②, ③ and ④, ⑤, ⑥ indicate the normalized inner products of spoofing SSV and the weights generated based on the proposed method with different weights generation approaches before and after authentic signal power maximization respectively, and Fig. 4b, the curves ①, ②, ③ and ④, ⑤, ⑥ indicate the normalized inner products of authentic SSV and the weights generated based on the proposed method with different weights generation approaches before and after authentic signal power maximization respectively.

It can be seen that the spoofing suppression performance of the method after signals despreading outperforms the methods before signals despreading, and the suppression performance improves as the spoofing power increasing. However, the mitigation of spoofing potentially causes attenuation of the authentic signals before power maximization, and after beamforming in the direction of the authentic signal, the impact to the authentic signal significantly reduced, meanwhile, the spoofing mitigation performance is almost unaffected.

5 Conclusion

In this paper, a spoofing mitigation methods by space processing after signal despreading is proposed. This approach is less affected by the power of spoofing, and a beam in the direction of the authentic signal can be formed to effectively enhance the power of authentic signals. But, because it is necessary to capture and track all received signals, and distinguish the authentic and spoofing signals, the computational complexity of the methods are much higher than those before despreading.

Acknowledgements This work was supported by the National Natural Science Foundation of China under grant No. 61403413, 41604016 and 61601485.

References

1. Broumandan A, Jafarnia-Jahromi A, Daneshmand S et al (2016) Overview of spatial processing approaches for GNSS structural interference detection and mitigation. Proc IEEE 104(6):1246–1257
2. McDowell CE (2007) GPS spoofer and repeater mitigation system using digital spatial nulling. US Patent 7250903 B1, 7 p, 2007
3. Daneshmand S, Jafarnia-Jahromi A, Broumandan A, Lachapelle G (2011) A low complexity GNSS spoofing mitigation technique using a double antenna array, GPS World magazine, December 2011, vol 22, no 12, pp 44–46

4. Daneshmand S, Jafarnia-Jahromi A, Broumandan A, Lachapelle G (2012) A low-complexity GPS anti-spoofing method using a multi-antenna array. In: Proceedings of the 25th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2012), 17–21 September, Nashville TN, 11 p
5. Daneshmand S, Jafarnia-Jahromi A, Broumandan A, et al (2014) A GNSS structural interference mitigation technique using antenna array processing. In: IEEE, sensor array and multichannel signal processing workshop, 2014, pp 109–112
6. Guo Y, Fan M, Kong M (2012) Spoofing interference mitigation using space-time process for GNSS receiver. In: International congress on image and signal processing, 2012, pp 1537–1541
7. Daneshmand S, Jafarnia-Jahromi A, Broumandan A, et al (2013) GNSS spoofing mitigation in multipath environments using space-time processing. In: European Navigation Conference 2013, Vienna, Austria
8. Montgomery PY, Humphreys TE, Ledvina BM (2009) Receiver-autonomous spoofing detection: experimental results of a multi-antenna receiver defense against a portable civil GPS spoofer. Proc. Int Tech Meet Inst Navig Itm 1(1):124–130
9. Dehghanian V, Nielsen J, Lachapelle G (2012) GNSS spoofing detection based on signal power measurements: statistical analysis. Int J Navig Obs 2012(7)
10. Jafarnia-Jahromi A, Broumandan A, Daneshmand S et al (2014) A double antenna approach toward detection, classification and mitigation of GNSS structural interference. NAVITEC, Noordwijk, The Netherlands
11. Compton RTJ (1979) The power-inversion adaptive array: concept and performance. IEEE Trans Aerosp Electron Syst AES-15(6):803–814
12. Nguyen HN (2002) Robust steering vector mismatch techniques for reduced rank adaptive array signal processing. Virginia Tech, Blacksburg, Virginia

Design of Tracking Loop for Single Antenna Satellite Navigation Receivers Under the Situation of Rotation

Dan Song, Shengqun Geng, Zhigang Huang and Shuaiyong Zheng

Abstract It is necessary to measure the real-time rotation speed and the roll angle of a projectile relative to the geodetic coordinate system in order to achieve accurate guidance. To obtain the rotation speed of the projectile, a single antenna satellite receiver is used to receive the satellite signal. Since the received signal can provide amplitude modulation information relative to the roll characteristic of projectile, the signal can be demodulated to determine the rotation speed of the projectile. According to this principle, a real-time tracking loop of the satellite receiver which consists of a rotation tracking loop and a traditional tracking loop is designed to track satellite signal. Simulations under different rotation speed of projectile with single antenna satellite receiver were carried out. The mean values of relative error of projectile rotation speed are 5 and 3%, respectively, when the rotation speed of the projectile is 20 and 100 r/s, which can meet general engineering requirements. It is suited for the rotation speed (angle) measurement of a rotating object. The main contributions of this paper are a reference for the receiver tracking loop de-sign of military artillery and a method to measure the rotation speed of a rotating object.

Keywords Projectile · Rotation speed · Tracking loop · Single antenna satellite receiver

1 Introduction

Projectile trajectory correction is a kind of technology which is used to correct trajectory of a projectile in real-time during its flight. It aims to improve the hit rate and firing accuracy of projectiles [1]. It has been widely used in modern battlefield to strike a target accurately.

D. Song · S. Geng · Z. Huang (✉) · S. Zheng
Beihang University, Xueyuan Road No. 37, Haidian District,
Beijing, China
e-mail: baahzg@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_38

As is known to all, the projectile often flies in high dynamic motion, in the meantime rotating at a certain speed. In detail, the rotation speed of rocket is less than 20 r/s. As for wing-stabilized projectiles, their rotation speed ranges from 4 to 40 r/s. As for the stabilized rotating projectiles like cannon, their rotation speed can reach up to 200–300 r/s [2]. Rotation of projectile can guarantee its stability. Meanwhile, the guidance system can adjust its canard according to its rotation speed to make itself drift sideways and correct the horizontal ordinate of touchdown. Therefore, the measurement of rotation parameters is one of the core technologies of trajectory correction.

With the development of satellite navigation, single satellite receiver antenna-based rotation parameter measurement has become a key method of the projectile trajectory correction. In principle, the single antenna of projectile receiver can receive satellite navigation signal which contains the information of carrier amplitude and phase modulation associated to the rotation (or roll) of the projectile. The rotation modulation signal has characteristics of cosine and periodicity which can be used to analyse the rotation speed of a projectile by tracking and demodulating this signal [3].

There are some single satellite receiver antenna-based methods of rotation parameter measurement of a target. In [4], the roll rate and roll angle of a spinning platform is measured by a single GPS antenna. In [5], Doctor Li proposed a method to determinate the rolling angle of a spinning vehicle according to single antenna GPS carrier amplitude. In addition, Rockwell Collins also released its Advance spinning vehicle navigation receiver (ASVN) in 2003.

Unfortunately, few of the existing methods can be used to measure the rotation angle of a projectile perfectly. In this paper, the authors propose a single satellite receiver antenna-based method to measure the rotation parameters of a projectile by design a tracking loop of the satellite receiver according to the characteristics of the rolling amplitude modulation signal. Compared to traditional receiver, a rolling demodulation processor is added to measure the rolling speed and rolling angle of the projectile. The main contributions of this paper are a reference for the receiver tracking loop design of military artillery and a method to measure the rotation speed of a rotating object.

2 Design of the Tracking Loop for Receivers Under the Situation of Rotation

First, the characteristics of amplitude modulated signal received by single antenna rotation are analysed, then the receiver tracking loop is researched, and the rolling demodulation loop is added to the traditional artillery-borne receiver to track and estimate the rolling speed of the projectile.

2.1 Model of Single Antenna Receiving Signal Under the Situation of Rotation

Figure 1 shows the schematic of single antenna receiving satellite signal, which can be used to analyse the effect of rotation.

The antenna is mounted on the side of the projectile, and the antenna coordinate system is defined as $o - xyz$. The ox axis is the rotating shaft of the projectile and oy axis points to the antenna phase center. The ox, oy, oz axis is mutually perpendicular following to the right hand rule. S_i is the satellite, and a straight line m is the projection of the satellite signal on the YOZ plane. α is the angle between the incident signal and the ox axis, which is called the pitch angle. β is the angle between the line m and the oz axis and called the azimuth.

As shown in Fig. 1, we can make a conclusion that the power density $P_{re}(\alpha, \beta)$ of the received signal can be expressed as [6]:

$$P_{re}(\alpha, \beta) = \left(\frac{\lambda}{4\pi r}\right)^2 G_1(\alpha, \beta) G_i P_{in} \tag{1}$$

where r denotes inter antenna distance, G_i denotes antenna gain, $G_1(\alpha, \beta)$ denotes gain direction function of receiving antenna, P_{in} denotes transmit antenna power, $P_{re}(\alpha, \beta)$ denotes receiving signal power, λ denotes signal wavelength.

From the formula (1), we can see that the intensity of the received signal $P_{re}(\alpha, \beta)$ is positively correlated with the gain of the receiving antenna $G_1(\alpha, \beta)$ knowing the parameters such as the transmitted signal power P_{in} and the transmit gain G_i. Taking the microstrip antenna as an example, the antenna pattern is related to the pitch angle and has nothing to do with the azimuth angle. Antenna electric field pattern function is:

$$G_1(\alpha) = \sin\alpha \cos\left(\frac{\theta L}{2\sin\alpha}\right) \tag{2}$$

where θ is the wave number, $\theta = 2\pi/\lambda$, L is the effective length of antenna.

It can be seen that the received signal power presents a periodic sine and cosine rule change during the rotation of the antenna, and the modulation characteristic of the signal is proportional to the change of the incident signal to the pitch angle.

If the antenna's direction is symmetry about the antenna axis, then the amplitude of the GPS signal is the largest. The amplitude of the signal will change as the

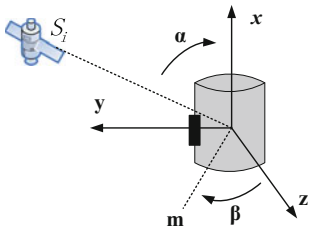


Fig. 1 Schematic diagram of receiving satellite signal by single antenna

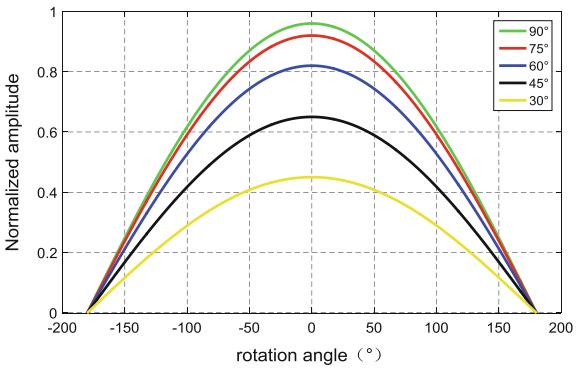


Fig. 2 The relationship between antenna rotation angle and the amplitude of received satellite signal

carrier rolling, and be minimized when the position of the antenna is just back to a satellite [7]. Figure 2 is the influence of single antenna rotation on the received signal amplitude; we can see that the satellite carrier signal amplitude will change as well as the rolling angle change. Signal modulation characteristic increases with the increase of the pitch angle of incident signal [8].

2.2 Design of Improved Tracking Loop

According to the characteristics of the amplitude modulation signal received under rotation condition, this paper improved the design of tracking loop. Upper the dotted line is the part of the traditional tracking loop, and below the dotted line is the part of the new rotation tracking loop. The input of rolling loop is the traditional tracking loop integral cleaner results, after DC integral rotation extraction the power signal, and then generating rolling loop I/Q signal. Rolling loop is consisting of rolling phase discriminator, loop filter and rolling NCO (Fig. 3).

In high dynamic environment, in order to improve the speed of carrier phase tracking, carrier loop use the second-order frequency locked loop (FLL) assisting the third-order phase locked loop (PLL). When the local carrier and the input signal frequency are in large difference, FLL help system implement fast frequency traction. Besides, the carrier tracking loop assists the code ring in real-time, to eliminate most of the dynamics [4, 9, 10]. Therefore, the order of the code loop does not need to be too high. The traditional tracking loop adopts the second order loop.

The amplitude modulation signal can be expressed as

$$u_i(t) = A\cos(\omega_r t + \theta_r) + \sin(\omega_i t + \theta_i) \tag{3}$$

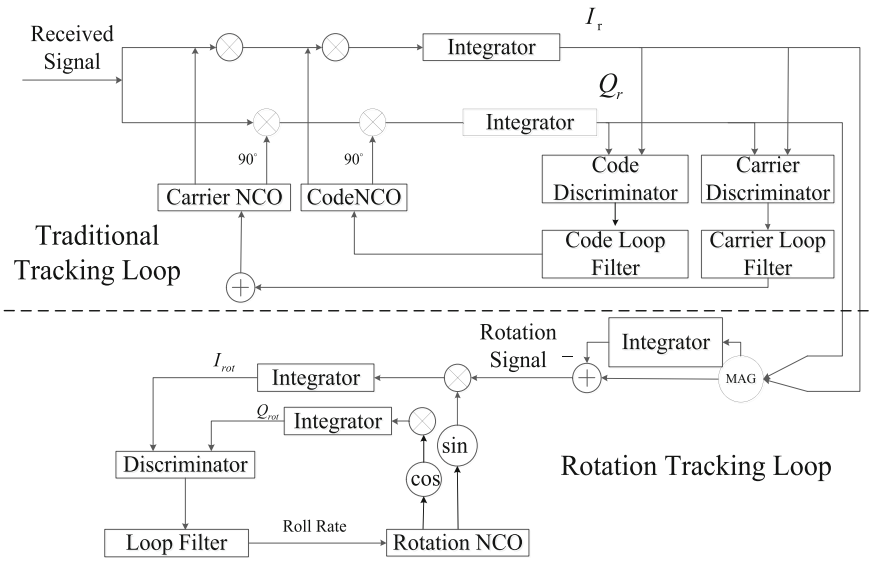


Fig. 3 Block diagram of tracking loop

$$u_{os}(t) = A_0 \sin(\omega_0 t + \theta_0) \tag{4}$$

$$u_{oc}(t) = A_0 \cos(\omega_0 t + \theta_0) \tag{5}$$

where $u_i(t)$ is received signal, A is the amplitude of signal, $\mathrm{w_i}$ is the frequency of signal, θ_i is initial phase of signal, $u_{os}(t)$, $u_{oc}(t)$ are local replicated branch signal and quadrature branch signal, A_0 is the amplitude of local replicated signal, ω_0 is the frequency of replication signal, θ_r is the initial phase of replication signal, ω_r is the modulation frequency of carrier signal, θ_r is the rolling phase. By mixing the received signal, and make integral operation with time T, we can get the results of coherent integral, where $K = A * A_0$.

$$I_r = \int_{t_1}^{t_1+T} -\frac{1}{2} * K \cos(\omega_r t + \theta_r)(\cos((\omega_i + \omega_0)t + \theta_i + \theta_0) - \cos((\omega_i - \omega_0)t + \theta_i - \theta_0))dt \tag{6}$$

$$Q_r = \int_{t_1}^{t_1+T} -\frac{1}{2} * K \cos(\omega_r t + \theta_r)(\sin((\omega_i + \omega_0)t + \theta_i + \theta_0) + \sin((\omega_i - \omega_0)t + \theta_i - \theta_0))dt \tag{7}$$

Through further calculation:

$$I_r \approx \frac{K}{\frac{1}{2}\omega_e T}\sin\left(\frac{1}{2}\omega_e T + \omega_r t\right)\cos\left((\omega_e + \omega_r)\left(t_1 + \frac{1}{2}T\right) + \theta_e\right) \tag{8}$$

$$Q_r \approx \frac{K}{\frac{1}{2}\omega_e T}\sin\left(\frac{1}{2}\omega_e T + \omega_r t\right)\sin\left((\omega_e + \omega_r)\left(t_1 + \frac{1}{2}T\right) + \theta_e\right) \tag{9}$$

$$\sqrt{I_r^2 + Q_r^2} = \frac{K\sqrt{2}}{\frac{1}{2}\omega_e T}\sin\left(\frac{1}{2}\omega_e T + \omega_r t\right) \approx \frac{K\sqrt{2}}{\frac{1}{2}\omega_e T}\sin(\omega_r t) \tag{10}$$

The I/Q integral power of the correlator is $\sqrt{I_r^2 + Q_r^2}$, however, because of the influence of rolling, the amplitude is no longer a fixed value. It reflects the speed and angle of the rotation. The power of integral cleaner can be used to extract so as to get power signal whose phase value is rolling angle information [11].

The initial phase of the local phase is calculated by aiding from the external data, rolling phase is mixed with real input signal after sine transform and cosine transform. And then it will be integrated to get I_{rot}, Q_{rot}. According to the principle of phase locked loop, rolling tracking loop tracks the rolling angle information to achieve closed-loop control. The rolling phase of the rolling PLL output is the signal generated by local phase mixer. With the loop tracking the input signal, the PLL makes the local output signal share the same frequency and phase with it, and then realizes the locking of the loop, and finally outputs the local rolling phase ϕ_{locked} and rolling rate $\hat{\phi}_{locked}$.

According to the characteristics of IQ signal, the phase detector is designed to two quadrant inverse tangent function of the phase detector which performs best both at high SNR and low SNR. In addition, output of the phase slope is the line and it has nothing to do with the amplitude of the signal. The phase discriminator output of the two quadrant inverse tangent function is:

$$\phi_e = \arctan(Q_{rot}, I_{rot}) \tag{11}$$

where I_{rot}, Q_{rot} denote the output results of the phase integrator and the orthogonal integrator.

Rolling tracking loop uses second-order PLL to track rolling phase, and the loop filter is second-order filter. The below figure is the design diagram of the filter. The loop filter input is phase error of discriminator, and the output is NCO frequency word. k_1, k_2 are the loop filter coefficient (Fig. 4).

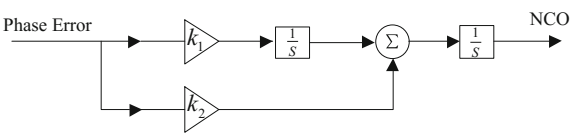


Fig. 4 Principle of tracking loop

$$k_1 = \omega_n^2 \tag{12}$$

$$k_2 = a\omega_n \tag{13}$$

$$\omega_n = \frac{8B_l\xi}{1+4\xi^2} \tag{14}$$

In the formula, B_l is the loop bandwidth, ξ is damping coefficient, ω_n is inherent frequency of the system. The design of rolling NCO is similar to the traditional carrier NCO, which is no longer described in detail.

3 Performance Analysis of the Tracking Loop

3.1 Simulation Design

The rolling modulated satellite signal is generated by simulator, and the new tracking loop is applied to adjust the loop coefficient to adapt to different speed measurement range. In the simulation, the pitching angle is set to 45°, the frequency of intermediate frequency is 9.55 MHz, the sampling frequency is 38.192 MHz, and the speed is designed to be 20 and 100 r/s. The traditional tracking loop integral time of the receiver is 1 ms. As shown in the modulation results under different rolling rates, the results of I/Q integral generated under rotation condition show modulation corresponding to the speed (Fig. 5).

The original rolling satellite signal is processed by software receiver and the generated IQ integral value is used as the input of the rolling demodulation module.

3.2 Verification of Result

The I/Q result is used as the input of the phase-locked loop, and different loop parameters are designed to tracking the speed of the projectile. Figure 6 shows the result of the satellite signal demodulation under different rolling rate. The abscissa axis represents the processing time and the longitudinal axis is the rate of rolling.

From the figure, we can know that the tracking loop designed in this paper can demodulate the rolling rate of the missile in real time. Through Error calculation formula (15) and (16), it can be calculated that the tracking error can be controlled within 5 and 3%, which can be well meet the requirements of most practical applications.

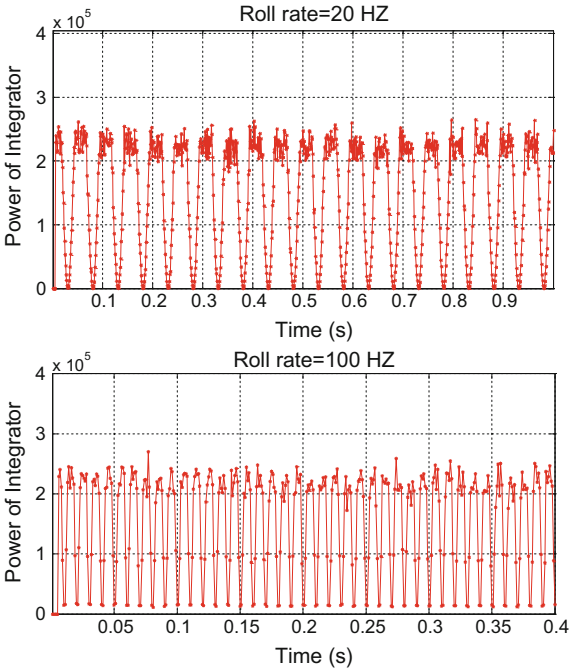


Fig. 5 Power output by integrator

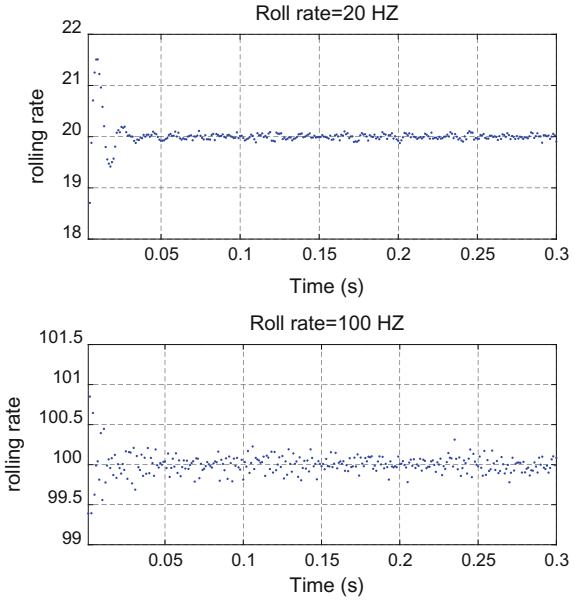


Fig. 6 Measurement of roll rate

$$\xi_i = \frac{|\hat{n}_i - n_i|}{n_i} \quad (i = 1, 2, 3 \ldots m) \tag{15}$$

$$\overline{\xi} = \frac{1}{m} \sum_{i=1}^{m} \xi_i \tag{16}$$

where n_i and $\hat{n}_i$ stand for the ground truth and estimation of rotation speed of projectile, respectively.

4 Conclusion

The projectile trajectory correction requires real-time information of rotation speed. According to the amplitude characteristics of the received signal, a real-time tracking loop of the satellite receiver which consists of a rotation tracking loop and a traditional tracking loop is designed to track satellite signal. Rotation tracking loop uses the power output by the integral cleaner of the traditional tracking loop as its input, and at the same time uses two order phase locked loop to track satellite signal. By this way, projectile speed can be analysed and output in real-time.

Simulations under different rotation speed of projectile with single antenna satellite receiver were carried out. The mean values of relative error of projectile rotation speed are 5 and 3% when the rotation speed of the projectile is 20 and 100 r/s. The improved tracking loop can achieve high accuracy positioning and effectively solve the rotation speed of the projectile. Results show that the performance of our tracking loop is satisfactory.

As for application, our method can be used to measure the rotation speed and rotation angle of the projectile whose rotation speed varies from 20 to 100 r/s. It can solve the problems of the rotation speed measurement of a rotating object. The main contributions of this paper are a reference for the receiver tracking loop design of military artillery and a method to measure the rotation speed of a rotating object.

References

1. Seidel W, Guischard F (2000) Method for autonomous guidance of a spin-stabilized artillery projectile and autonomously guided artillery projectile for realizing this method: US, US 6135387 A[P]
2. Doty JH, Mc Graw GA (2003) Spinning-vehicle navigation using apparent modulation of navigational signals: U.S. Patent 6,520,448. 2003-2-18
3. Liu XD, Zhao JX (2009) A generation method of GPS if signal under carrier rotating conditions. J Spacecraft TT&C Technol 2009(28):91–94
4. Luo J, Vander Velde WE, Tseng H, Upadhyay TN (2016) Single antenna GPS measurement of roll rate and roll angle of spinning platform. U.S. Patent: US 9,429,660 B2

5. Li Y, Zhang J, Fei T, Cui N, Shang Y (2014) Single antenna GPS carrier amplitude based rolling angle determination of spinning vehicle. Fire Control Radar Technol 43(3) (Series 169)
6. Shen Q, Wang M, Li DG (2009) Analysis on frequency and phase of received GPS signal in condition of spinning. Trans Beijing Inst Technol 29(1):35–37
7. Zeng GY (2015) Navigation and roll attitude determination technology by non-omnidirectional antenna under rotation condition. Beijing Institute of Technology
8. Fenton PC, Kunysz W (2000) Method and apparatus using GPS to Determine Position and Attitude of a rotation vehicle. U.S. Patent: 6128557,200,8,3
9. Lan YF, Guo CJ, Cao JS (2014) GPS tracking loop design for the spinning missiles. Sci Technol Eng
10. Alexander SB, Redhead R (2014) Systems and methods for tracking power modulation. U.S. Patent: US 8,711,035 B1
11. Li Y, Zhang J, Fei T, et al (2015) Single antenna GPS carrier amplitude based rolling angle determination. In: Guidance, navigation and control conference. IEEE, pp 206–211

A Stable DOA Tracking Method Using FDPM and Kalman Filter

Gangqiang Guan, Deyong Xian, Zhou Yi and Weihua Xie

Abstract To solve the problems that eigenvalue decomposition of the covariance matrix is computationally intensive and the target moving can cause a spatial spectrum spread in the GNSS (global navigation satellite system) interference detection and localization system using antenna array, in this paper we propose a low-complexity GNSS signal DOA tracking algorithm using a Kalman filter implementation. This method firstly utilizes an improved FDPM (fast data projection method) for noise subspace tracking, and sets the DOA estimation predicted by Kalman filter as the initial value of Newtown iteration, then using the iteration result as a new DOA observation the sate vector of the Kalman filter is updated and calibrated, therefore a stable and effective DOA tracking loop is achieved. Simulation results show that using the method the performance of DOA tracking can be significantly improved while the computational burden is reduced simultaneously.

Keywords Subspace tracking · DOA tracking · Newton iteration Kalman filter

1 Introduction

Interference detection and location is crucial for high reliability GNSS applications in complex electromagnetic environments. When the interference is detected, how to achieve DOA estimation and DOA tracking of the moving target efficiently and stably is still worth further studying. Until now there are a large number of literatures on this subject, such as multiple signal classification (MUSIC) algorithm [1] and estimation of signal parameters via rotational invariance technique (ESPRIT) algorithm [1, 2], etc. Although super resolution can be achieved using these methods, they are difficult to implement in engineering for their heavy computation

G. Guan (✉) · D. Xian · Z. Yi · W. Xie
Beijing Satellite Navigation Center, Beiqing Road. 22, 100094 Beijing, China
e-mail: closetoqiang@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_39

burden. Moreover, when the target signal is rapidly maneuvered, multiple received data snapshot to estimate its covariance matrix may cause spatial spectral spread, which seriously reduces the DOA estimation performance [3]. Some low-complexity subspace tracking algorithms such as Projection Approximate Subspace (PAST) algorithm and its improved algorithm [4–10], data projection algorithm (DPM) and so on can significantly reduce the complexity of subspace real-time computation [11]. The convergence rate of these algorithms is not suitable for DOA tracking of fast maneuvering sources, and it also faces the problem of spatial spectrum extension. In order to solve these problems, this paper presents a combined Kalman filter and an improved fast data projection algorithm (Wang-FDPM). Finally, the performance of this algorithm is verified through several typical application scenarios.

2 Signal Model for DOA Tracking Using Kalman Filter

As shown in Fig. 1, the DOA tracking model of the Kalman filter implementation structure is mainly divided into the prediction and correction processes. The DOA observations used in the correction process are provided by the Newton iteration algorithm, while the subspace estimation required by the initialized Newton iteration algorithm through the Wang-FDPM algorithm. The initial value of the required signal DOA is obtained by the prediction process [8–10]. The Kalman filter can make the corrected DOA state vector have the least mean square error by synthesizing and balancing the predicted and observed DOA signals.

The DOA state vector in the Kalman filter is defined as $\boldsymbol{y}_k = \left[\theta_k, \dot{\theta}_k, \ddot{\theta}_k\right]^{\mathrm{T}}$ where θ_k, $\dot{\theta}_k$ and $\ddot{\theta}_k$ are signal DOA, angular velocity and angular acceleration at time k respectively. Then the state equation can be written as

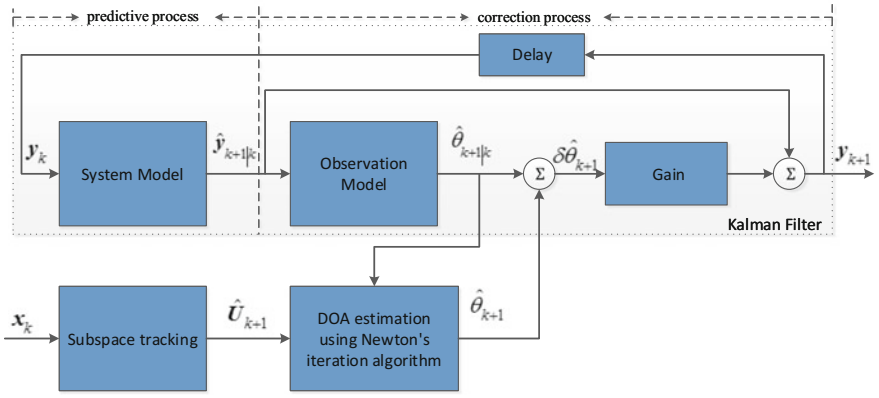


Fig. 1 DOA tracking signal model based on Kalman filter

$$\boldsymbol{y}_{k+1} = \boldsymbol{F}\boldsymbol{y}_k + \boldsymbol{w}_k \tag{1}$$

where $\boldsymbol{w}_k$ is the process noise vector with zero mean and variance $\boldsymbol{Q}_k$, which represents the random noise error generated within the system [8, 9]. The state transition matrix is

$$\boldsymbol{F} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where T is the update cycle. The observation equation of Kalman filter is

$$\theta_{k+1} = \boldsymbol{h}\boldsymbol{y}_{k+1} + v_{k+1} \tag{3}$$

In the above formula, h = [1 0 0] is the observation matrix, v_{k+1} is the measurement error and observation noise with 0 mean at the k + 1 time.

Let $\hat{\boldsymbol{y}}_{k|k}$ and $\boldsymbol{P}_{k|k}$ be the state vector estimate time and its error covariance matrix at the k, respectively. Then the predictive process of the Kalman filter which predicts the state value at the next moment using the state equation based on the state estimate of the previous moment [12] can be expressed as

$$\begin{aligned} \hat{\boldsymbol{y}}_{k+1|k} &= \boldsymbol{F}\hat{\boldsymbol{y}}_{k|k} \\ \boldsymbol{P}_{k+1|k} &= \boldsymbol{F}\boldsymbol{P}_{k|k}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{Q}_k \end{aligned} \tag{4}$$

where $\hat{\boldsymbol{y}}_{k+1|k}$, $\boldsymbol{P}_{k+1|k}$ represents the estimated value of the state vector and the error covariance matrix at time k + 1. The correction process uses the DOA measurements to update the prior estimate and the error covariance matrix [12] predicted by the previous step and can be written as

$$\begin{aligned} \boldsymbol{K}_{k+1} &= \boldsymbol{P}_{k+1|k}\boldsymbol{h}^{\mathrm{T}}\left(\boldsymbol{h}\boldsymbol{P}_{k+1|k}\boldsymbol{h}^{\mathrm{T}} + \sigma^2_{v,k+1}\right)^{-1} \\ \hat{\boldsymbol{y}}_{k+1|k+1} &= \hat{\boldsymbol{y}}_{k+1|k} + \boldsymbol{K}_{k+1}\left(\hat{\theta}_{k+1} - \boldsymbol{h}\hat{\boldsymbol{y}}_{k+1|k}\right) \\ \boldsymbol{P}_{k+1|k+1} &= (\boldsymbol{I} - \boldsymbol{K}_{k+1}\boldsymbol{h})\boldsymbol{P}_{k+1|k} \end{aligned} \tag{5}$$

In the above formula, $\boldsymbol{K}_{k+1}$ is the gain of Kalman filter, $\hat{\theta}_{k+1}$ is the DOA observation calculated by the subspace tracking algorithm and Newton iteration algorithm at the k + 1 time.

3 DOA Observation Estimation Based on Low Complexity Subspace Tracking Algorithm

3.1 Newton's Iteration Algorithm

The Kalman filter for DOA tracking needs subspace tracking algorithm and Newton iteration algorithm to provide accurate observations $\hat{\theta}_{k+1}$. Otherwise the Kalman filter gain matrix and state vector estimation parameters will not reach their optimal value, or even cause the filter to diverge. The objective function when used MUSIC algorithm for DOA estimation is

$$f(\theta) = \boldsymbol{a}^{\mathrm{H}}(\theta)\boldsymbol{\Pi}\boldsymbol{a}(\theta) \tag{6}$$

where $\boldsymbol{a}(\theta)$ is the steering vector of the array and $\boldsymbol{\Pi}$ is the projection matrix of the noise subspace. Provided that $f'(\hat{\theta}) = 0$ the optimal DOA estimation can be achieved by minimizing objective function. As a linearization method, Newton's iterative algorithm can be used to solve the DOA estimation that satisfies the derivative constraints to simplify the computation of the objective function search process. At the first moment, the Newton's iterative formula is [11]

$$\hat{\theta}_{k+1} = \hat{\theta}_k - f'\left(\hat{\theta}_k\right) \Big/ f''\left(\hat{\theta}_k\right) \tag{7}$$

where $f'(\hat{\theta}_k)$ and $f''(\hat{\theta}_k)$ are the first derivative and the second derivative of the objective function f at $\hat{\theta}_k$ respectively. Supposing the signal subspace at the time k is $\hat{\boldsymbol{U}}_{s,k}$, the corresponding noise projection matrix estimated is $\hat{\boldsymbol{\Pi}}_k = \boldsymbol{I} - \hat{\boldsymbol{U}}_{s,k}\hat{\boldsymbol{U}}_{s,k}^{\boldsymbol{H}}$, then

$$f'\left(\hat{\theta}_k\right) = \mathrm{Re}\left[\boldsymbol{d}^{H}(\theta)\hat{\boldsymbol{\Pi}}_k\boldsymbol{a}(\theta)\right]\Big|_{\theta=\hat{\theta}_k} \tag{8}$$

$$f''\left(\hat{\theta}_k\right) = \mathrm{Re}\left[\boldsymbol{d}'^{H}(\theta)\hat{\boldsymbol{\Pi}}_k\boldsymbol{a}(\theta)\right] + \boldsymbol{d}^{H}(\theta)\hat{\boldsymbol{\Pi}}_k\boldsymbol{d}(\theta)\Big|_{\theta=\hat{\theta}_k} \tag{9}$$

where $\boldsymbol{d}(\theta) = \frac{\partial \boldsymbol{a}(\theta)}{\partial \theta}$, $\boldsymbol{d}'(\theta) = \frac{\partial^2 \boldsymbol{a}(\theta)}{\partial^2 \theta}$. Although Newton's iterative method has higher convergence rate, it requires more accurate initial conditions. In order to ensure the convergence of the iterative process, the updated signal subspace $\hat{\boldsymbol{U}}_{s,k+1}$ and the a priori DOA estimation $\hat{\theta}_{k+1|k} = \boldsymbol{h}\hat{\boldsymbol{y}}_{k+1|k}$ of the Kalman filter prediction process are used as the iterative initial values. In this case, the calculation formula of the DOA observation using the Newton iteration can be written as

$$\hat{\theta}_{k+1} = \hat{\theta}_{k+1|k} - \frac{\mathrm{Re}\left[\boldsymbol{d}^{\boldsymbol{H}}(\theta)\hat{\boldsymbol{\Pi}}_{k+1}\boldsymbol{a}(\theta)\right]}{\mathrm{Re}\left[\boldsymbol{d}'^{\boldsymbol{H}}(\theta)\hat{\boldsymbol{\Pi}}_{k+1}\boldsymbol{a}(\theta)\right] + \boldsymbol{d}^{\boldsymbol{H}}(\theta)\hat{\boldsymbol{\Pi}}_{k+1}\boldsymbol{d}(\theta)}\Big|_{\theta=\hat{\theta}_{k+1|k}} \tag{10}$$

3.2 Subspace Tracking Algorithm Based on Wang-FDPM

Using the receive data vector to update signal or noise subspace estimation in real time can effectively reduce the computational complexity of subspaces. In particular, the low-complexity subspace tracking algorithm has become the hot topic in array signal processing, and many improved algorithms with excellent performance appear. PAST algorithm, PASTd algorithm can effectively track the signal subspace vector, but they are only suitable for signal subspace tracking with a computational complexity of $3NL + O(L^2)$, where N is the number of array elements, L is the number of signals. While the PAST algorithm can not guarantee the standard orthogonalization of subspace basis, the improved OPAST algorithm orthogonalizes the subspace in each iteration, which guarantees the orthogonality of subspace and improves the convergence performance of the algorithm. However, The computational complexity is increased to $4NL + O(L^2)$ [9, 10].

In order to further improve the robustness and real-time performance of the DOA observations in the Kalman filter, as shown in Table 1, this section adopts the improved data projection algorithm proposed in [14], i.e. the Wang-FDPM algorithm, for subspace tracking. The algorithm can track the signal subspace or the noise subspace simultaneously, and is insensitive to the rounding error and has numerical stability. The computational complexity is 3NL + O (N) [11, 14].

Therefore, the noise subspace projection matrix after subspace tracking is

$$\hat{\boldsymbol{\Pi}}_{k+1} = \hat{\boldsymbol{U}}_{n,k+1}\hat{\boldsymbol{U}}^{\boldsymbol{H}}_{n,k+1} = \boldsymbol{I} - \hat{\boldsymbol{U}}_{s,k+1}\hat{\boldsymbol{U}}^{\boldsymbol{H}}_{s,k+1} \tag{11}$$

where $\hat{\boldsymbol{U}}_{s,k+1}$ and $\hat{\boldsymbol{U}}_{n,k+1}$ are signal subspace and noise subspace estimation respectively.

4 Simulation

In this section, the performance of the DOA tracking algorithm based on Kalman filter is studied using a uniform linear array with 12 omnidirectional elements spaced at half the carrier wavelength. The simulation process temporarily considers

Table 1 Wang-FDPM algorithm

Suppose that the subspace estimation at the current moment is U_k and the data vector is x_k
1. $\boldsymbol{r}(k) = \boldsymbol{U}_k^{\mathrm{H}}\boldsymbol{x}(k)$
2. $\boldsymbol{g}(k) = \boldsymbol{U}_k^{\mathrm{H}}\boldsymbol{r}(k)$
3. $\boldsymbol{w}(k) = \boldsymbol{g}(k)\|\boldsymbol{r}(k)\|_2^{-1} + \alpha\boldsymbol{x}(k)\|\boldsymbol{r}(k)\|_2$
4. $\boldsymbol{q}(k) = \boldsymbol{w}(k)\|\boldsymbol{w}(k)\|_2^{-1} - \boldsymbol{g}(k)\|\boldsymbol{r}(k)\|_2^{-1}$
5. $\boldsymbol{U}_{k+1} = \boldsymbol{U}_k + \boldsymbol{q}(k)\boldsymbol{r}^{\mathrm{H}}(k)\|\boldsymbol{r}(k)\|_2^{-1}$

where α is the step factor, $\alpha > 0$ is set to track the signal subspace, otherwise to track the noise subspace.

only the DOA tracking of a single interference source, assuming that the signals DOA of two consecutive moments are known as $\hat{\theta}_{-1}$ and $\hat{\theta}_0$, then the Kalman filter system state vector can be initialized to

$$\hat{\boldsymbol{y}}_{0|\,0} = \left[\hat{\theta}_0, T^{-1}\left(\hat{\theta}_0 - \hat{\theta}_{-1}\right), 0\right]^{\boldsymbol{T}} \tag{12}$$

The corresponding initial error covariance matrix of the state vector $\boldsymbol{P}_{0|\,0}$ and the process noise covariance matrix $\boldsymbol{Q}_k$ are initialized by

$$\begin{aligned} \boldsymbol{P}_{0\,|\,0} &= \begin{bmatrix} 1 & T^{-1} & 0 \\ T^{-1} & 2T^{-2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \sigma_v^2 \\ \boldsymbol{Q}_k &= \begin{bmatrix} T^4/4 & T^3/2 & T^2/2 \\ T^3/2 & T^2 & T \\ T^2/2 & T & 1 \end{bmatrix} \sigma_w^2 \end{aligned} \tag{13}$$

Assuming the noise variance $\sigma_v^2 = 0.1$, $\sigma_w^2 = 0.0001$, the initial signal DOA is 0°. Moreover, the rate of signal DOA change is 1° per second due to the target moving, the sampling interval of the array receiving data is 20 ms, and the number of the receiving data snapshots is 60. The step factor of Wang-FDPM algorithm in the simulation process is set as $\alpha = 0.01$ and the state vector of the Kalman filter is $\hat{\boldsymbol{y}}_{0|\,0} = [\theta_0, 0, 0]^{\boldsymbol{T}}$.

The simulation time is set to 60 s and the number of snapshots of received data per second is 10. The performance of the DOA tracking algorithm is studied by using two simulation scenarios and the DOA of two signals incident on the array in the simulation scenario 1 is set at (20°, 80°)–(80°, 20°). The DOA of the third signal is sinusoidal with an initial value of 30°. The DOA of signal 1 in scenario 2 remains unchanged at 18°, but the Kalman filter has an initial estimate of 16° and signal 2 has a DOA of 16° with a jump of 3° at 15 and 20 s. Simulation results show that the DOA tracking results based on the Wang-FDPM subspace tracking algorithm are shown in Figs. 2 and 3, respectively. It can be seen from the figure that the algorithm can realize the tracking of signals of different fast trajectories DOA is stable, and its DOA can be tracked when the signal DOA changes or the initial DOA has error. Figure 4 shows the RMSE of DOA tracking Kalman filter when using different subspace tracking algorithms. It can be seen from the figure that the combination of the Wang-FDPM algorithm and the Kalman filter can improve the DOA tracking performance at a low SNR.

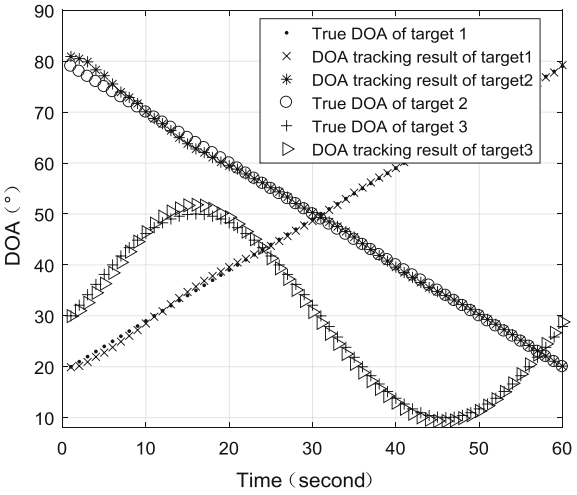


Fig. 2 DOA tracking results using Wang-FDPM method

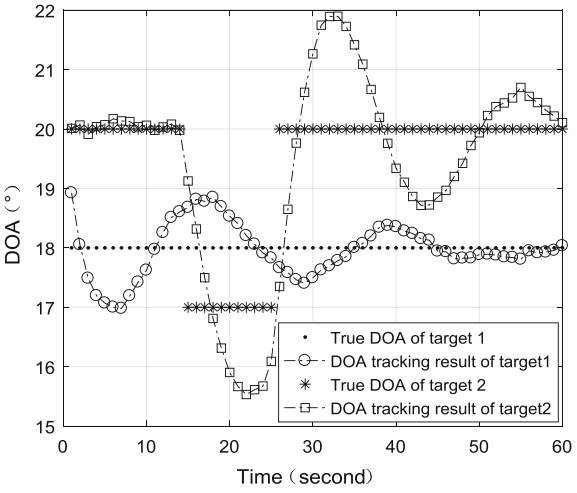


Fig. 3 DOA tracking result using Wang-FDPM method when a DOA jump appears

It can be seen from the above simulation results that the DOA tracking technology based on Kalman filter proposed in this paper can achieve the stable tracking of signal DOA and the tracking error of DOA using Wang-FDPM subspace tracking algorithm is smaller than OPAST algorithm, The algorithm can reduce the amount of computation while maintaining subspace tracking orthogonality.

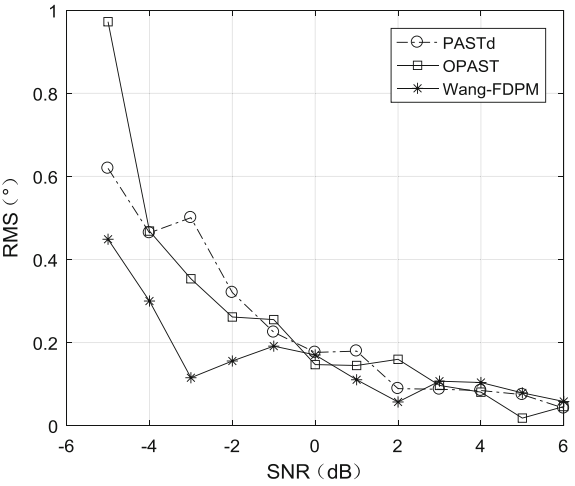


Fig. 4 RMS curve of DOA tracking result when using different subspace tracking methods

5 Conclusion

In the engineering application of array antenna to detect and locate GNSS interference sources, DOA estimation obtained by subspace tracking technology and Newton iteration algorithm is taken as the observed value of DOA tracking Kalman filter. This method can realize the DOA tracking stably and efficiently. The simulation verification process in this paper compares the DOA tracking performance of two commonly used low-complexity subspace tracking algorithms, Wang-FDPM and OPAST. The simulation results show the effectiveness of the proposed method.

References

1. Wang YL et al (2004) Spatial spectrum estimation theory and algorithm. Tsinghua University Press, Beijing, pp 2–13
2. Cheng Q, Lei H, So HC (2014) Computationally efficient ESPRIT algorithm for direction-of-arrival estimation based on Nystrom method. Signal Process 94:74–80
3. Jian FG, Chan SC, Wei PZ et al (2013) Joint DOA estimation and source signal tracking with Kalman filtering and regularized QRD RLS algorithm. IEEE Trans Circ and Syst 60(1):46–50
4. Abed-Meraim K, Chkeif A, Hua Y (2000) Fast orthonormal PAST algorithm. IEEE Signal Process Lett 7(3):60–62
5. Bartelmaos S, Abed-Meraim K (2008) Fast principal component extraction using givens rotations. IEEE Signal Process Lett 15:369–372
6. Liu C, Wang G, Xin J, Wang J, Zheng N, Sano A (2012) Low complexity subpace-based two-dimensional direction-of-arrivals tracking of multiple targets. In: The 11th IEEE international conference on signal processing, Beijing, China, 21–25 Oct 2012
7. Chan SC, Wu HC, Tsui KM (2012) Robust recursive eigendecomposition and subspace-based algorithms with application to fault detection in wireless sensor networks. IEEE Trans Instrum Meas 61(6):1703–1717

8. Javier SA, Sylvie M (1999) An efficient PASTd-algorithm implantation for multiple direction of arrival tracking. IEEE Trans Signal Process 47(8):2321–2324
9. Amir V, Mahmood K, Omid O (2007) An efficient algorithm for multiple direction of arrival tracking based on the constrained projection approximation approach and Kalman filter. In: IEEE international conference on acoustics, speech, and signal processing, Hawaii, USA, 15–20 Apr 2007
10. Wang C (2011) Research on key adaptive interference suppression in satellite navigation receiver. Xidian University, Xi'an
11. Yuan XD (2012) Research on key technologies of GNSS interference suppression in direct data domain. National University of Defense Technology, Changsha
12. Xie G (2009) Principles of GPS and receiver design. Publishing House of Electronics Industry, Beijing, pp 127–150
13. Ouyang S, Ching PC, Lee T (2003) Robust adaptive quasi-Newtown algorithms for eigensubspace estimation. IEE Proc Vis Image Signal Process 150(5):321–330
14. Rong W, Minli Y, Daoming Z, Hongxing Z (2012) A novel orthonormalization matrix based fast and stable DPM algorithm for principal and minor subspace tracking. IEEE Trans Signal Process 60(1):466–472

GNSS Spoofing Detection with Single Moving Antenna Based on the Correlation of Satellite Transmit Time Residual

Zhiyuan Chen, Hong Li and Mingquan Lu

Abstract GNSS spoofing detection method has been widely discussed recently. Most of the detection methods assume that spoofing signal is transmitted by a single source rather than distributed sources. The detection methods mainly fall into three categories: multi-antenna, moving antenna and cloud based method. The existing moving antenna method is based on monitoring the variation of measurements. In this paper, a spoofing detection method based on the pairwise correlations of satellite transmit time residual with a single moving antenna is proposed. Authentic GNSS signals with various elevation and azimuth angles have different propagation paths. However, it should be noted that counterfeit GNSS signals emitted by a single source have the same propagation path from the spoofing emitter to the receiver antenna. Consequently, the correlation results between the spoofing measurements with the same delay are higher than those of the authentic ones. The detection metric to measure the correlation is offered in this paper. The detection metric is the mean value of all the pairwise correlation coefficients. This method combines the PVT solution tracking level and the satellite transmit time in navigation level. Therefore, this spoofing detection method can cover a wide range of situations.

Keywords Spoofing detection · Moving antenna · Transmit time Correlation coefficient

1 Introduction

Global Navigation Satellite Systems (GNSS) provide position and timing for numerous critical infrastructures. The low power of GNSS signals means that they are vulnerable. Therefore, defending against jamming and spoofing becomes an essential element. Jamming declines accuracy or denies service. What's more,

Z. Chen · H. Li (✉) · M. Lu
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
e-mail: lihongee@tsinghua.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_40

Spoofing counterfeits position, velocity, and timing (PVT) solution. It is a hazard that generally receivers cannot realize they are counterfeited. The spoofing signals may be generated by a simulator or repeater. Limited by cost and technical feasibility, spoofing is usually transmitted by a single source rather than distributed sources [1].

The single source spoofing detection techniques utilize the spatial feature between different satellite signals. Research on various approaches to acquire spatial feature mainly includes multi-antenna, moving antenna and cloud based method [2]. The multi-antenna method needs to use antenna array and afford parallel processing for each antenna. The receiver checks the consistency of direction of arrival (DOA) between estimation values measured by antenna array and calculation values derived by ephemeris [8]. However, the antenna array is large and expensive and the algorithm is complicated. The moving antenna method involves space-time processing as the receiver is moving. The spoofing detection techniques can be implemented in navigation level and tracking level. In Ref. [5], a spoofing detection technique in navigation level is proposed. This proposed method is based on analyzing the clock bias variation. In addition, the performance relies heavily on the frequency stabilization. Current methods need to learn clock parameters in advance. In Ref. [3], a spoofing detection technique in tracking level is proposed. This proposed method is based on monitoring the amplitude and Doppler correlations of different satellite signals. It does not consider the phase correlation. In Ref. [9], the carrier phase is taken into consideration. The spoofing detection method is effective depending on an intentional high-frequency antenna motion. These tracking level methods cannot handle the situation in which the spoofer fabricates counterfeit signals with random variations of amplitude, position, and velocity. The cloud based method needs communication links to connect the spatially distributed receivers. It works by checking the consistency of the double differences of observations or PVT solutions. In Ref. [6], the double difference of carrier phase is considered. This measurement is time-variant in an authentic scene, but time-invariant in a spoofing scene.

In this paper, we propose a spoofing detection method which can be implemented in a receiver with a single moving antenna based on the correlation of satellite transmit time residual. The satellite transmit time residual is defined as the difference between the measurement and estimation. The measurement is the satellite transmit time. The estimation is calculated by the clock drift and the distance between the satellite and the receiver. The single-source counterfeit GNSS signals mean that the GNSS signals of different PRNs are combined before the spoofing emitter. The fundamental is that counterfeit GNSS signals emitted by a single source have the same extra propagation path from the spoofing emitter to the receiver antenna. Therefore, the correlations of satellite transmit time residual have the distinguishable characteristics between the authentic and spoofing scenes.

The main contributions of this paper are as follows. The correlation of satellite transmit time residual is proposed as a supplement to amplitude and Doppler correlation. A new elimination approach of the satellite transmit time has been proposed. This improvement enables the spoofing detection method to still work

when the spoofer fabricates counterfeit signals with random variations of measurements. In addition, this method can detect the partial channel spoofing for either moving scene or static scene. The partial channel spoofing means that both the authentic signals and the spoofing signals are in the tracking channels. This method can actually be used in a receiver which extracts the satellite transmit time from the code phase. This paper is organized as follows: Sect. 2 details the fundamental of measurements and metrics, Sect. 3 introduces the experiment results. Finally, Sect. 4 summarizes the paper.

2 Fundamental

2.1 Satellite Transmit Time Model

A simplified authentic model for the ith satellite transmit time at receiver time t can be expressed as Eq. (1.1) [4].

$$t^i_{tran}(t) = t - \Delta t_u(t) - \left\{\Delta t^i_r(t) + d^i_r(t)/c\right\} + n^i_r(t) \tag{1.1}$$

where $\Delta t_u(t)$ is the user receiver clock bias. $\Delta t^i_r(t)$ is the ith satellite clock bias. $d^i_r(t)$ is the distance between the ith satellite transmit antenna and the user receiver antenna. c is the speed of light in the atmosphere. $n^i_r(t)$ is the ith satellite transmit time noise. The ith satellite signal troposphere and ionospheric delays are small and change slowly. Therefore, it is reasonable to ignore them.

A similar spoofing model for the ith satellite transmit time can be written as Eq. (1.2).

$$t^i_{tran}(t) = t - \Delta t_u(t) - \left\{\Delta t^i_r(t) + d^i_r(t)/c\right\} - \left\{\Delta t_s(t) + d_s(t)/c\right\} + n^i_r(t) \tag{1.2}$$

where $\Delta t_u(t)$, $d^i_r(t)$ and $n^i_r(t)$ in Eq. (1.2) are the spoofing variables corresponding to each in Eq. (1.1). $\Delta t_s(t)$ is the spoofer clock bias. $d_s(t)$ is the distance between the spoofing transmit antenna and the user receiver antenna.

Authentic GNSS signals with various elevation and azimuth angles have different propagation paths. These features can be simulated in spoofer before broadcasting the signals. However, it should be noted that the single-source counterfeit GNSS signals have the same propagation path from the spoofing emitter to the receiver antenna. As time goes on, the satellite transmit time changes along with not only the clock features but also the satellite and receiver motions in the authentic scene. In addition, the satellite transmit time also changes along with the spoofer and receiver motions in the spoofing scene.

The variable $d^i_r(t)/c$ in Eq. (1.2) is caused by the geometric positions of the ith satellite and user receiver. The variable $d_s(t)/c$ in Eq. (1.2) is introduced by the

relative position between spoofer antenna and receiver antenna. In order to highlight the influence which is introduced by the relative position between the spoofer and receiver, we should eliminate the influence which is introduced by the relative position between the satellite and receiver. In Ref. [9], the observation is the carrier phase difference between two antennas. It works well for two-antenna spoofing detection method, but we need another elimination approach for a single-antenna receiver. The GPS satellite motion is an ellipse with an orbital period of 11 h and 58 min. The variation mapping to the satellite transmit time or carrier phase is a curve with small radian in a short time. In Ref. [10], the observation is the difference between the carrier phase and the 2-order polynomial curve fitting of carrier phase. However, the carrier phase residual changes along with both the relative position of the spoofer and the relative position of the satellite. It means that this approach may not work as the spoofer fabricates counterfeit signals with random variations of position.

2.2 Satellite Transmit Time Residual Model

The authentic model of *i*th satellite transmit time residual is defined and shown as Eq. (1.3).

$$\begin{aligned}\varepsilon^i_{tran}(\tau, t_0) = &\left\{t^i_{tran}(t_0+\tau) - t^i_{tran}(t_0)\right\} \\ &- \left\{\hat{d}^i_r(t_0+\tau) - \hat{d}^i_r(t_0)\right\}/c \\ &- \left\{\Delta\hat{t}_u(t_0) + \Delta\hat{t}^i_r(t_0) + 1\right\}\cdot\tau\end{aligned} \tag{1.3}$$

where t_0 is the reference time. τ is the elapsed time since the reference time. $\hat{d}^i_r(t_0+\tau)$ and $\hat{d}^i_r(t_0)$ are the estimated distances between the *i*th satellite transmit antenna and the user receiver antenna at different time. The real user receiver position is replaced by the estimated value. The estimated receiver position can be obtained from the received pseudo-range using the iterative method. $\Delta\hat{t}_u(t_0)$ is the receiver clock drift. Similarly, the estimated one can be obtained from the received Doppler frequency using the iterative method. $\Delta\hat{t}^i_r(t_0)$ is the *i*th satellite clock drift. The SV Clock Correction a_{f1} can replace the satellite clock drift and it is given in the navigation message [7].

Substituting Eq. (1.1) into Eq. (1.3), we obtain the authentic model, which is given in Eq. (1.4).

$$\begin{aligned}\varepsilon^i_{tran}(\tau,t_0) = &\{\Delta t_u(t_0+\tau)-\Delta t_u(t_0)\}-\Delta\hat{\dot{t}}_u(t_0)\cdot\tau\\ &+\left\{\Delta t^i_r(t_0+\tau)-\Delta t^i_r(t_0)\right\}-\Delta\hat{\dot{t}}^i_r(t_0)\cdot\tau\\ &+\left\{d^i_r(t_0+\tau)-\hat{d}^i_r(t_0+\tau)\right\}/c\\ &+\left\{d^i_r(t_0)-\hat{d}^i_r(t_0)\right\}/c\\ &+\left\{n^i_r(t_0+\tau)-n^i_r(t_0)\right\}\end{aligned} \tag{1.4}$$

The authentic model is simplified as Eq. (1.5)

$$\varepsilon^i_{tran}(\tau,t_0) = n^i_r(\tau,t_0) \tag{1.5}$$

where $n^i_r(\tau,t_0)$ represents the expression on the right side of Eq. (1.4). It is the sum of various errors and noises. It contains the error between the real receiver clock bias and the estimated one, the error between the real ith satellite clock bias and the estimated one, the error between the real distance and the estimated one, and the difference of the two noises at different time. $n^i_r(\tau,t_0)$ is supposed to be additive white Gaussian noise (AWGN) and is distributed as $N\left(0,\sigma^2_i\right)$.

Substituting Eq. (1.2) into Eq. (1.3), we obtain the spoofing model, which is given in Eq. (1.6).

$$\begin{aligned}\varepsilon^i_{tran}(\tau,t_0) = &\{\Delta t_u(t_0+\tau)-\Delta t_u(t_0)\}-\Delta\hat{\dot{t}}_u(t_0)\cdot\tau\\ &+\{\Delta t_s(t_0+\tau)-\Delta t_s(t_0)\}\\ &+\left\{\Delta t^i_r(t_0+\tau)-\Delta t^i_r(t_0)\right\}-\Delta\hat{\dot{t}}^i_r(t_0)\cdot\tau\\ &-\left\{d^i_s(t_0+\tau)-d^i_s(t_0)\right\}/c\\ &+\left\{d^i_r(t_0+\tau)-\hat{d}^i_r(t_0+\tau)\right\}/c\\ &+\left\{d^i_r(t_0)-\hat{d}^i_r(t_0)\right\}/c\\ &+\left\{n^i_r(t_0+\tau)-n^i_r(t_0)\right\}\end{aligned} \tag{1.6}$$

where $\Delta\hat{\dot{t}}_u(t_0)$ is the calculated clock drift. However, it is the estimation of $\{\Delta\dot{t}_u(t_0)+\Delta\dot{t}_s(t_0)\}$. $\Delta\dot{t}_u(t_0)$ is the receiver clock drift and $\Delta\dot{t}_s(t_0)$ is the spoofer clock drift. They cannot be distinguished in the calculated value.

Similarly, the spoofing model can be approximated as Eq. (1.7)

$$\varepsilon^i_{tran}(\tau,t_0) = t_d(\tau,t_0)+n^i_r(\tau,t_0) \tag{1.7}$$

where $t_d(\tau,t_0)$ represents $-\left\{d^i_s(t_0+\tau)-d^i_s(t_0)\right\}/c$. It is the propagation delay of spoofing signals. $n^i_r(\tau,t_0)$ in Eq. (1.6) which is similar to the one in Eq. (1.5), represents the rest of the expression on the right side. It is the sum of various errors and noises. The spoofing signals which are generated by the simulator and repeater both can use the above signal model for the simple spoofing attack mode. The extra

delay keeps common among all satellites for the single-source spoofer. It is impossible to compensate propagation delay in non-cooperative mode. The satellite transmit time residuals are usually random variables with zero-mean in the authentic scene, but with an extra common delay in the spoofing scene. The common delay is changing over time when the spoofer antenna and receiver antenna is moving. The spoofing correlation of different satellite transmit time residual is higher than the authentic one.

2.3 Statistical Detection Metric

The statistical detection metric in Ref. [3] is also suitable for this paper. The pairwise correlation coefficients between the *i*th satellite and the *j*th satellite is defined as Eq. (1.8).

$$\rho_{ij} = \frac{E\left\{\left[\varepsilon_{tran}^{i} - E\left(\varepsilon_{tran}^{i}\right)\right]\left[\varepsilon_{tran}^{j} - E\left(\varepsilon_{tran}^{j}\right)\right]\right\}}{\sqrt{E\left\{\left[\varepsilon_{tran}^{i} - E\left(\varepsilon_{tran}^{i}\right)\right]^{2}\right\}E\left\{\left[\varepsilon_{tran}^{j} - E\left(\varepsilon_{tran}^{j}\right)\right]^{2}\right\}}} \tag{1.8}$$

where $E\{\cdot\}$ is the symbol of expectation and $E\left\{[\cdot - E(\cdot)]^2\right\}$ is the symbol of variance. ε_{tran}^{i} is the random variable of the *i*th satellite transmit time residual. The mean value of all the pairwise correlation coefficients ρ_{ij} is regarded as the detection metric in Eq. (1.9).

$$metric = \frac{\sum_{i=1}^{M}\sum_{j=1}^{i-1}\rho_{ij}}{\sum_{m=1}^{M-1} m} \tag{1.9}$$

The detection metric is obvious and natural to measure the correlation of satellite transmit time residual in this paper. It can detect all channel spoofing scene as well as the partial channel spoofing scene. In next section, the experiments are designed and implemented to validate the proposed metric.

3 Experiment

GPS L1 C/A signals are employed to test and verify the proposed method in the experiments. A real-time hardware receiver with 12 independent channels is adopted. The receiver clock is an oven-controlled crystal oscillator (OCXO) with a frequency stability 10^{-9} s at one second. The down converter (DC) and analog to digital converter (ADC) use the same clock. This condition is demanded by the

carrier phase smoothing. In order to reduce the jitter of PVT solution, the carrier phase smoothing pseudo-range is used to solve PVT solution. The measurement sample rate is 1 Hz and the processing interval is 1 s. The detection processing interval is set as 60 s.

The experiment configurations are depicted in Fig. 1. The experiment equipment is depicted in Fig. 2. In the authentic scene, the signals were received by a randomly moving antenna outdoors. Seven satellite signals including PRN10, PRN15, PRN18, PRN20, PRN21, PRN24, and PRN32 were in tracking channels. In the all channel spoofing scene, the signals were generated by an Agilent E4438C simulator and received by a randomly moving antenna indoors. Seven satellite signals including PRN02, PRN04, PRN05, PRN13, PRN23, PRN25, and PRN27 were in tracking channels. In the partial channel spoofing scene, the signals were received by a static antenna outdoors. There are five authentic signals (PRN04, PRN16, PRN23, PRN26 and PRN31) and four spoofing signals (PRN08, PRN09, PRN21 and PRN27) with 10 μs delay in tracking channels. The random motion is in the horizontal plane. The distance of the back and forth movement is about 10 m. The distance variation of 10 m is equivalent to the time variation of 3.3356 ns, as electromagnetic waves travel with the speed of light.

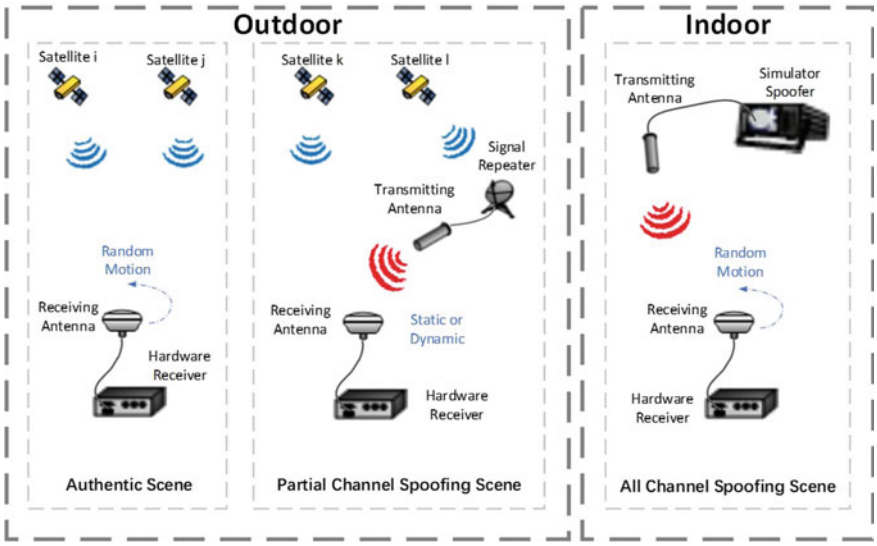


Fig. 1 Experiment configurations

Fig. 2 Experiment equipments

3.1 Authentic Scene

Figure 3 shows the satellite transmit time residuals and the pairwise correlation coefficients for different PRNs in the authentic scene. The satellite transmit time residuals for different PRNs have different variations in (a). The pairwise correlation coefficients for different PRNs are small in (b). The value of detection metric is 0.032643. It means that there is few correlation the measurements of between different PRNs.

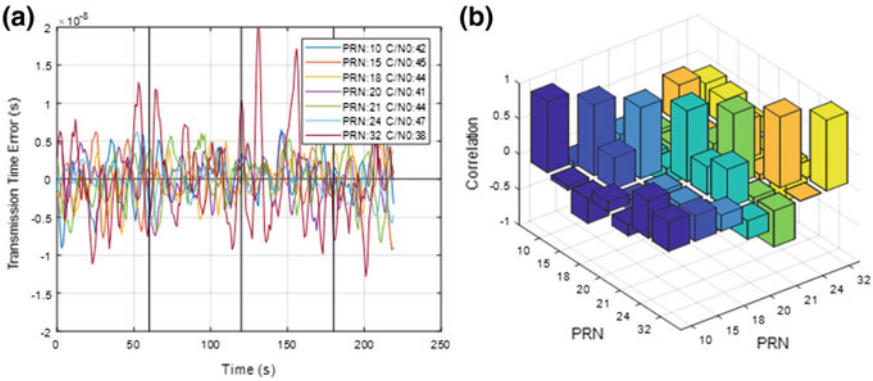


Fig. 3 **a** The satellite transmit time residuals and **b** the pairwise correlation coefficients for different PRNs in the authentic scene

3.2 All Channel Spoofing Scene

Figure 4 shows the satellite transmit time residuals and the pairwise correlation coefficients for different PRNs in the all channel spoofing scene. The satellite transmit time residuals for different PRNs have similar variations in (a). The pairwise correlation coefficients for different PRNs are large in (b). The value of detection metric is 0.94438. It means that there is a strong correlation between the measurements of different PRNs and the correlation is significant.

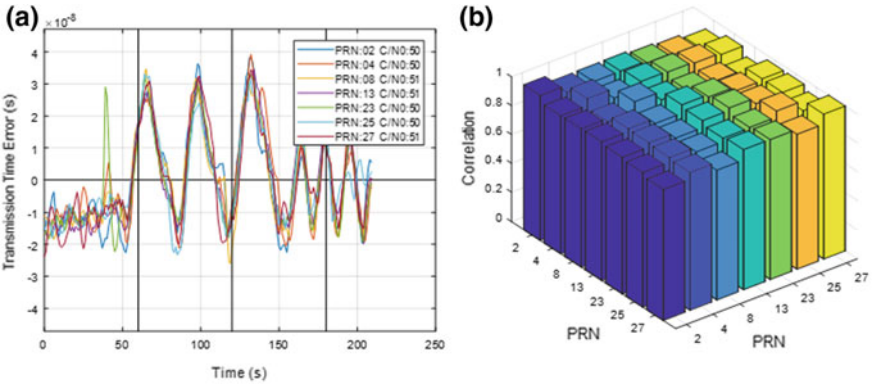


Fig. 4 **a** The satellite transmit time residuals and **b** the pairwise correlation coefficients for different PRNs in the all channel spoofing scene

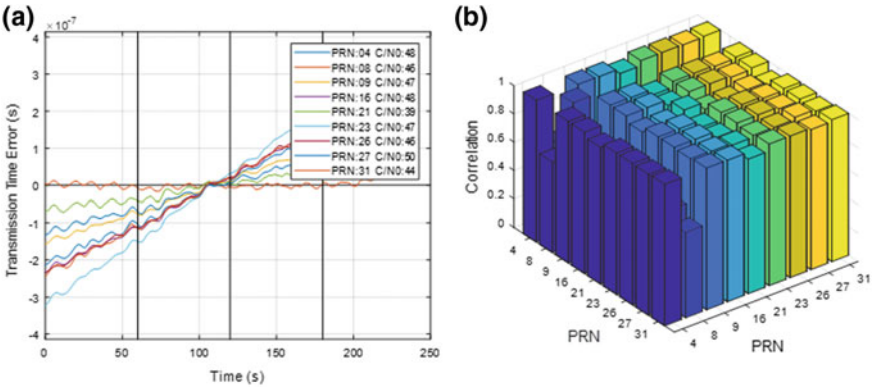


Fig. 5 **a** The satellite transmit time residuals and **b** the pairwise correlation coefficients for different PRNs in the partial channel spoofing scene

3.3 Partial Channel Spoofing Scene

Figure 5 shows the satellite transmit time residuals and the pairwise correlation coefficients for different PRNs in the partial channel spoofing scene. The satellite transmit time residuals for different PRNs have constant slopes in (a). The pairwise correlation coefficients for different PRNs are not small in (b). The value of detection metric is 0.9104. It means that there is a significant correlation between the measurements of different PRNs and the slope of the satellite transmit time residuals (a) is caused by the inconsistency of the PVT solution.

4 Conclusions

This paper proposes a method based on the pairwise correlations of satellite transmit time residual, which can be implemented in a single moving antenna receiver. This method combines the information of tracking level and navigation level. The PVT solution obtained in navigation level and the satellite transmit time obtained in tracking level are jointly utilized to calculate the satellite transmit time residuals. This operation helps receiver to get rid of clock parameters learning. Besides, this method is still valid when the spoofer fabricates counterfeit signals with random variations of position. The detection metrics represent the correlation of satellite transmit time residuals which extracted from the code phase measurements. The detection metrics in the spoofing scene are higher than the one in the authentic scene. The proposed method is an effective and reliable spoofing detection algorithm for the single-source spoofing attack. This method can be used in not only the all channel spoofing scene but also the partial channel spoofing scene.

Acknowledgments This work was supported by the National Natural Science Foundation of China (Grant No. 61571255).

References

1. Amin MG, Closas P, Broumandan A, Volakis JL (2016) Vulnerabilities, threats, and authentication in satellite-based navigation systems [scanning the issue]. Proc IEEE 104 (6):1169–1173
2. Broumandan A, Jafarnia-Jahromi A, Daneshmand S, Lachapelle G (2016) Overview of spatial processing approaches for GNSS structural interference detection and mitigation. Proc IEEE 104(6):1246–1257
3. Broumandan A, Jafarnia-Jahromi A, Dehghanian V, Nielsen J, Lachapelle G (2012) GNSS spoofing detection in handheld receivers based on signal spatial correlation. In: Paper presented at the position location and navigation symposium
4. Dana PH (1997) Global positioning system (GPS) time dissemination for real-time applications. Real-Time Syst 12(1):9–40
5. Jafarnia-Jahromi A, Daneshmand S, Broumandan A, Nielsen J, Lachapelle G (2013) PVT solution authentication based on monitoring the clock state for a moving GNSS receiver. In: Paper presented at the Vienna, Austria, presented at the European navigation conference (ENC2013)
6. Jahromi AJ, Broumandan A, Daneshmand S, Sokhandan N, Lachapelle G (2014) A double antenna approach toward detection, classification and mitigation of GNSS structural interference. In: Paper presented at the proceedings of NAVITEC 2014
7. Kaplan E, Hegarty C (2005) Understanding GPS: principles and applications, Artech house
8. Meuer M, Konovaltsev A, Cuntz M, Hättich C (2012) Robust joint multi-antenna spoofing detection and attitude estimation using direction assisted multiple hypotheses RAIM. In: Paper presented at the proceedings of the 25th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS 2012)
9. Psiaki ML, O'Hanlon BW, Powell SP, Bhatti JA, Wesson KD, Humphreys TE, Schofield A (2014) GNSS spoofing detection using two-antenna differential carrier phase. In: Paper presented at the proceedings of the 27th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS+ 2014)
10. Psiaki ML, Powell SP, O'hanlon BW (2013) GNSS spoofing detection using high-frequency antenna motion and carrier-phase data. In: Paper presented at the proceedings of the 26th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS+ 2013)

The Divergence Analysis of Kalman Filter Phase Lock Loop and Unbiased Correction of Frequency

Liwei Zhang, Dongqing Zhao, Shaolei Peng, Caijie Zhu and Zhongpan Li

Abstract In software receiver, the signal tracking is an important pre-step in demodulation of the satellite signals and positioning. The signals tracking is composed of the carrier tracking loop and the C/A code tracking loop, which can be used to track the carrier component and the C/A component. In this paper, the Kalman filter used in carrier tracking loop is studied in depth, and aimed at the phenomenon of divergence caused by the increasing phase-difference in the Kalman phase lock loop with the increase of the tracking time, the causes of this phenomenon is analyzed and the divergence function that can be used to predict the size of the divergence is posed. Then the tracking loop filter under Kalman adjusts the gain model frequency correction unbiased algorithm is put forward. Experiments show that the divergence function can correctly predict the divergence of the original tracking loop, and the divergence problem of standard Kalman filter can be solved. The same frequency and phase tracking of high stability and high accuracy is always kept in the carrier tracking loop by this method.

Keywords Kalman · PLL · Tracking · Divergence function · Unbiased correction Software receiver

1 Introduction

In the process of satellite signals acquisition, the software receiver can obtain the accurate carrier frequency and code phase through coarse and fine acquisition. In tracking stage, the pseudo-code tracking loop and the carrier tracking loop are tightly coupled by carrier frequency and code phase. The stripping operation of pseudo-code and carrier is implemented by continuously updating loop parameters. Finally, the bit data of the navigation message can be obtained by In-Phase

L. Zhang (✉) · D. Zhao · S. Peng · C. Zhu · Z. Li
Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450001, China
e-mail: zhangliwei2015@foxmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_41

(I) branch of the tracking loop. At the same time, the carrier phase observations can be calculated by Doppler frequency. Research suggests that the code loop is more robust than the carrier loop. Therefore, the carrier tracking loop is important for the positioning accuracy and dynamic performance of the GNSS software receiver [1].

With the development of society, GNSS receivers are used widely. Now there are great demand of positioning in a variety of severe environment, such as weak signal environment, high dynamic environment, etc. The traditional tracking loop has good performance in low dynamic and high signal-to-noise ratio environment, but the tracking performance is worse in hash environments. Kalman filter algorithm can be employed in the harsh environment [2, 3]. Zeng used Kalman filter tracking algorithm for tracking in weak signal environment, and the result indicated a lower 4 dB signal tracking compared to the traditional PLL/DLL tracking loop [4], another adaptive Kalman tracking PLL is proposed by Zhang which can be applied in the realization of signal tracking in high dynamic environment by adjusting the observation noise. But there is a defect in this algorithm that long time stable tracking cannot be realized. As the tracking time increases, divergence and losing lock will appear [5, 6]. In 2015, Gao analyzed the two algorithms of fading Kalman filtering and adaptive Kalman filtering, and proposed an improved adaptive extended Kalman filter tracking algorithm for residuals. The algorithm is effective in suppressing the divergence phenomenon, but the divergence problem is not solved well [7]. Therefore, I this paper the cause of the problem is theoretically analyzed and an unbiased model is solved to the problem.

2 Kalman Filter Phase Lock Loop

The carrier tracking loop is divided into PLL (Phase Lock Loop) and FLL (Frequency Locked Loop). Carrier phase measurement value is more accurate by PLL compared with FLL, but its stability is poor and it is easy to loss lock. FLL is not very sensitive to the frequency difference. Different frequency discriminator has different frequency pulling range, usually from 50 to 100 Hz. In the beginning of tracking stage, FLL is utilized in frequency pulling to lessen the frequency difference, then the signal tracking of the same frequency and phase is achieved by PLL.

2.1 The Equation of Kalman Filter

The equation of state and measurement equation of Kalman filter algorithm is written as [8]:

$$\begin{cases} X_k = \Phi X_{k,k-1} + \Gamma_{k,k-1} W_{k-1} \\ Z_k = H_k X_k + V_k \end{cases} \tag{1}$$

where X_k is the state vector at k moment, Z_k is the observation vector at k moment, Φ is the state transition matrix of the system, V_k is the vector of observation noise, W_k is the vector of system process noise, H_k is the observation matrix, Γ_k is the noise input matrix.

Five basic equations are used to estimate the state of the next time:

(a) State prediction equation:

$$\hat{X}_{k,k-1} = \Phi_{k,k-1} \hat{X}_{k-1} + BU(k) \tag{2}$$

(b) Prediction equation of error:

$$P_{k,k-1} = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + \Gamma_{k,k-1} Q_{k-1} \Gamma_{k,k-1}^T \tag{3}$$

(c) Filter gain matrix:

$$K_k = P_{k,k-1} H_k^T \left[H_k P_{k,k-1} H_k^T + R_k \right]^{-1} \tag{4}$$

(d) Estimation the variance matrix of error:

$$P_k = [I - K_k H_k] P_{k,k-1} \tag{5}$$

(e) Estimation state:

$$\hat{X}_k = \hat{X}_{k,k-1} + K_k [Z_k - H_k \hat{X}_{k,k-1}] \tag{6}$$

According to these five basic equations, the state estimation vector $\hat{X}_k$ of k moment can be obtained as long as the observation value Z_k of k moment is known.

2.2 State Model and Measurement Model of Third-Order PLL

At k moment, the phase difference between the locality carrier and the satellite signal carrier is x_φ. The Doppler frequency of satellite signals is x_ω. The Doppler frequency acceleration of the satellite signal is x_α. Then the state vector of k moment is $X_k = [x_\varphi, x_\omega, x_\alpha]_k^T$. When the sampling interval Δt is very small, the relationship between this time interval can be expressed as follows:

$$x_\omega(k+1) = x_\alpha(k)\Delta t + x_\omega(k) \tag{7}$$

$$x_\varphi(k+1) = x_\varphi(k) + 2\pi x_\omega(k)\Delta t + \pi x_\alpha(k)\Delta t^2 - 2\pi x_N(k)\Delta t \tag{8}$$

where $x_\omega(k+1)$, $x_\varphi(k+1)$ is the Doppler frequency and phase difference of the satellite signal at k + 1 moment respectively, and Δt is 1 ms.

Therefore, state transition equation of carrier phase is:

$$\begin{bmatrix} x_\varphi \\ x_\omega \\ x_\alpha \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & \Delta t & \Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_\varphi \\ x_\omega \\ x_\alpha \end{bmatrix}_k - \begin{bmatrix} \Delta t \\ 0 \\ 0 \end{bmatrix} x_N \tag{9}$$

where variables x_φ, x_ω, x_α, x_N are measured by rad.

In contrast to formula (2):

$$\Phi = \begin{bmatrix} 1 & \Delta t & \Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

$$B = \begin{bmatrix} \Delta t \\ 0 \\ 0 \end{bmatrix} \tag{11}$$

$$U = x_N \tag{12}$$

So the state transition equation can be expressed as:

$$\hat{X}_{k,k-1} = \Phi_{k,k-1}\hat{X}_{k-1} + BU(k) \tag{13}$$

The measured value of the measurement model is computed by the phase difference value between the local signal and the satellite signal, and the phase difference value is obtained by the two quadrant arctan function because the function is not affected by the navigation data bit flip.

$$\varphi_k = -\arctan(Q_P/I_P) \tag{14}$$

where φ_k is the phase difference value between the local signal and the satellite signal at k moment, Q_P, I_P is orthogonally integral prompt value and the in-phase integral prompt value after the mixing.

The phase difference value can also be expressed as:

$$\varphi = x_\varphi + x_\omega \frac{\Delta t}{2} + x_\alpha \frac{\Delta t^2}{6} - x_N \frac{\Delta t}{2} \tag{15}$$

So the measurement model of the PLL can be written as:

$$\varphi_{k+1} = \begin{bmatrix} 1 & \frac{\Delta t}{2} & \frac{\Delta t^2}{6} \end{bmatrix} \begin{bmatrix} x_\varphi \\ x_\omega \\ x_\alpha \end{bmatrix}_k - \frac{\Delta t}{2} x_N + v_{k+1} \tag{16}$$

The parameters are set up as follows [9]: $x_{\varphi 0}$, $x_{\omega 0}$ is obtained in the acquisition stage, $x_{\alpha 0}$ is set to 0 Hz/s, the standard deviation σ_v of measuring white noise v_k is 0.114 rad.

3 Unbiased Correction of Doppler Frequency in Kalman Filtering PLL

Kalman filtering algorithm is an optimal estimation algorithm for system state, which minimizes the mean square deviation between the predicted value and the state vector, that is, $X_k - \hat{X}_k$ reaches the minimum value. Therefore Kalman filtering algorithm has better tracking performance compared with the traditional PLL.

When the state vector X_{k-1} of the k − 1 moment is obtained, the frequency value of the X_{k-1} state vector is exploited to update the frequency of the local carrier, that is $x_N(k) = x_\omega(k-1)$. The value of x_α is considered unchanged within 2 Δt of the adjacent loop update cycle. However, due to the hysteresis of the filtering algorithm, there is a tiny difference between the frequency of the local carrier and the frequency of the input signal carrier. This difference depends on acceleration of frequency x_α. This tiny frequency difference value will cause minimal phase difference in one cycle. When the x_α remains the same, the phase difference will unceasingly accumulate as time goes on until losing lock. The phase difference $\Delta\varphi$ can be represented by a divergent function (the x_φ, x_ω, x_α, and x_N in this section are expressed in cycle).

$$\Delta\varphi = \int_0^{\Delta t} \int_0^{\Delta t} x_\alpha dt dt = \frac{1}{2} x_\alpha \Delta t^2 \tag{17}$$

From the divergence function, we can see that when the PLL is used to track for short time, the tracking can proceed normally, and the navigation message can also be demodulated, but the integral value of the in-phase will slowly decrease. When using the PLL for long time tracking, the phase difference between the local signal and the satellite signal will continue to expand at a minimum speed, and the speed magnitude in the normal observation environment is about $10'/\mathrm{s}$. After long time tracking, with the decrease of the integral value of the in-phase, the tracking loop will lose lock. In addition, because of the difference of frequency, the carrier phase observation will be more and more inaccurate.

In order to solve this problem, this paper novelly uses feedback adaptive under-regulation gain to correct the frequency estimation value obtained by Kalman filter PLL to ensure that the phase difference obtained at the next phase discrimination will not continue to expand and stabilize in a reasonable range.

The phase difference obtained by the phase discriminator is used as the input signal of the gain correction, and the correction model is expressed as follows:

$$\Delta f = \alpha \cdot q \cdot \frac{\Delta\varphi}{\Delta t} \tag{18}$$

where α is an under-regulation factor; q is a gain factor.

$$q = \begin{cases} |\Delta\varphi|/3\sigma & |\Delta\varphi| > 3\sigma \\ |\Delta\varphi|/\sigma & |\Delta\varphi| < \sigma \\ 1 & \cdots \end{cases} \tag{19}$$

where σ is the standard deviation of the phase difference.

After the state vector X_{k-1} is obtained at k − 1 moment, the frequency of the local replicating carrier at k moment is modified to $x_\omega(k-1) + \Delta f$ from the original algorithm's $x_\omega(k-1)$ (The under-regulation factor α is set to 0.01).

4 Experiment

Digital intermediate frequency (IF) sampler was used outdoors for data sampling (IF is 2.5 MHz, sampling frequency is 10.0 MHz) in November 9, 2017. The sampled data is processed by the software receiver. The Doppler frequency and the code phase of the satellite are obtained in the process of acquisition. The acquisition results are shown in Figs. 1 and 2. Take PRN10 satellite as example, its code phase is 882, and its Doppler frequency is 57.86 Hz.

In order to compare the phase difference between two tracking loops more intuitively, the phase difference cosine value is used instead of phase difference to map, that is, $\cos(\Delta\varphi)$. In Fig. 3, we present the phase difference cosine output diagram of the standard Kalman filter tracking algorithm and the unbiased Kalman filter tracking algorithm.

As we all know, the output value of the in-phase I in the tracking stage [10] is

$$I(k) \approx ADT\cos(\Delta\varphi) \tag{20}$$

where A is the signal magnitude, D is the navigation message bit, T is the integral time. With the increase of $\Delta\varphi$, the output value of the branch I will become smaller and smaller.

It can be concluded from Fig. 3 that after 300 s tracking, the tracking loop based on standard Kalman algorithm has been divergent, and the tracking loop based on

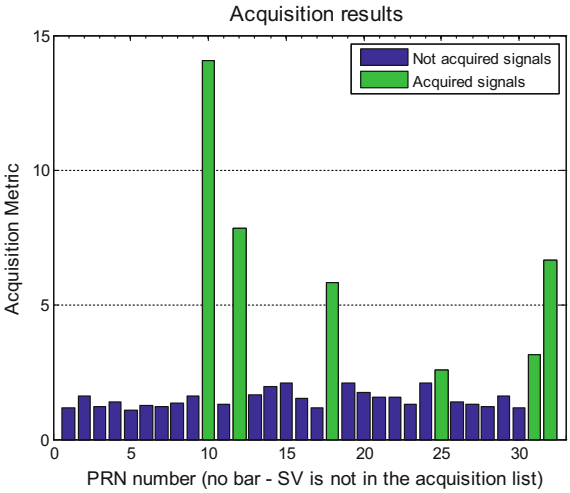


Fig. 1 Acquisition results

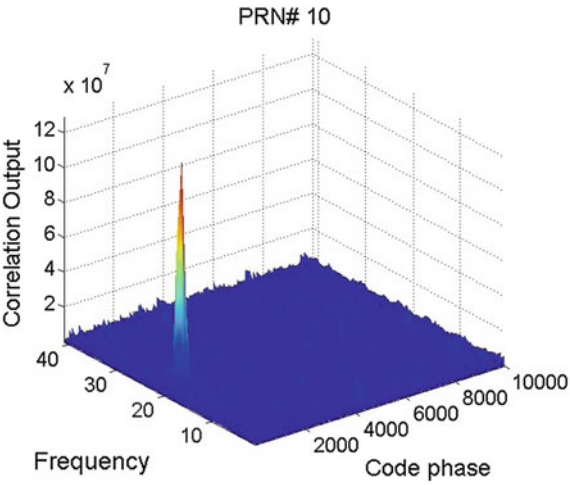


Fig. 2 Acquisition results of the PRN 10

correction Kalman filtering algorithm proposed in this paper can still achieve stable tracking of the satellite signal and did not show any signs of divergence. From Fig. 4, it is clearly shown that the integral value of branch I is getting smaller and smaller, because the phase deviation is larger and larger with the increase of tracking time.

Figure 5 shows the carrier frequency at each time. From the figure, it can be calculated that the frequency changes about −180 Hz in 300 s, and the change rate is about −0.6 Hz/s (−180 Hz/300 s). From the divergent functional expression (17) $\Delta\varphi = \frac{1}{2}x_\alpha \Delta t^2$, It can be obtained that the phase variation caused by the frequency deviation is $6\pi \times 10^{-7}$ in 1 ms updating cycle period, and the phase difference

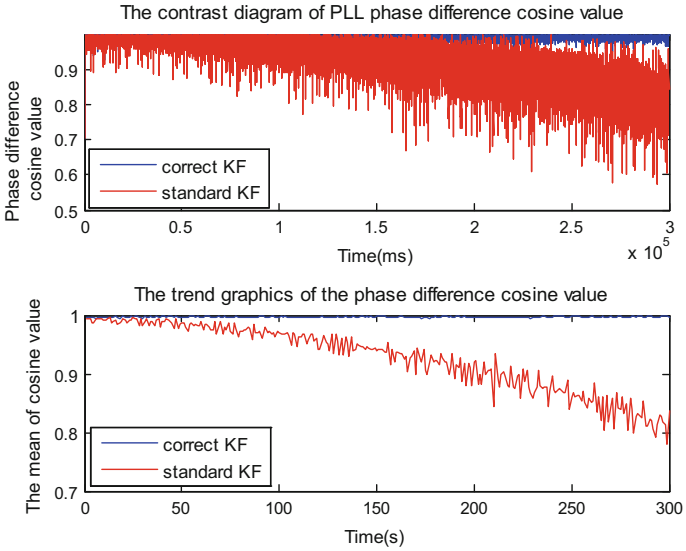


Fig. 3 Phase difference cosine value of two algorithms

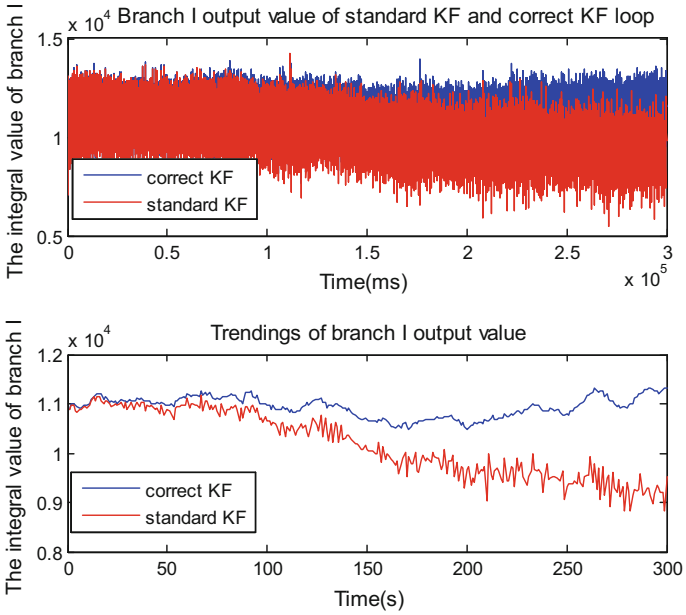


Fig. 4 Branch I value of two tracking loops

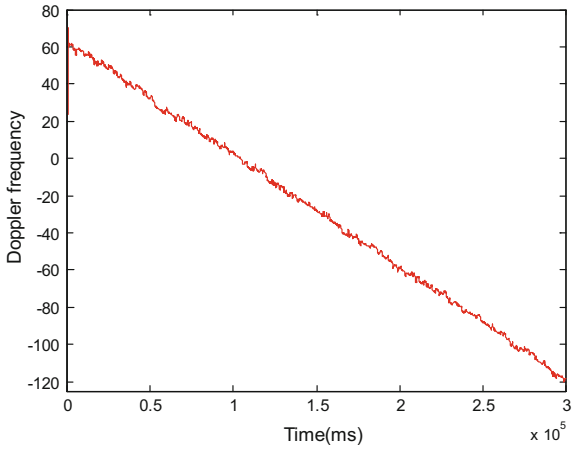


Fig. 5 The changes of carrier frequency

Table 1 The between the output values of the unbiased tracking algorithm and the traditional tracking algorithm

300 s	Correction KF	Standard KF	Predicted value	The difference between the predicted value and the measured value (%)
Phase deviation cosine value	0.9973	0.8369	0.8377	0.1
Branch I output value	1.133	0.9519	0.9517	0.02

caused by the slight frequency offset is 0.18π in 300 s. Therefore, phase difference cosine value of the tracking algorithm based on standard KF should be $\cos(0.18\pi)$, that is about 0.84. Accordingly, the branch I output value and the phase difference cosine value obtained by the standard KF algorithm should be 0.84 times of the output value of the unbiased correction algorithm.

It can be seen from Table 1: that at the 300 s, the phase difference cosine obtained by unbiased correction tracking algorithm is 0.9973 and the phase difference cosine obtained by standard KF is about 0.8369. The difference is only 0.1% from the predicted value 0.8377 (0.9973×0.84). The integral value of the branch I is 1.133×10^4 through unbiased correction algorithm. The branch I integral value of the uncorrected tracking algorithm is 0.9519×10^4, the difference is only 0.02% from the predicted value 0.9517×10^4. Therefore, it can be considered that the branch I integral output value and the phase difference cosine value obtained by the standard KF algorithm fully conform to the theoretical prediction value.

5 Conclusions

In this paper, the third-order Kalman filter PLL algorithm is studied deeply. The problem of divergence is analyzed with the increase of tracking time and the cause of divergence is found. After that, the divergence function is derived through theoretical analysis. Then, this paper proposes a self-adaptive Kalman filtering PLL unbiased correction algorithm based on under-adjustment gain model. Finally, the experimental results show that the algorithm can solve the divergence problem of standard Kalman filtering algorithm and more stable tracking of the satellite signal is achieved.

Acknowledgments The National Natural Science Foundation of China (Grant No. 41774037).

References

1. Zhang J (2012) The key technologies of the high-dynamic carrier tracking loop in GNSS software receiver. Dalian Maritime University
2. Psiaki ML, Jung H, Jung H (2002) Extended Kalman filter methods for tracking weak GPS signals. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, pp 2539–2553
3. Hinedi S (1988) An extended Kalman filter based automatic frequency control loop. The Telecommunications and Data Acquisition Report, pp 219–228
4. Zeng X (2010) Study on the algorithm for weak signal acquisition and tracking in GPS software receiver. PLA Information Engineering University
5. Yu KKC, Watson NR, Arrillaga J (2005) An adaptive Kalman filter for dynamic harmonic state estimation and harmonic injection tracking. IEEE Trans Power Deliv 20(2):1577–1584
6. Jwo DJ, Wang SH (2007) Adaptive fuzzy strong tracking extended Kalman filtering for GPS navigation. IEEE Sens J 7(5):778–789
7. Gao T (2015) Research on carrier tracking technology of high dynamic navigation signal based on Kalman filter. Hunan University
8. Fu M, et al (2010) Kalman filtering theory and its application in navigation system. Science Press
9. Psiaki ML (2010) Smoother-based GPS signal tracking in a software receiver. IEEE Trans Aerosp Electron Syst
10. Lu Y (2009) The principle and implementation technology of the BDS/GPS dual mode software receiver. Beijing Electron Industry Press, pp 82–84

A Method of Carrier Phase Multipath Mitigation Based on Punctual Code Correlation Reference Waveform

Chunjiang Ma, Xiaomei Tang, Yingxiang Liu, Zhibin Xiao and Guangfu Sun

Abstract Carrier phase multipath is one of the main sources of error for high-precision GNSS applications. In an urban environment with severe multipath fading, multipath effects can cause centimeter-scale shifts in carrier phase measurements. The carrier multipath is related to the location of the receiver, the variation of static multipath error is mainly affected by the satellite motion, showing the deviation characteristic, which is difficult to eliminate by using the existing differential technique or multipath error model. This paper defines the concept of carrier multipath peak-to-peak ratio in carrier correlator and uses it to measure the effect of multipath signal on carrier phase measurement. Considered the characteristics that multipath signals always lag behind direct signals, this paper presents a method of reference waveform design for on-time codes. The asymmetric cross-correlation function constructed by this method can suppress the cross-correlation of lag signals and improve the anti-carrier phase multipath capability. Through theoretical analysis and simulation, the proposed method of carrier phase multipath reduction can reduce the envelope area of carrier phase multipath error by about 49% when the gate width of the reference waveform is 1/4 chip, effectively increasing the receiver carrier multipath rejection.

Keywords Carrier phase · Multipath peak to peak ratio · Code reference waveform · GNSS

1 Introduction

In Global Navigation Satellite Systems (GNSS), high-precision positioning technologies such as Real Time Kinematic (RTK) and Precise Point Position (PPP) provide users with centimeter-level positioning services by using carrier

C. Ma · X. Tang · Y. Liu · Z. Xiao · G. Sun (✉)
College of Electronic Science, National University of Defense Technology, Changsha, Hunan, People's Republic of China
e-mail: gfsunmail@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_42

phase [1]. Carrier phase multipath is one of the main sources of errors in high-precision positioning, especially in urban environments where the multipath fading is severe. Multipath effects can cause centimeter-scale dispersion in carrier phase measurements [2].

High-precision positioning technology can reduce the influence of tropospheric delay, ionospheric delay and clock error on carrier phase measurement through differential or modeling. However, the multipath errors are not the same due to the same satellite signals received at different locations or different satellite signals received at the same location. Therefore, the differential and modeling methods are generally effective in dealing with multipath errors. For static or low dynamic users, the variation of multipath error is mainly affected by the satellite motion, showing the characteristics of slow time-varying deviation [3]. This makes it difficult to repair carrier multipath errors in a short period of time even with data filtering methods.

At present, there are lots of literatures which researched on pseudo-range multipath [4–10]. Compared with pseudo range multipath, the effect of carrier phase multipath is small, but the processing is more difficult, so research on carrier multipath is still relatively small. However, the current high-precision positioning technology mainly uses the carrier phase, so the influence of carrier multipath error on high-precision users is particularly prominent.

The anti-multipath antenna is mainly suitable for stationary high-precision users, which can effectively suppress the multipath error of the carrier phase. For receiving antennas placed in a more open environment, the multipath signal is mainly incident from a negative elevation angle. Patch antennas [4] and Fire Wheel Antennas [5], etc. are all aimed at reducing the antenna's negative gain as much as possible while ensuring the visibility of the antenna to satellites above horizontal.

The Maximum Likelihood Estimation (MLE) method [6] estimates the amplitude and phase of multipath signals based on the maximum likelihood principle, and compensates for the carrier phase multipath errors by using the signal parameter estimation. In addition, in order to improve the computational efficiency, a series of parameter estimation methods based on maximum likelihood are proposed, such as the Multipath Estimation Delay Lock Loop (MEDLL) [7], the Maximum Likelihood Delay Estimation Reduced Searching Space Maximum Likelihood (RSSML) technology [8] and Fast iteration Maximum-Likelihood Algorithm (FIMLA) algorithm [9]. Parameter estimation method can estimate the parameters of multipath signals accurately, but its calculation is relatively complex and sensitive to noise. Some algorithms have fuzzy estimation, and additional information needs to be introduced to assist decision.

In 1999, Gary's proposed High Resolution Correlator (HRC) technique [10] is a multi-correlator carrier multipath suppression technique that uses three sets of correlators (lead, punctual and lag) to track the carrier. However, this method not only increases the thermal noise variance of the carrier phase but also is severely affected by multipath signals near one chip.

Based on the idea of HRC technology, this paper proposes a carrier phase multipath suppression method based on Punctual Code Reference Waveform

(PCRW). The concept of multipath peak to peak ratio (MPPR) is proposed and MPPR is used to guide the design of reference waveform of punctual code according to the theoretical formula of multipath error of carrier phase. Compared with the HRC method, the PCRW method has better carrier phase multipath suppression capability under the same SNR loss.

2 Carrier Phase Multipath Error

Multipath error refers to the deviation of the measurement of the direct signal due to the received reflected or scattered replica signal. For N-path multipath model, ignoring the effects of noise and other disturbances, the received signal can be expressed as:

$$r(t) = a_0 \mathrm{e}^{-j\theta_0}\left[x(t-\tau_0) + \sum_{i=1}^{N} \tilde{a}_i \mathrm{e}^{-j\tilde{\theta}_i} x(t-\tau_0-\tilde{\tau}_i)\right] \tag{1}$$

where a_0 is the signal amplitude, $\tilde{\theta}_0$ is the initial phase of the signal carrier phase, τ_0 is the signal delay, $\tilde{a}_i$, $\tilde{\theta}_i$ and $\tilde{\tau}_i$ are the amplitude attenuation, phase offset and delay of the multipath signal, respectively.

When the receiver performs the matching correlation, the correlation function of the local reference signal $\mathrm{e}^{-j\hat{\theta}}x(t)$ and the received signal $r(t)$ of the on-time tributary may be expressed as:

$$\bar{R}(\tau) = a_0 \mathrm{e}^{-j\left(\theta_0-\hat{\theta}\right)}\left(R(\tau-\tau_0) + \sum_{i=0}^{N} \tilde{a}_i \mathrm{e}^{-j\tilde{\theta}_i} R(\tau-\tau_0-\tilde{\tau}_i)\right) \tag{2}$$

Among them,

$$R(\tau) = \int x(t)x(t+\tau)dt \tag{3}$$

Assuming that the difference between the local carrier phase and the received signal carrier phase is $\varphi\left(=\theta_0-\hat{\theta}\right)$, that is, the true value of the carrier phase measurement. Due to multipath signal caused by carrier phase measurement deviation, that is, carrier phase multipath error φ_e can be expressed as:

$$\varphi_e\left(\tilde{\mathbf{a}}, \tilde{\theta}, \tilde{\tau}\right) = \arctan\left[\frac{\sum_{i=0}^{N} \tilde{a}_i \sin\left(\tilde{\theta}_i\right) R(-\tilde{\tau}_i)}{R(0) + \sum_{i=0}^{N} \tilde{a}_i \cos\left(\tilde{\theta}_i\right) R(-\tilde{\tau}_i)}\right] \tag{4}$$

It is assumed here that the code phase tracking error is negligible. Deformation on the type, we can get:

$$\varphi_e\left(\tilde{\mathbf{a}},\tilde{\theta},\tilde{\tau}\right) = \arctan\left[\frac{\sum_{i=0}^{N}\tilde{a}_i\sin\left(\tilde{\theta}_i\right)R(-\tilde{\tau}_i)/R(0)}{1+\sum_{i=0}^{N}\tilde{a}_i\cos\left(\tilde{\theta}_i\right)R(-\tilde{\tau}_i)\Big/R(0)}\right] \tag{5}$$

The ratio of the multipath peak-to-peak ratio of the carrier multipath is defined as the ratio of the correlation peak of the multipath signal to the direct signal, and then the multipath peak-to-peak ratio of the multipath of the multipath signal of the *i*-th path can be calculated as follows:

$$r(\tilde{\tau}_i) = R(-\tilde{\tau}_i)/R(0), \tilde{\tau}_i > 0 \tag{6}$$

According to Eq. (5), the carrier phase multipath error is directly related to the size of carrier multipath peak-to-peak ratio. The larger the carrier multipath peak-to-peak ratio, the more serious the multipath signals will affect the carrier phase.

The carrier phase multipath error envelope is composed of the maximum and minimum values of carrier phase multipath error, which can reflect the average effect of different delayed multipath signals on carrier phase measurement. The maximum and minimum values of carrier phase multipath errors are calculated as follows:

$$\varphi_{\min}(\tilde{\mathbf{a}},\tilde{\tau}) = \min_{\substack{\tilde{\theta}_i\in(0,2\pi]\\(i=1,\cdots,N)}} \varphi_e\left(\tilde{\mathbf{a}},\tilde{\theta},\tilde{\tau}\right) \tag{7}$$

$$\varphi_{\max}(\tilde{\mathbf{a}},\tilde{\tau}) = \max_{\substack{\tilde{\theta}_i\in(0,2\pi]\\(i=1,\cdots,N)}} \varphi_e\left(\tilde{\mathbf{a}},\tilde{\theta},\tilde{\tau}\right) \tag{8}$$

The carrier phase error envelope can be given by $[\varphi_{\min}, \varphi_{\max}]$.

3 Punctual Code Reference Waveform

In order to reduce the impact of multipath signals on the carrier phase measurement, this paper designs the code reference waveform of the punctual branch. By reducing the multipath peak-to-peak ratio of the punctual branch (carrier) correlator, the receiver can improve the carrier multipath mitigation ability.

As shown in Fig. 1, the local code waveform of the traditional all-match receiving method is exactly the same as that of the spread spectrum. The HRC

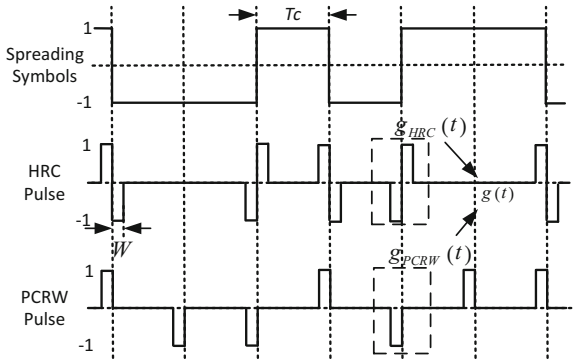


Fig. 1 Example of locally generated PRN code and tailored pulses

method for the multiple correlator combination can be equivalent to the HRC code reference waveform given in the figure. The carrier phase multipath method proposed in this paper is achieved by designing the PCRW code reference waveform.

The reference code waveform can be described by the following formula:

$$w(t) = \sum_{k=0}^{\infty} g\left(t - kT_C + \frac{1}{2}T_C\right)c_k(t) \tag{9}$$

where $g(t)$ is the reference waveform symbol, $c_k(t)$ is the spreading code symbol, T_C is the spreading code symbol width, and W is the reference waveform gate width.

Receiver block diagram using PCRW technology shown in Fig. 2. The PCRW reference code generator will generate a local code reference waveform $w(t)$ of the on-time branch based on the local replica code generator.

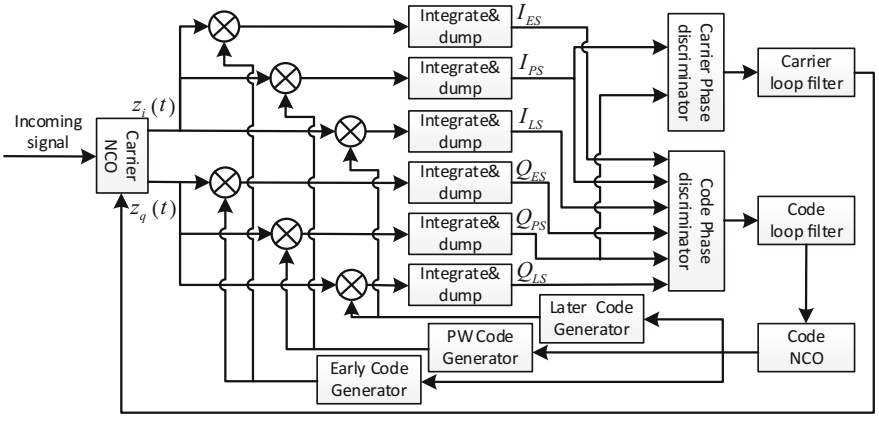


Fig. 2 Block diagram of receiver code loop of proposed scheme

Since the early code and the late code are directly derived from the output of the local code generator and are independent of the PCRW reference code. Therefore, PCRW technology can only suppress carrier multipath and has no direct effect on pseudorandom multipath. While using PCRW technology to suppress carrier phase multipath, code phase multipath can still be suppressed by using narrow correlation (NC) or Code Cross Reference Waveform (CCRW).

Where, C represents the amplitude of the received signal, N_{IW} and N_{QW} are Gaussian noise, the noise variance is:

$$\sigma_{XW}^2 = E\left(N_{IW}^2\right) = E\left(N_{QW}^2\right) = R_{WW}(0)N_0/T \tag{13}$$

where N_0 is the noise power spectral density of the input signal and T is the coherent integration time.

Because multipath signals always lag behind the direct signals in time, the PCRW technique designs an asymmetric cross-correlation curve to suppress the correlation values of the lagging portions and reduce the influence of multipath signals on the direct signals in the correlation domain.

Under the condition of infinite bandwidth, the cross matching function of full match reception, HRC method and PCRW method proposed in this paper is shown in Fig. 3, where the gate width W of the punctual reference waveform takes 1/4 chip. As can be seen from the figure, the PCRW method can suppress multipath signal correlation values delayed by more than 1/4 chip, while the HRC method is sensitive to multipath signals near one chip delay.

In addition, as can be seen from Fig. 3, since the HRC method and the PCRW method only partially match the reference code, there is a certain amount of related loss, which is analyzed in detail in Part 4 of the article.

The MPPR of the cross-correlation curve between the received signal and the local reference waveform directly determines the multipath suppression capability

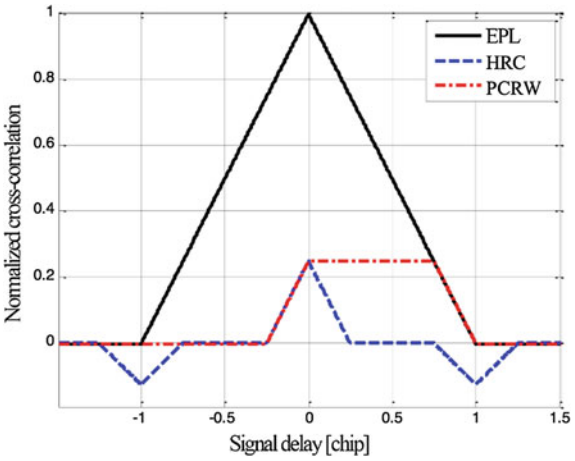


Fig. 3 Several reference waveforms correspond to the cross-correlation function (B = ∞, W = 1/4)

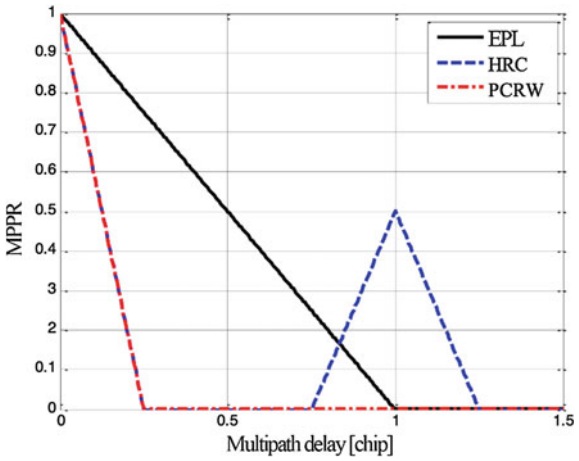


Fig. 4 The multipath peak-to-peak ratio of several reference waveforms (B = ∞, W = 1/4)

of the carrier correlator. Figure 4 shows the MPPR corresponding to several reference waveforms. It can be seen that the PCRW method is consistent with the anti-multipath performance of HRC for multi-path signals with 0–0.75 chips delay, and both are superior to fully matched receiver method. The PCRW method completely suppresses multipath with a signal delay greater than 1/4 chip, while the HRC method is more sensitive to multipath signals with a delay of 0.75–1.25 chips.

The received signal is orthogonally demodulated by a carrier to obtain an in-phase component z_i and a quadrature component z_q, respectively, both of which are related to a local code signal $w(t)$, and then integrated to obtain I_{PS} and Q_{PS}.

The cross-correlation function that defines the received signal and the PCRW reference code is:

$$S_{IF} = \sqrt{2C}X(t)\exp[j((\omega_0 + \omega_d)t + \theta_e)] + n(t) \tag{10}$$

The in-phase integrated value $I_W(\varepsilon)$ and the quadrature integrated value $Q_W(\varepsilon)$ of the on-time correlator can be expressed as:

$$I_W(\varepsilon) = \sqrt{2C}R_{XW}(\varepsilon)\cos\varphi + N_{IW} \tag{11}$$

$$Q_W(\varepsilon) = \sqrt{2C}R_{XW}(\varepsilon)\sin\varphi + N_{QW} \tag{12}$$

To facilitate the analysis of the problem, consider only one case of signal multipath and assume that the multipath to direct ratio (MDR) of the multipath signal is −6 dB. Figure 5 shows the multipath error envelopes for several reference waveforms under infinite bandwidth conditions. The full-match receiving method has the largest multipath error envelope area, and the PCRW method proposed in this paper has the smallest multipath error envelope area. The PCRW method reduces the envelope area of multipath error by 75% compared with the full match receiving method, and decreases by 49% when compared with the HRC method.

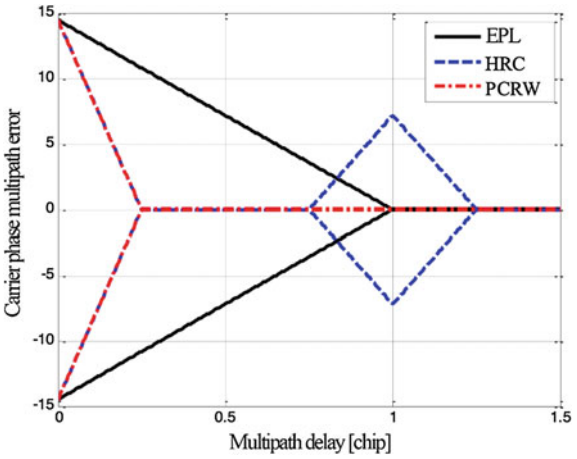


Fig. 5 Carrier phase multipath error envelope corresponding to several reference waveforms (B = ∞, W = 1/4, MDR = −6 dB)

In the case of infinite bandwidth, reducing the gate width W of the reference waveform can further reduce the error envelope area of carrier phase multipath and improve the suppression of carrier multipath. For an actual receiver, the limited bandwidth will distort the signal correlation peak, reduce the peak-to-peak ratio of carrier multipath, and reduce the suppression of carrier phase multipath. Figure 6 shows the carrier phase multipath error envelope corresponding to the PCRW reference codes of different gate widths when receiving the GPS C/A code when the front-end bandwidth of the radio frequency is 5 MHz. It can be seen that when using the PCRW method for receiving, it is also necessary to combine the specific radio frequency front-end bandwidth and select a suitable gate width.

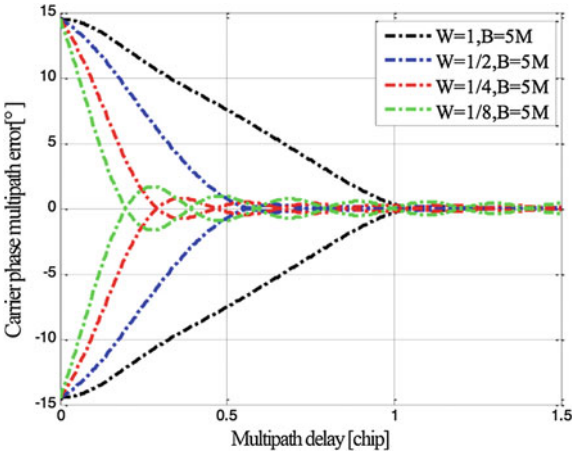


Fig. 6 Carrier multipath error envelopment with different gate widths under limited bandwidth (Fc = 1.023 MHz, MDR = −6 dB)

4 Tracking Jitter

Since both PCRW and HRC are a partially matched reception method, there is energy loss when the local code is correlated with the received signal.

In the following, a multiplicative phase detector is taken as an example to analyze the loop thermal noise of the carrier tracking loop based on the PCRW method. Combined with Eqs. (11) and (12), the carrier phase error can be expressed as:

$$\begin{aligned} d(\theta_e) &= I_W(0) * Q_W(0) \\ &= CR_{XW}^2(0)\sin 2\varphi + \sqrt{2C}R_{XW}(0)N_{QW}\cos\varphi \\ &\quad + \sqrt{2C}R_{XW}(0)N_{IW}\sin\varphi + N_{IW}N_{QW} \end{aligned} \tag{14}$$

When the carrier loop is locked, $\varphi \approx 0$, therefore, $\cos\varphi \approx 1$. The phase identification function is linearized:

$$d(\varphi) = g_0\varphi + N \tag{15}$$

where g_0 is the phase-contrast gain, and N is a Gaussian white noise with mean zero. g_0 and N can be expressed as:

$$g_0 = 2CR_{XW}^2(0) \tag{16}$$

$$N = \sqrt{2C}R_{XW}(0)N_{QW} + N_{IW}N_{QW} \tag{17}$$

Among them, the variance of noise N can be expressed as:

$$\sigma_N^2 = 2CR_{XW}^2(0)R_{WW}(0)N_0/T + R_{WW}^2(0)(N_0/T)^2 \tag{18}$$

According to the linearization model of carrier phase discrimination function, the normalized steady-state carrier tracking error is:

$$\sigma^2 = \frac{2B_L T\sigma_N^2}{g_0^2} = \frac{B_L R_{WW}(0)}{R_{XW}^2(0)C/N_0}\left(1 + \frac{R_{WW}(0)}{2R_{XW}^2(0)C/N_0T}\right) \tag{19}$$

where B_L represents the unilateral noise bandwidth of the code tracking loop.

Under the condition of infinite bandwidth, the auto-correlation peak of PRWR and the cross-correlation peak of PRWR with the input signal are the same in size and equal to the gate width of the reference waveform, that is, the correlation peak can be expressed as:

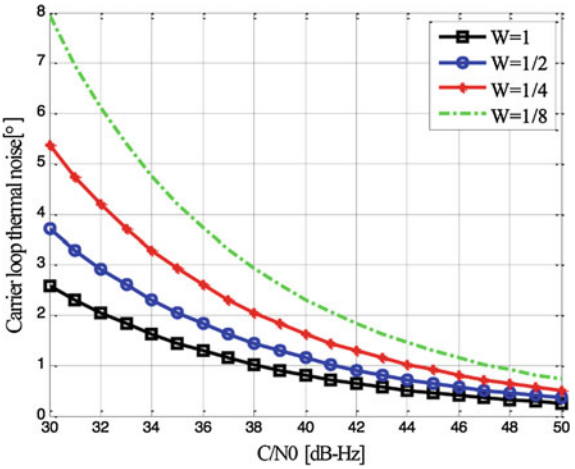


Fig. 7 The standard deviation of loop thermal noise based on different gate widths (BL = 2 Hz, T = 20 ms)

$$R_{WW}(0) = R_{XW}(0) = GW \tag{20}$$

The fully matched receive method is equivalent to the PCRW method with a gate width equal to one chip, whereas the thermal noise performance of the HRC method is exactly the same as the PCRW method. Figure 7 shows the variation of the thermal noise of PCRW carrier phase with the carrier-to-noise ratio of the loop under W = 1, W = 1/2, W = 1/4 and W = 1/8 respectively.

It can be seen that the smaller the gate width of the reference waveform, the greater the variance of the carrier loop thermal noise. Reduce gate width by half the size, which can be equivalent to loss of 3 dB input signal power.

After the above analysis shows that the carrier multipath suppression performance improvement at the expense of receiving SNR at the expense. However, for stationary or low dynamic receivers, the relatively fixed carrier phase multipath drift has a greater impact on the positioning result, while the loop thermal noise can still be suppressed by increasing the integration time or data filtering.

5 Conclusion

In this paper, a carrier phase multipath suppression method based on PCRW is proposed. The concept of MPPR is defined, and MPPR is used to guide the design of local code reference waveform by deriving a formula of carrier multipath error. Analysis shows that when the gate width of the local punctual branch code reference waveform is 1/4 chip, the reception SNR loss of the PCRW receiving method is the same as that of the HRC method, but the anti-multipath performance is improved by about 49%.

Acknowledgements This work was supported by the National Natural Science Foundation of China under grant No. 61601485, 61403413 and 41604016.

References

1. Kouba J, Héroux P (2010) GPS precise point positioning using IGS orbit products. GPS Solutions 5(2):12–28
2. Xu Ch, Tang X, Huang Y, Wang F (2016) Signal carrier frequency variation approach for GEO satellite multipath mitigation in the BDS. IEICE Trans Commun E99-B(11)
3. Lau L, Cross P (2006) A new signal-to-noise-ratio based stochastic model for GNSS high-precision carrier phase data processing algorithms in the presence of multipath errors. In: Proceedings of the ION GNSS 2006, 26–29 September, Fort Worth, TX, pp 276–285
4. Boccla L, Amendola G, Massa GD, Glullcchl L (2001) Shorted annular patch antennas for multipath rejection in GPS-based attitude determination systems. Microwave Opt Technol Lett 28(1):47–51
5. Tranquilla JM, Carr JP (1994) Analysis of a choke ring ground-plane for multipath control in global positioning system applications. IEEE Trans Antennas Propag 42(7):905–911
6. Bryan R, Townsend B, Richard DJ, et al (1995) Performance evaluation of the multipath estimating delay lock loop. In: Proceedings of the ION national technical meeting
7. Van Nee Richard DJ (1992) The multipath estimating delay lock loop. In: IEEE 2nd international symposium on spread spectrum techniques and applications (ISSTA'92), Yokohama, Japan, pp 39–42
8. Bhuiyan MZH, Lohan ES (2010) Advanced multipath mitigation techniques for satellite-based positioning applications. Int J Navig Obs 412393:1–15
9. Sahmoudi M, Amin MG (2008) Fast iterative maximum-likelihood algorithm (FIMLA) for multipath mitigation in the next generation of GNSS receivers. Trans Wireless Commun 7 (11):4362–4374
10. McGraw G, Braasch M (1999) GNSS multipath mitigation using gated and high resolution correlator concepts. In: Proceedings of the national technical meeting of the institute of navigation, San Diego, California, January 1999, pp 333–342

Part III
Multi-source Fusion Navigation Technology

Maximum Correntropy Criterion Based Robust Kalman Filter

Liansheng Wang, XingWei Gao and Lijian Yin

Abstract In this paper, a novel filter based on maximum correntropy criterion (MCC) and M-estimate theory is proposed. The MCC based Kalman filter (MKF) is developed by combining the MCC and the M-estimate theories into the framework of the classical Kalman filter (CKF). In this way, the measurement outliers can be suppressed. In addition, since the dynamic model of the filter is consistent with the real movement of the vehicle, the proposed MKF performs better than Huber function based Kalman filter (HKF) and the CKF in state estimations. The performance of the proposed filer is proved by the numerical simulation of a moving vehicle with constant velocity and real GNSS static positioning results.

Keywords Robust Kalman filter · Maximum correntropy criterion
M-estimate theory · Huber function

1 Introduction

The Kalman filter (KF) type estimators have been extensively applied in the area of navigation. However, a problem of these kinds of estimators is that they are sensitive to outliers, because they rely on weighted least-squares (WLS) criteria or ℓ_2 norm estimation procedure which is susceptible to outliers. Thus robust Kalman filter should be applied to process the data contaminated by outliers.

Among the robust estimators, the most popular one is the M-estimate theory proposed by Huber in [1, 2]. From a Bayesian maximum likelihood perspective,

L. Wang (✉) · X. Gao
Allystar Technology (Beijing) CO. Limited, Beijing, China
e-mail: wanglsh@allystar.com

X. Gao
e-mail: gaoxw@allystar.com

L. Yin
School of Automation, Beijing Institute of Technology, Beijing, China
e-mail: yinlijian00@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_43

by introducing the Huber function to modify the cost function of the classical Kalman filter (CKF), the outliers' bad effect from both observations and updated parameters can be mitigated by appropriate initial values based iterations [3, 4]. In [5], the robust estimators are categorized as: M-LS filter when outliers only exist in observations, LS-M filter when outliers only exist in updated parameters, and M-M filter when outliers exist both in updated parameters and observations. However, to employ the robustness of the filter, the dimension of measurements should be larger than that of states.

To suppress the bad effect of contaminated measurement noises or outliers, the information theoretical tool called correntropy was proposed and a maximum correntropy criterion (MCC) was proposed and applied in engineering applications [6]. In [7], a robust information filter based on MCC is proposed. Although the proposed filter is also robust to measurement outliers and can reduce the estimation errors, unfortunately, only the robustness to contaminated Gaussian measurement noises is highlighted. Additionally, in order to pursue the Gaussian efficiency, the more complicated information filter has to be employed.

In this paper, the maximum correntropy criterion (MCC) combining the M-estimate theory is applied to form a novel kind of robust Kalman filter. When the state model of the filter is consistent with the real movement of the vehicle, its robustness to measurement outliers and performance in estimations are investigated.

2 Brief Review of M-Estimate Theory and Maximum Correntropy Criterion

2.1 Overview of M-Estimate Theory

In this section, the linear regression model of the form is considered:

$$\mathrm{y} = Hx + w \tag{2.1}$$

where $y \in R^m$ is the is the measurements with dimension m, $\mathrm{x} \in R^n$ is the variable to be estimated with dimension n, $\mathrm{H} \in R^{m \times n}$ is the design matrix, **w** represents the random measurement errors.

To solve the problem, Huber introduced the generalized maximum likelihood method in 1964, also known as the M-estimation theory [1]. In this method, the following cost function of residuals is sought to be minimized:

$$J(x) = \sum_{i=1}^{m} \rho(e_i) \tag{2.2}$$

where $\mathbf{e} = \mathbf{Hx} - \mathbf{y}$ is the measurement residuals, ρ is a differentiable function. The solution to the generalized maximum likelihood problem can be found from the implicit equation

$$\sum_{i=1}^{m} \phi(e_i) \frac{\partial e_i}{\partial \mathbf{x}} = 0 \tag{2.3}$$

where $\phi(e_i) = \rho'(e_i)$. By defining the function $\psi(e_i) = \phi(e_i)/e_i$, and the matrix $\Psi = diag[\psi(e_i)]$, the implicit Eq. (2.3) can be written in matrix form as

$$\mathbf{H}^T \Psi(\mathbf{Hx} - \mathbf{y}) = 0 \tag{2.4}$$

In generalized maximum likelihood estimation, the function ρ can be chosen to yield an estimator $\hat{\mathrm{x}}$ with certain desirable properties. The ρ, ϕ, ψ functions are sometimes known as the score, influence and weight functions, respectively.

2.2 Review of Maximum Correntropy Criterion

The correntropy between two arbitrary scalar random variables X and Y is defined by Liu et al. [8]

$$V_\sigma(X, Y) = \mathbf{E}[\kappa_\sigma(X - Y)] \tag{2.5}$$

In practice, the joint probability density function (PDF) is unknown and only a finite number of data $(x_i, y_i)_{i=1}^N$ are available, leading to the sample estimator of correntropy

$$\hat{V}_{N,\sigma}(X, Y) = \frac{1}{N} \sum_{i=1}^{N} \kappa_\sigma(x_i - y_i) \tag{2.6}$$

where $\kappa_\sigma(x_i - y_i)$ is the kernel function, σ is the kernel width or bandwidth. Correntropy of the following error can be used as a new cost function, which is called maximum correntropy criterion (MCC) [8]:

$$\min J_{MCC} = \min \sum_{i=1}^{N} \rho_{MCC}(e_i) \tag{2.7}$$

Here $\rho_{MCC}(e_i) = \left(1 - \frac{1}{\sqrt{2\pi}\sigma}\right) \exp\left(-\frac{e_i^2}{2\sigma^2}\right)$ is the score function of MCC. $\kappa_\sigma = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{e_i^2}{2\sigma^2}\right)$ is known as Gaussian kernel function.

3 Derivation of the Proposed Filter

The solution of CKF consists of two parts: the prediction from dynamic model and the correction from measurements.

$$\begin{aligned} \mathbf{x}_k &= \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_{k-1} \\ \mathbf{y}_k &= \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k \end{aligned} \tag{3.1}$$

where $\mathbf{x} \in R^n$ is the state variable, $\mathbf{y} \in R^m$ is the measurement variable, $\mathbf{F}$ is the state transformation matrix, $\mathbf{H}$ is the measurement matrix,$\mathbf{v}$ and $\mathbf{w}$ are uncorrelated Gaussian white noises, whose covariance is $\mathbf{Q}$ and $\mathbf{R}$ respectively. The solution of CKF from a Bayesian maximum likelihood estimation perspective can be obtained by minimizing the following cost function:

$$\hat{x}_k = \arg\min(\left\|x_k - \hat{x}_{k|k-1}\right\|^2_{P^{-1}_{k/k-1}} + \|H_k x_k - y_k\|^2_{R_k^{-1}}) \tag{3.2}$$

where $\|\mathbf{x}\|^2_{\mathrm{A}} = \mathbf{x}^T\mathbf{A}\mathbf{x}$, $\hat{x}_k$ and $\hat{x}_{k/k-1}$ are predictive state and updated state respectively. $\mathbf{P}_{k/k-1}$ is the covariance matrix of $\hat{x}_{k/k-1}$. Let $\mathbf{e}_{x,k} = \mathbf{P}^{-\frac{1}{2}}_{k/k-1}(x_k - \hat{x}_{k/k-1})$ and $\mathbf{e}_{y,k} = \mathbf{R}_k^{-\frac{1}{2}}(H_k x_k - y_k)$, the Eq. (3.2) can be written as following

$$\hat{x}_k = \arg\ \min\left(\sum_{i=1}^{n} \mathbf{e}^2_{x,k,i} + \sum_{j=1}^{m} \mathbf{e}^2_{y,k,j}\right) \tag{3.3}$$

where $\mathbf{e}_{x,k,i}$ is the ith component of $\mathbf{e}_{x,k}$ and $\mathbf{e}_{y,k,j}$ is the jth component of $\mathbf{e}_{y,k}$. From Eq. (3.3), it can be noted that the ρ function for both prediction and measurement part of CKF is ℓ_2 norm:

$$\rho_{\ell_2}(e_i) = e_i^2 \tag{3.4}$$

To suppress the measurement outliers, the measurement part of the cost function above is modified with MCC as:

$$\hat{x}_k = \arg\ \min\left(\sum_{i=1}^{n} \rho_{\ell_2}(\mathbf{e}^2_{x,k,i}) + \sum_{j=1}^{m} \rho_{MCC}(\mathbf{e}^2_{y,k,j})\right) \tag{3.5}$$

According to the M-estimate theory, the solution can be obtained by following formula

$$\mathbf{P}^{-1}_{k/k-1}(\mathbf{x}_k - \hat{x}_{k/k-1}) + \mathbf{H}_k^T\mathbf{R}_k^{-T/2}\Psi_y\mathbf{R}_k^{-1/2}(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k) = 0 \tag{3.6}$$

where

$$\Psi_y = diag[\psi(\mathbf{e}_{y,k,j})] \quad \psi(\mathbf{e}_{y,k,j}) = \frac{1}{\sqrt{2\pi\sigma^3}}\exp\left(-\frac{\mathbf{e}_{y,k,j}^2}{2\sigma^2}\right) \tag{3.7}$$

Define

$$\breve{\mathrm{R}}_k = \mathbf{R}_k^{\frac{1}{2}}\Psi_y^{-1}\mathbf{R}_k^{\frac{T}{2}} \tag{3.8}$$

Equation (3.6) can be reformed as:

$$\mathbf{P}_{k/k-1}^{-1}(x_k - \hat{x}_{k/k-1}) + \mathbf{H}_k^T\breve{\mathrm{R}}_k^{-1}(H_k x_k - y_k) = 0 \tag{3.9}$$

which is the solution for the following minimization problem:

$$\hat{x}_k = \arg\min(\left\|x_k - \hat{x}_{k/k-1}\right\|^2_{\mathbf{P}_{k/k-1}^{-1}} + \|H_k x_k - y_k\|^2_{\breve{\mathrm{R}}_k^{-1}}) \tag{3.10}$$

It can be seen that Eq. (3.10) is similar with Eq. (3.2). The only difference is the reformed measurement error covariance matrix $\breve{\mathrm{R}}_k$ expressed in (3.8). The explicit solution of Eq. (3.10) is that:

$$\hat{x}_k = \hat{x}_{k/k-1} + \mathbf{K}_k[\mathbf{y}_k - H_k\hat{x}_{k/k-1}]\mathbf{K}_k \tag{3.11}$$

$$\mathrm{K}_k = \mathbf{P}_{k/k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k/k-1}\mathbf{H}_k^T + \breve{\mathrm{R}}_k)^{-1}\widehat{\mathbf{P}}_k \tag{3.12}$$

$$\hat{\mathrm{P}}_k = (\mathbf{I}_n - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k/k-1} \tag{3.13}$$

It can be observed that the solution from (3.11) to (3.13) is similar with the solution from CKF except the modified measurement covariance matrix $\breve{\mathrm{R}}_k$ obtained by MCC, which is robust against measurement outliers. Therefore, in this paper, we call this robust filter as MKF. When the Ψ_y equals to identity matrix, the MKF degrades to the CKF. Note that, in calculating the weighting matrix Ψ, a solution of the filter should be know in advance. In this paper, the first solution is set as $\hat{\mathrm{x}}_{k/k-1}$ at the assumption that the dynamic model is consistent with the real movement of the vehicle. By iterations, a converged solution can be obtained because the ψ function is nonincreasing [9].

4 Performance Analysis of MKF, HKF and CKF

The popular Huber function is as follows:

$$\rho_{Huber}(e) = \begin{cases} 0.5e^2, & |e| \le \gamma \\ \gamma|e| - 0.5\gamma^2, & |e| > \gamma \end{cases} \tag{4.1}$$

where γ is the modulating parameter, which is generally set between 1.0 and 2.0 [1]. It can be seen that the Huber function is a combination of ℓ_1-norm and ℓ_2-norm. Recalling from the M-estimation theory, the weighting function of Huber function is

$$\psi_{Huber}(e) = \begin{cases} 1, & |e| \le \gamma \\ \frac{\gamma}{|e|}, & |e| > \gamma \end{cases} \tag{4.2}$$

Recalling the score function of MCC and the ℓ_2 norm function employed by CKF, their weighting functions are

$$\psi_{MCC}(e) = \frac{1}{\sqrt{2\pi\sigma^3}} \exp\left(-\frac{e^2}{2\sigma^2}\right) \tag{4.3}$$

$$\psi_{l_2}(e) = 1 \tag{4.4}$$

The parameter $\gamma = 1.345$ is suggested when the measurement noise is Gaussian. To pursue the Gaussian efficiency, σ is set to be 3.75. However, in this situation, the weight is very small even if the measurement error is zero. And when the error become a little large, the weight is so tiny that the more complicated information filter has to be employed. In this paper, to make the weight value to be 1 when the error is zero, $\sigma = (2\pi)^{-1/6}$ is suggested according to Eq. (4.3).

Suppose only one measurement is obtained at a time. Figure 1 shows the weighting functions of MCC, Huber and ℓ_2 norm with different errors in which $\gamma = 1.345$, $\sigma = (2\pi)^{-1/6}$. Respectively. From Fig. 1, it can be seen that, with the increasing of errors, the weight function of MCC decreases, which means that the filter can reduce the weight of the measurements with large errors. So does HKF when the errors exceed the bound γ. That's why MKF and HKF have robustness against measurement outliers. One advantage of MKF over HKF in resisting measurement outliers is that the MKF can totally discarding the measurements with large errors, while the HKF still employs partially the measurements with large errors. As for the weight function of ℓ_2 norm function, it remains unchanged no matter what value of the errors. Therefore the ℓ_2 norm function based CKF is susceptible to measurement outliers. In addition, it can be seen that the weight of MCC is smaller than that of Huber function and ℓ_2 norm function with same errors all the time. It means that the measurement part of MKF contributes less to the estimation results than that of HKF and CKF. Looking backward of Eq. (3.11), the

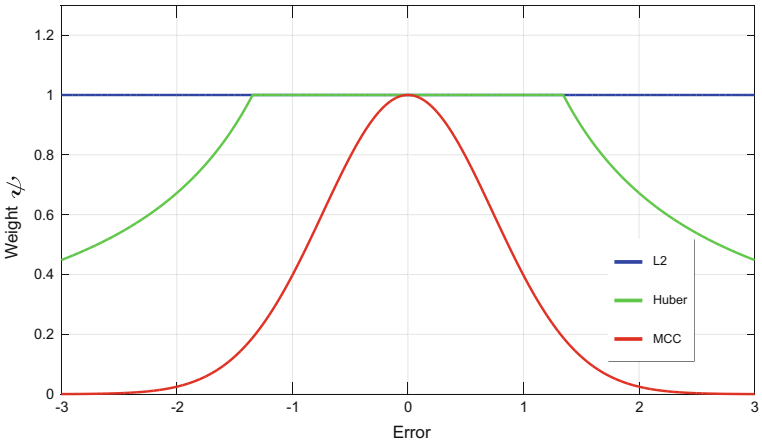


Fig. 1 Weighting functions of MCC, Huber and $\ell 2$ norm

filtering results rely on two parts: prediction and measurement. So the MKF will have a better performance in estimations when the dynamic model is consistent with the real movement of the vehicle.

5 Simulation and Tests

In this section, the performance of the MKF is verified by numerical simulations and GNSS static positioning results. The constant velocity (CV) model is used. Under CV model, the state variable $\mathbf{x} = [p, v]$ which means position and velocity, the only measurement is the position. The state transformation matrix $\mathbf{F}$ and the measurement matrix $\mathbf{H}$ is therefore expressed as:

$$\mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad \mathbf{H} = [1 \quad 0] \tag{5.1}$$

where T is the sampling intervals.

5.1 Numerical Simulations

In some engineering applications, the measurement noise is contaminated to be not Gaussian distribution, but a mixture of Gaussian distribution or heavy-tailed distribution defined by the probability density function (PDF) like:

$$f(e_k) = \left(\frac{1-\varepsilon}{\sigma_1\sqrt{2\pi}}\right)\exp\left[-\left(\frac{e_k^2}{2\sigma_1^2}\right)\right] + \left(\frac{\varepsilon}{\sigma_2\sqrt{2\pi}}\right)\exp\left[-\left(\frac{e_k^2}{2\sigma_2^2}\right)\right] \tag{5.2}$$

where σ_1 and σ_2 are the standard deviations of the individual Gaussian distributions, and ε is a perturbing parameter that represents error model contamination. In this part, this kind contaminated noises are modeled as measurement errors to test the effectiveness of the proposed filter. Here let $\sigma_1 = 1$ and $\sigma_2 = 5$. To model the worst case, ε is set to be 0.5. Under the PDF (5.2), the parameter $\gamma = 0.4363$ is optimal for HKF [2, 7].

A vehicle moves 1500 s with constant velocity is modeled. The prior measurement covariance is set to be 9, which means that the measurement error is 3 m. 100 Monte Carlo simulations are carried out. The root mean square error (RMSE) of every epoch about position and velocity is presented as in Fig. 2. It can be obtained that the MKF is better than HKF and HKF performs better than CKF in estimation performance. The reason is explained in Sect. 4.

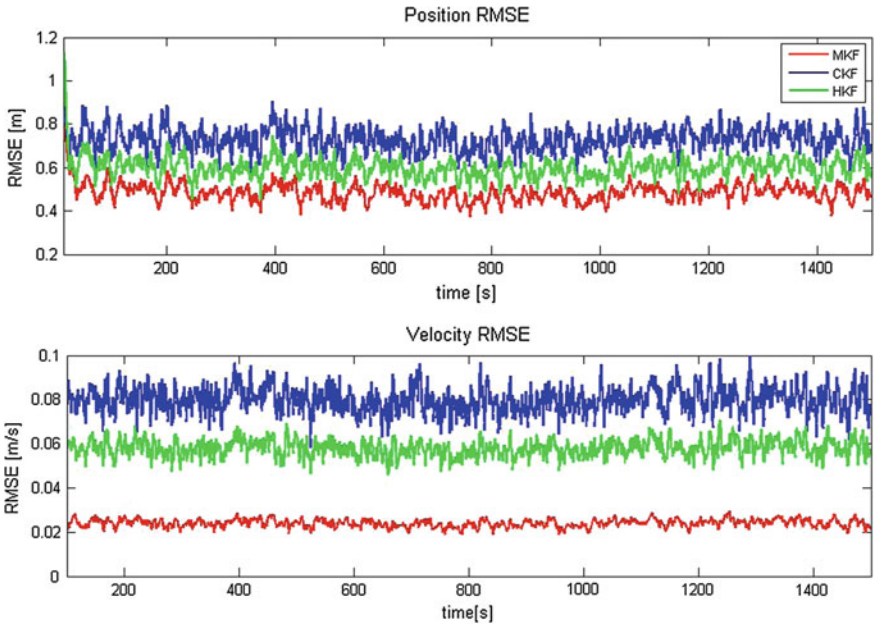


Fig. 2 RMSE comparison of the filters under mixture Gaussian distribution

5.2 GNSS Static Positioning Tests

In this part, the data was collected from HD8020 chip of Allystar Technology (Beijing) CO. Limited. The HD8020 chip supports GPS and GNSS multiple systems positioning. The antenna is located on the roof of the building of HuaDa Electronics Design CO. Limited. The available measurements are C/A code with 1 s sampling rate. The single point positioning results are processed. For simplicity, only the x-axis result is presented. To investigate the outliers' resistance of the proposed algorithm, four measurement outliers are added to the measurements from single point positioning results: $y_{100} = y_{100} - 25$, $y_{500} = y_{500} + 35$, $y_{900} = y_{900} - 45$, $y_{1300} = y_{1300} + 55$. The positioning results obtained by CKF, HKF and MKF are depicted as in Fig. 3, from which it can be observed that the MKF (red line) has better performance than HKF (green line) and CKF (blue line) in both position and velocity estimations. And the MKF and HKF has robustness against measurement outliers (in 100, 500, 900 and 1300 s), while the CKF is susceptible to measurement outliers. In addition, it can be observed that, the positioning results of HKF almost coincide with that of CKF except the measurement outliers epoch. This is expected because when the measurement standard errors are in the scope of $\gamma = 1.345$, the weight equals to 1, which can be seen from Fig. 1. In this situation, the HKF degrades to the CKF.

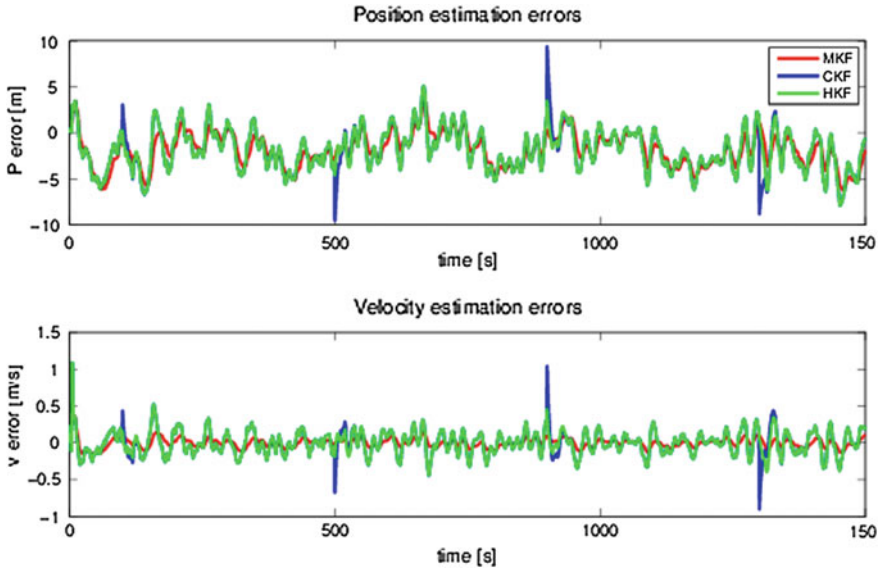


Fig. 3 Comparison of the filters' performance for static GNSS positioning

6 Conclusion

In this paper, the MCC and M-estimate theory based MKF is proposed for state estimations. The performance of MKF, HKF and CKF is analyzed and why MKF performs better is explained. By simulations and real GNSS static positioning, the performance of the proposed filter is illustrated.

References

1. Huber PJ (1964) Robust estimation of a location parameter. Annal Mathem Stat 35(2):73–101
2. Huber PJ (1981) Robust statistics. Wiley, New York, pp 43–106, 153–198
3. Yang YX (1991) Robust Bayesian estimation. Bull Geod 65:145–150
4. Koch KR, Yang YX (1998) Robust Kalman filter for rank deficient observation models. J Geod 72:436–441
5. Yang Y, He H, Xu G (2001) Adaptively robust filtering for kinematic geodetic positioning. J Geod 75:109–116
6. Principe JC (2010) Information theoretic learning-Renyi's entropy and kernel perspectives. Springer, New York, pp 385–455
7. Wang YD, Zheng W, Sun SM, Li L (2016) Robust information filter based on maximum correntropy criterion. J Guid Control Dyn 39(5):1126–1131
8. Liu WF, Pokharel PP, Jose CP (2007) Correntropy: properties and applications in non-Gaussian signal processing. IEEE Trans Signal Proc 5(11):5286–5298
9. Karlgaard Christopher D, Hanspeter S (2007) Huber-based divided difference filtering. J Guid Control Dyn 30(3):885–891

Gradient Boost Decision Tree Fingerprint Algorithm for Wi-Fi Localization

Yanxu Liu, Zhongliang Deng and Lu Yin

Abstract Location Based Services (LBS) is require the indoor and outdoor seamless positioning that providing real-time, stable and high-accuracy localization and navigation. Mobile devices can be positioning by using Wi-Fi signals based on correlation between Wi-Fi signals strength and coordinates. And Wi-Fi signals are common in modern buildings, so there needn't deploy equipment. But there are still some drawbacks, such as poorly positioning accuracy and too long online computing time during using Wi-Fi signals to localization. For this reason, we proposed a Gradient Boosting Decision Tree (GBDT) fingerprint algorithm for Wi-Fi localization, this algorithm adopt a linear combination of multiple decision trees to obtain an approximate model of the coordinates and received signal strength (RSS). Experiment shows that about 13% increases in positioning accuracy and 65% reduces in online computation time compares with AdaBoost-based algorithm.

Keywords LBS · RSSI · Fingerprint algorithm · Gradient boosting decision tree

1 Introduction

Nowadays, the development of smart devices (e.g., tablet computer, smartphones) plays an important part in the growth of fields, such as Location Based Services (LBS) and Internet of Things Applications (IOTA) [1]. About three-quarters (74%) of smartphone owners are active uses of LBS [2] (Pew Research 2013). Indoor localization information plays an important part in people's daily life, which improving people's quality of lifestyle.

However, indoor positioning technologies are still difficult in providing the same level of positioning accuracy, continuity and reliability as outdoor [3]. Global Navigation Satellite Systems (GNSS) are frequently used for outdoor positioning

Y. Liu (✉) · Z. Deng · L. Yin
School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
e-mail: liuyx@bupt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_44

and it work well. Unfortunately, GNSS technology is hardly operational indoor due to heavy signal attenuation and block by the buildings. Such flaws will lead to poor positioning accuracy and meet the demand of indoor positioning impossibly. And because of disturbed by furniture, pedestrians and walls, the indoor signals are easier interfering by non-line-of-sight (NLOS) and multi-path. Therefore, GNSS cannot achieve the seamless localization of indoor and outdoor. Most public places such as shopping malls, stations, schools etc. have deployed multiple Wi-Fi nodes, so there is no necessary for additional equipment to locate by Wi-Fi signals.

Bahl and Padmanabhan [4] proposed the RADAR system, which use the access points (AP) in wireless local area networks (WLANs), is the first system using fingerprint to positioning. The RADAR system adopts the nearest neighbor method to localize and the accuracy of positioning is poor [5]. Feng et al. proposed the improved adaptive boosting (AdaBoost) fingerprint algorithm for Wi-Fi indoor localization, was improved positioning accuracy by removing the isolated points from the training set data [6]. Simon et al. from the bottom layer of the wireless fingerprint positioning, compared the method of fingerprint map, grid size and the density of the AP on the positioning accuracy [7].

To address aforementioned shortcomings, we proposed Gradient Boosting Decision Tree algorithm to realize the matching of the RSS with the coordinate. Compared with the AdaBoost-based method, GBDT has lower sensitivity of noise and shorter online computing time.

The remainder of this paper is organized as follows: A detailed description of fingerprint-based indoor localization system is presented in Sect. 2. In Sect. 3, the experimental results and evaluation are listed. Finally, Sect. 4 concludes the paper.

2 Fingerprint-Based Indoor Localization System

The proposed fingerprint algorithm consists of two phase: an offline phase and an online phase. During the offline phase, RSS readings and its orientation that the mobile device are facing are collected on grid of reference points(RP), and then the RSS readings and its orientation are stored in fingerprint database. The online phase consists of the mobile devices measuring RSS, and positioning estimation by proposed algorithm.

2.1 Building Fingerprint Map

In the offline training phase, the RSS readings are collected at known positions, which referred to as the reference points (RPs), by pointing the mobile devices to four diverse orientations (north, south, east, west). The raw set of RSSI samples collected form AP i at RP j with orientation o denote as $\gamma_{i,j}^{(o)}(t)$, $t = 1, 2, \ldots, T$,

where the T representing the total number of time RSSI collected. Then, the average of total samples at one RP in diverse orientations is calculated and stored in fingerprint map. Such fingerprint map can be represented by $Map^{(o)}$:

$$Map^{(o)} = \begin{bmatrix} \gamma_{1,1}^{(o)} & \gamma_{1,2}^{(o)} & \cdots & \gamma_{1,N}^{(o)} \\ \gamma_{2,1}^{(o)} & \gamma_{2,2}^{(o)} & \cdots & \gamma_{2,N}^{(o)} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{L,1}^{(o)} & \gamma_{L,2}^{(o)} & \cdots & \gamma_{L,N}^{(o)} \end{bmatrix}$$

where $\gamma_{i,j}^{(o)} = \frac{1}{T}\sum_{t=1}^{q} \gamma_{i,j}^{(o)}(t)$ denotes the average of RSSI over time domain from AP i at RP j with orientation o, for $i = 1, 2, \ldots, L$, $j = 1, 2, \ldots, N$, and $o \in \{north, sourth, east, west\}$. L is the total number of RPs in the localization area and the N is the number of APs that can be detected in the building. The algorithm chooses the maximum RSSI value in RP at diverse orientation as training set during the offline training phase. During the online localization phase, RSSI measurement vector is denote as:

$$\mathbf{X}_r = [X_{1,r}, X_{2,r}, \ldots, X_{L,r}]$$

where $\{\mathbf{X}_{k,r}, k = 1, 2, \ldots, L\}$ is the RSS readings at RP r. Then, we choose the maximum RSS as the GBDT algorithm prediction of standard.

2.2 GBDT Algorithm

Classifier ensemble and feature selection are two common methods for improving classification performance in machine learning [8]. And the Gradient Boosting Decision Trees (GBDT) belong to the former. In theory, Classifier ensemble can boost strong learning algorithm from weak learning algorithm, and the weak leaning algorithm is only a little bit better than random guessing. And strong learner is linear combination of a series of weak learners. GBDT algorithm adopt decision tree as basis function, and it is a simple yet very effective method for learning linear and non-linear function by linear combination of a series of decision trees [8]. The model of GBDT is represent by a series of decision trees additive model [9]:

$$f_M(\mathbf{x}) = \sum_{m=1}^{M} DT_m(\mathbf{x})$$

where $DT_m(\mathbf{x})$ denotes the decision tree, and $\mathbf{x} = [x_1, x_2, \ldots, x_L]$ denotes the RSS vector. M is the number of decision tree.

During the offline training phase, the GBDT algorithm takes as input training set $\mathbf{T}$. We assumed $\mathbf{T} = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \ldots, (\mathbf{x}_N, \mathbf{y}_N)\}$, $\mathbf{x}_i \in \chi \subseteq \mathbb{R}^L$, $\mathbf{y}_i \in \zeta \subseteq \mathbb{R}^2$,

each $\mathbf{x}_i$ belongs to input space χ that spanned by column vectors of fingerprint map. $\mathbf{y}_i$ belongs to output space ζ and represents position coordinate of the 2 dimensional plane. and if the input space χ is divided into K disjoint regions $R_1, R_2, \ldots, R_K$, and each region has a constant output c_k, then the decision tree can be expressed as:

$$DT(\mathbf{x}) = \sum_{k=1}^{K} c_k I_{\mathrm{R_k}}(\mathbf{x})$$

where $I_{\mathrm{R_k}}(\mathbf{x})$ is an indicator function, defined as:

$$I_{R_k}(\mathbf{x}) = \begin{cases} 1 & if \quad \mathbf{x} \in R_k \\ 0 & else \end{cases}$$

We recursively construct a binary decision tree to generate each decision tree by Least Squares Regression Tree algorithm in the Classification and regression tree (CART) algorithm. A regression tree corresponds to a partition R_k of input space and the output value c_k on a partition cell. And we take square error $\sum_{\mathbf{x}_i \in R_k} (\mathbf{y}_i - f(\mathbf{x}_i))(\mathbf{y}_i - f(\mathbf{x}_i))^T$ as loss function after determining input space, and then finding the optimal output value on each cell with making square error minimum. Obviously the best value of $\hat{c}_k$ on a cell R_k is the average of the output $\mathbf{y}_i$ corresponding to all the input instances $\mathbf{x}_i$ of the cell R_k, so the $\hat{c}_k$ becomes:

$$\hat{c}_k = \frac{1}{K} \sum_{i=1}^{K} (\mathbf{y}_i | \mathbf{x}_i \in R_k)$$

We choose the jth input data $\mathbf{x}^{(j)}$ and the value s as a splitting variable and splitting point, and the two regions define as:

$$R_1(j, s) = \{\mathbf{x} | \mathbf{x}^{(j)} \leq s\}, R_2(j, s) = \{\mathbf{x} | \mathbf{x}^{(j)} > s\}$$

Then the optimal splitting variable and splitting point is found by solving Eq. (1).

This leads to the following algorithm for Least squares regression tree:

Algorithm 1: Least squares regression tree
Input: training set D
Output: regression tree $DT(\mathbf{x})$

(1) Traverse the variable j, scanning splitting point s by fixing splitting variable j, and choose (j, s) to make the Eq. (1) minimum:

$$\min_{j,s}\left[\min_{c_1}\sum_{x_i\in R_1(j,s)}(y_i-c_1)(y_i-c_1)^T+\min_{c_2}\sum_{x_i\in R_2(j,s)}(y_i-c_2)(y_i-c_2)^T\right] \tag{1}$$

(2) Dividing the region by (j, s) and calculating the output value

$$R_1(j,s)=\{\mathbf{x}|\mathbf{x}^{(j)}\le s\}, R_2(j,s)=\{\mathbf{x}|\mathbf{x}^{(j)}>s\}$$

$$\hat{c}=\frac{1}{N_k}\sum_{\mathbf{x}_i\in R_k(j,s)}\mathbf{y}_i,\ \mathbf{x}\in R_k(j,s),\ k=1,2$$

(3) Continue with two sub-regions using steps (1), (2) until the stop condition is met.
(4) The input space is divided into $R_1, R_2, \ldots, R_K$,and the decision tree as:

$$DT(\mathbf{x})=\sum_{k=1}^{K}\hat{c}_k I_{R_k}(\mathbf{x})$$

end Algorithm

At each iteration of the algorithm, a new decision tree model is built based on the residuals of the previous decision tree. The residual can be calculated by Eq. (2). But for the general loss function, the algorithm uses the value of the negative gradient of the loss function in the current model as an approximation of the residual, as shown in Eq. (3). The GBDT algorithm flow is shown in Algorithm 2.

$$\mathbf{r}=\mathbf{y}-f_{m-1}(\mathbf{x}) \tag{2}$$

Algorithm 2: Gradient Boosting Decision Trees

$$f_0(\mathbf{x})=arg\min_{c}\sum_{i=1}^{N}L(\mathbf{y}_i,c)$$

For m = 1 to M do:

$$\mathbf{r}_{mi}=-\left[\frac{\partial L(\mathbf{y}_i,f(\mathbf{x}_i))}{\partial f(\mathbf{x}_i)}\right]_{f(\mathbf{x})=f_{m-1}(\mathbf{x})},\ i=1,2,\ldots,N \tag{3}$$

$DT_m(\mathbf{x}) = \sum_{k=1}^{K} c_{mk} I_{R_{mk}}(\mathbf{x})$, (using algorithm 1 to calculate c_{mk} and R_{mk} by input training set is $(\mathbf{x}_i, \mathbf{r}_i), i = 1, 2, \ldots, N$)

$$f_m(\mathbf{x}) = f_{m-1}(\mathbf{x}) + DT_m(\mathbf{x})$$

endFor
end Algorithm

During the online localization phase, the L-dimensional vector $\mathbf{X}_r$ of the RSSI collected by the mobile terminal is substituted into the $f_M(\mathbf{x})$, and the $f_M(\mathbf{X}_r)$ is the current position coordinates.

3 Experiment Result and Evaluation

3.1 Experiment Platform Building

During the offline training phase, fingerprint map was collected by the Signal-acquiring device shown in Fig. 1, which measures the RSSI of APs by RAK476 module. The positioning area was selected at the room 1015, National University Science Park, Beijing University of Posts and Telecommunications, the total dimension of the room was 5.6 × 7.2 m. The room is equipped with Wi-Fi environment. In this experiment environment, we collected 13 reference point in the centrality of each grid, and select the maximum value. In 4 orientations, the RSSI was collected from 6APs for a period of 120 s over 13 RPs with the average grid spacing of 1.7 m. And determine the orientation of the Signal-acquiring device by

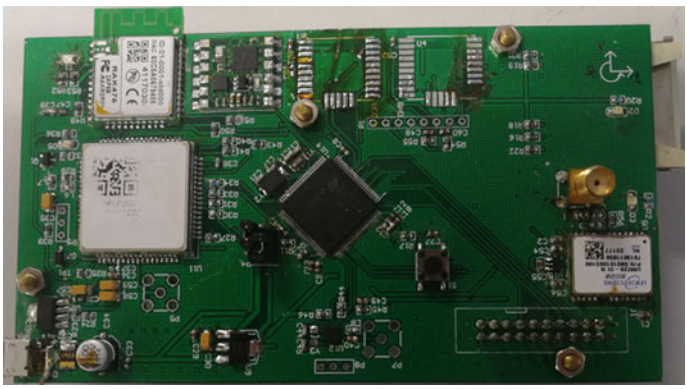


Fig. 1 Signal-acquiring device

compass sensor of JY901 module. In the online localization phase, 30 independent known points were select as test points to evaluate the performance of the GBDT algorithm by Python 2.7 on Tsinghua Tongfang Computer (operating system Arch Linux, i5-6400 CPU).

3.2 *Performance of the GBDT Algorithm*

In this paper, we compared the computing time and positioning accuracy of GBDT with NN, KNN and AdaBoost algorithms in the same data. We collected 20 RSSI data at each test point.

Figure 2 illustrates that the cumulative distribution function (CDF) of the positioning errors for the four algorithms. From Fig. 2 and Table 1, we can get the localization accuracy of GBDT fingerprint algorithm with the probability of 67% at 2.05 m, and the localization accuracy is the highest among the four algorithms. And compared with AdaBoost algorithm, the localization accuracy of GBDT algorithm is improved by about 13%.

Table 2 shows the calculation time of the four algorithms in the offline phase and the online phase.

It can be seen that the GBDT algorithm takes longer computing time at offline phase relative to the other three algorithms. On the contrary, computing time at

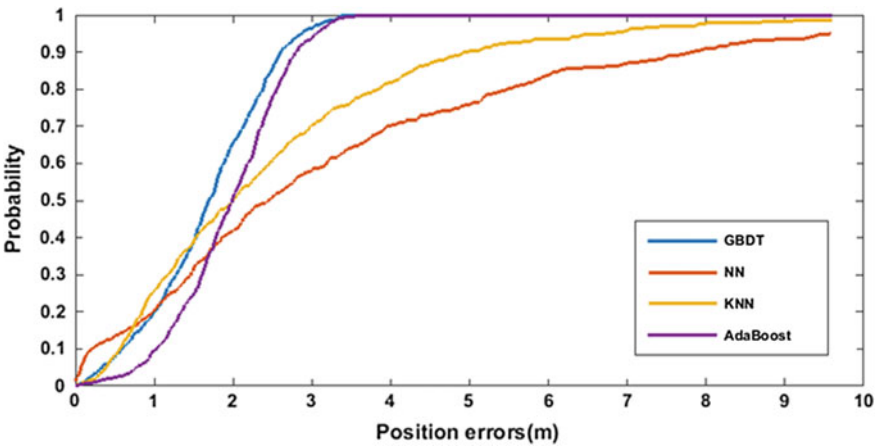


Fig. 2 CDF of localization errors with diverse algorithms

Table 1 Localization accuracy with the probability of 67%

	GBDT	AdaBoost	KNN	NN
Localization errors (m)	2.05	2.306	2.787	3.748

Table 2 Calculate of the time consumption

	Offline phase (s)	Online phase (ms)
GBDT	35.5	35.8
AdaBoost	9.39	55.1
KNN	0.0	16.1
NN	0.0	9.3

online localization phase is shorter than other algorithms. The online computation time of the GBDT algorithm is greater than NN, KNN algorithms, but the corresponding localization accuracy has greatly improved.

4 Conclusion

We proposed GBDT fingerprint algorithm for WiFi localization in this paper, that improved the accuracy of the position estimation. The proposed algorithm boosts the decision tree to a strong predictor with stronger predictive ability. Compared with the based on AdaBoost fingerprint algorithm for positioning, location accuracy of the algorithm improved about 13%, and online positioning calculation time reduced to about 60% of the former. In the future, we plan to improve the algorithm to use in the bigger fingerprint map and more complex building with many floors.

Acknowledgements The work was supported by The National Key Research and Development Program of China (Grant: 2016YFB0502001).

References

1. Huang X, Guo S, Wu Y, Yang Y (2017) A fine-grained indoor fingerprinting localization based on magnetic field strength and channel state information. Pervasive Mob Comput 41:150–165
2. Zickuhr K (2013). Location-based services [online]. Pew Research Center, Internet, Science & Tech. Available at: http://www.pewinternet.org/2013/09/12/location-based-services/. Accessed on 30 Nov 2017
3. Basiri A, Lohan E, Moore T, Winstanley A, Peltola P, Hill C, Amirian P, Figueiredo e Silva P (2017) Indoor location based services challenges, requirements and usability of current solutions. Comput Sci Rev 24:1–12
4. Bahl P, Padmanabhan V (n.d.) RADAR: an in-building RF-based user location and tracking system. In: Proceedings IEEE INFOCOM 2000 conference on computer communications, nineteenth annual joint conference of the IEEE computer and communications societies (Cat. No. 00CH37064)
5. Dawes B, Chin K (2011) A comparison of deterministic and probabilistic methods for indoor localization. J Syst Softw 84(3):442–451
6. Feng Y, Minghua J, Jing L, Xiao Q, Ming H, Tao P, Xinrong H (2014) Improved AdaBoost-based fingerprint algorithm for Wi-Fi indoor localization. In: 2014 IEEE 7th joint international information technology and artificial intelligence conference

7. Yiu S, Dashti M, Claussen H, Perez-Cruz F (2017) Wireless RSSI fingerprinting localization. Sig Process 131:235–244
8. Zhang C, Liu C, Zhang X, Almpanidis G (2017) An up-to-date comparison of state-of-the-art classification algorithms. Expert Syst Appl 82:128–150
9. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 25:1189–1232

The Geomagnetic Filtering Algorithm Based on Correlative Probability Density Add-Weight

Yang Chong, Hongzhou Chai, Wei Liu, Yulong Kong, Zongpeng Pan and Jie Chen

Abstract To overcome the linearizated errors from the state equation and observation equation based on Extended Kalman Filter (EKF), a Unscented Kalman Filter (UKF) matching algorithm using geomagnetic anomaly based on probability weighted was proposed. For the problem that the quasi observations might be arose by choosing the same weight coefficient as the UT transformation, the geomagnetic anomaly UKF filtering algorithm associated with probability density function to assign weight for the sampled observation has been researched. The two experiments have been carried out in the South China Sea from the Earth Magnetic Anomaly Grid 2 (EMAG2), it is shown that the problem mentioned above could be overcome based on probability weighted, the drifting errors of inertial navigation system in longitude and latitude can be reduced by the modified algorithm, and the positioning accuracy and reliability of the modified algorithm is obviously superior to that of the traditional UKF algorithm and the Inertial Navigation System (INS).

Keywords Correlative probability density · Geomagnetic aided inertial navigation Unscented kalman filtering · Add-weight · Nonlinear

1 Introduction

Autonomous navigation technology of long duration and high-precision becomes the main factor that restricts the development of autonomous underwater vehicle (AUV). Inertial navigation system (INS) has been core navigation equipment for

Y. Chong (✉) · H. Chai · W. Liu · Y. Kong · Z. Pan
Information Engineering University, Zhengzhou 450001, China
e-mail: chongyang_geodesy@outlook.com

Y. Chong
State Key Laboratory of Geo-Information Engineering, Xi'an, China

J. Chen
Department of Navigation, Naval University of Engineering, Wuhan, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_45

AUV. However, due to the existing shortcomings that the navigation error increases with time, the INS cannot meet the AUV's mission. Geomagnetic matching technology can provide passive external resource of correction information for AUV, which improves the INS's long voyage navigation accuracy. Hence this technology has been widely studied by many researchers in recent years. It is focused on key techniques of underwater geomagnetic matching navigation, including geomagnetic field mapping construction, correction of geomagnetic surveying data, match-filtering algorithm, and suitability evaluation on geomagnetic map. Matching algorithm is the core of the geomagnetic aided inertial navigation system. The observation noise is caused by uncertain problems, and higher demand in the robustness and usability of matching algorithm put forward [1, 2].

Filter matching algorithm has been put into practical in gravity matching navigation, such as general gravity module (UGM) in matching computer. The essence is a kind of optimal estimation algorithm based on extended kalman filter. Extending Kalman filtering is the most common method for solving nonlinear filter problems [3]. EKF algorithm was first used by Psiaki to conduct low orbit spacecraft based on magnetic scalar strength [4]. NASA Goddard space center carried out a lot of work on underwater geomagnetic navigation, and completed the verification test on the ground in great quantities. EKF algorithm was used to estimate the attitude and position of the spacecraft, and the precision reaches 1 and 4 km [5]. Armstrong, from the Idaho University, has adopted EKF algorithm to improve the navigation accuracy during geomagnetic disturbance, and verified the effectiveness of the algorithm through a series of simulations and experiments [6]. Wang and Tian [7] used Matlab to simulate the autonomous navigation system based on the geomagnetic field, and introduced EKF algorithm, so that the navigation accuracy of the system with geomagnetic field vector was significantly improved. However, when the nonlinear system is more serious, the estimation of EKF algorithm is not accurate, and the computation amount is relatively large. The Unscented Kalman filtering is also an extension of Kalman filter, which is an improved algorithm based on EKF algorithm. The error effect caused by the linearization of EKF was reduced and the linear distribution was approximated by sampling method [8]. In recent years, UKF algorithm and adaptive UKF algorithm have been successfully applied to GPS/INS combined navigation and gravity anomaly matching navigation [9–13].

When considering predicted observations, using the same weight coefficients as state prediction, the selection criteria of weight coefficient in the state prediction are based on the Sigma sampling sample and state random quantity have the same mean and variance. In the field of geomagnetic anomaly matching, the analytical model of geomagnetic field usually does not have a fixed form, which would lead to the predicted observation value being the quasi observation. According to the above problem, from the perspective of probability density, this paper puts forward a geomagnetic anomaly UKF filtering algorithm with weighted sampling observation value of correlation probability density, which improves the accuracy of the UKF filtering solution.

2 Methodology of UKF Algorithm

By the inertial navigation error equation, the filter equation is given,

$$\boldsymbol{X}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{X}_{k-1} + \boldsymbol{W}_k \tag{1}$$

where state vector is $\boldsymbol{X} = [\delta\varphi,\delta\lambda,\delta h,\delta v_e,\delta v_n,\delta v_u,\delta\theta,\delta\gamma,\delta\psi]^T$, containing position errors of Autonomous Underwater Vehicles (AUVs) ($\delta\varphi,\delta\lambda,\delta h$), three direction of the velocity error ($\delta v_e,\delta v_n,\delta v_u$), and attitude error ($\delta\theta,\delta\gamma,\delta\psi$). $\boldsymbol{\Phi}_{k,k-1}$ is the state transition matrix, $\boldsymbol{W}_k$ is matrix of process noise.

To establish the observation equation related to the geomagnetic anomaly, it is important to establish analytical function mode of local geomagnetic field. It is a good choice that double quadric model is proposed to approximate the local geomagnetic field [14], and the observation equation of filtering as below

$$\Delta T(\varphi,\lambda) = \sum_{n=0}^{2}\frac{1}{n!}\sum_{r=0}^{n} a_{nr}C_n^r(\varphi-\varphi_0)^{n-r}(\lambda-\lambda_0)^r \tag{2}$$

where (φ_0,λ_0) is waiting-for-matched point. The coefficients are obtained by least square method.

The basic idea of UKF algorithm is that the amount of some special sample points (Sigma point) are produced by the mean and variance of estimated state quantity. Symmetric sampling strategy of sigma points are $2m+1$, and m is the dimensions of state vector.

Calculation steps of UKF as below,

(1) State parameter initialization

$$\hat{\boldsymbol{X}}_0 = E[\boldsymbol{X}_0] \tag{3}$$

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{X}}_0} = E[(\boldsymbol{X}_0-\hat{\boldsymbol{X}}_0)(\boldsymbol{X}_0-\hat{\boldsymbol{X}}_0)^T] \tag{4}$$

(2) Calculation of Sigma point

$$\boldsymbol{\chi}_{k-1} = [\hat{\boldsymbol{X}}_{k-1},\hat{\boldsymbol{X}}_{k-1}+\sqrt{(m+\lambda)\boldsymbol{\Sigma}_{\hat{\boldsymbol{X}}_{k-1}}},\hat{\boldsymbol{X}}_{k-1}-\sqrt{(m+\lambda)\boldsymbol{\Sigma}_{\hat{\boldsymbol{X}}_{k-1}}}] \tag{5}$$

where m is the number of state parameter, $\lambda = a^2(m+\kappa)-m$ is scale factor. a is the distance between Sigma point and mean value. κ is constant, and its value is $\kappa = 3-m$.

(3) Time update

$$\bar{\boldsymbol{\chi}}_k = f(\boldsymbol{\chi}_{k-1}, \boldsymbol{w}_k) \tag{6}$$

$$\bar{\boldsymbol{X}}_k = \sum_{i=0}^{2m} W_i^m \bar{\boldsymbol{\chi}}_{k,i} \tag{7}$$

$$\boldsymbol{\Sigma}_{\bar{\boldsymbol{X}}_k} = \sum_{i=0}^{2m} W_i^c (\bar{\boldsymbol{\chi}}_{k,i} - \bar{\boldsymbol{X}}_k)(\bar{\boldsymbol{\chi}}_{k,i} - \bar{\boldsymbol{X}}_k)^T + \boldsymbol{\Sigma}_{\boldsymbol{w}_k} \tag{8}$$

where $\bar{\boldsymbol{X}}_k$ is forecast values of state vector, and $\boldsymbol{\Sigma}_{\bar{\boldsymbol{X}}_k}$ is covariance matrix of state vector. Weight coefficient of UT conversion should meet the given constraint conditions.

$$W_0^m = \lambda/(m+\lambda) \tag{9}$$

$$W_0^c = \lambda/(m+\lambda) + (1 - \alpha^2 + \beta) \tag{10}$$

$$W_i^m = W_i^c = 0.5/(m+\lambda) \quad i = 1, \ldots, 2m \tag{11}$$

(4) Measurement update

$$\boldsymbol{y}_k = g(\bar{\boldsymbol{\chi}}_k, \boldsymbol{e}_k) \tag{12}$$

$$\hat{\boldsymbol{L}}_k = \sum_{i=0}^{2m} \beta_i y_{k,i} \tag{13}$$

$$\boldsymbol{\Sigma}_{\tilde{L}_k \tilde{L}_k} = \sum_{i=0}^{2m} \beta_i (y_{k,i} - \hat{L}_k)(y_{k,i} - \hat{L}_k)^T + \boldsymbol{\Sigma}_k \tag{14}$$

$$\boldsymbol{\Sigma}_{X_k L_k} = \sum_{i=0}^{2m} \beta_i (\bar{\boldsymbol{\chi}}_{k,i} - \bar{\boldsymbol{X}}_k)(y_{k,i} - \hat{L}_k)^T \tag{15}$$

$$\hat{\boldsymbol{X}}_k = \bar{\boldsymbol{X}}_k + \boldsymbol{K}_k(\boldsymbol{L}_k - \hat{\boldsymbol{L}}_k) \tag{16}$$

$$\boldsymbol{\Sigma}_{\hat{X}_k} = \boldsymbol{\Sigma}_{\bar{X}_k} - \boldsymbol{K}_k \boldsymbol{\Sigma}_{\tilde{L}_k \tilde{L}_k} \boldsymbol{K}_k^T \tag{17}$$

3 Probabilistic Data Association Matching Algorithm

In the probabilistic data association, effective echo of given target at any time is often more than one under the noise environment, considering the influence of random factors. Each echoes are likely from the right echo, and the only distinction is the possibility of correct echo. Therefore, probability density function can indicate the nondeterminacy. From the Probability Data Association Filter (PDAF), state variable turn into state samples values after sigma sampling [15–17]. The determined forecast quantity of geomagnetic anomaly observation is based on the probability. Assuming that the collections of geomagnetic anomaly prediction in epoch k are defined as:

$$\Delta T(k) = \left\{\bar{L}_j(k)\right\}_{j=1}^{s(k)} \tag{18}$$

Meanwhile, assuming that $\theta_j(k)$ denotes $g_j(k)$ is correct event derived from target, and $\theta_0(k)$ denotes no one is correct event. Therefore, probability density function of geomagnetic anomaly prediction $\bar{L}_j(k)$ derived from certain event is as below.

$$\beta_j(k) = P(\theta_j(k)\big|\Delta T(k)) \tag{19}$$

$\left\{\theta_0(k), \theta_1(k), \ldots, \theta_{s(k)}(k)\right\}$ is an incompatible complete event of event space, and we obtain:

$$\sum_{j=0}^{s(k)} \beta_j(k) = 1 \tag{20}$$

$\bar{L}_j(k)$ is derived from expectation estimate of real observation under the condition that $\hat{L}_j(k|k)$ denotes event $\theta_j(k)$ occurrence.

$$\hat{L}_j(k|k) = E(L(k)|\theta_j(k), \Delta T(k)) \tag{21}$$

From law of total probability,

$$\begin{aligned} \hat{L}(k|k) &= E(L(k)|\Delta T(k)) \\ &= \sum_{i=0}^{s(k)} E(L(k)|\theta_j(k), \Delta T(k) P(\theta_j(k)\big|\Delta T(k)) \\ &= \sum_{j=0}^{s(k)} \hat{L}_j(k|k)\beta_j(k) \end{aligned} \tag{22}$$

where β_j denotes the probability of determined sampling values from actual measured value. In practical engineering applications, each predicted observations may be derived from measured value, and the reason that leads to this difference should be the possibility. Related degree of two geomagnetic anomalies are chose to determine the probability value β_j. Relative value is smaller, the corresponding sampling values from the probability of measured values is higher.

From Eq. (20) we obtain:

$$s\beta_j = 1/p\_J(\lambda_j, \varphi_j) \tag{23}$$

$$sums\beta_j = \sum_{j=1}^{s(k)} s\beta_j \tag{24}$$

$$\beta_j = s\beta_j / sums\beta_j \tag{25}$$

where $p\_J(\lambda_j, \varphi_j)$ is related degree of observation forecast $\bar{L}_j$ and observation L_k, which is determined by Mean Absolute Deviation (MAD) algorithm.

$$p\_J_{MAD}(\lambda_j, \varphi_j) = \left\| \bar{L}_j - L_k \right\| \tag{26}$$

Therefore, based on the weighted forecast observation from the above, it can ensure the reliability of observation forecast in a certain extent. The problem that the false observations might be arose by choosing the same weight coefficient as the UT transformation is mitigated.

4 Experiments and Results Analysis

4.1 Geomagnetic Anomaly Grid Map

Geomagnetic background databases are the most fundamental elements of the geomagnetism aided navigation system. In this research, geomagnetic anomaly grid map is selected as the databases. The Earth Magnetic Anomaly Grid (EMAG2) was released by National Geographic Data Center (NGDC) in March 2009; EMAG2 is specified as a global 2-arc-minute resolution grid of the anomaly of the magnetic intensity. It was compiled from satellite, marine, aeromagnetic and ground magnetic surveys [18]. Therefore, the EMAG2 geomagnetic anomaly grid map with $2' \times 2'$ resolution was chosen for geomagnetism aided navigation in this paper. And the experimental region is selected from 10° to 15° in latitude, and from 108° to 113° in longitude which is situated in the South China Sea. Geomagnetic anomaly maps for simulation is illustrated in Fig. 1.

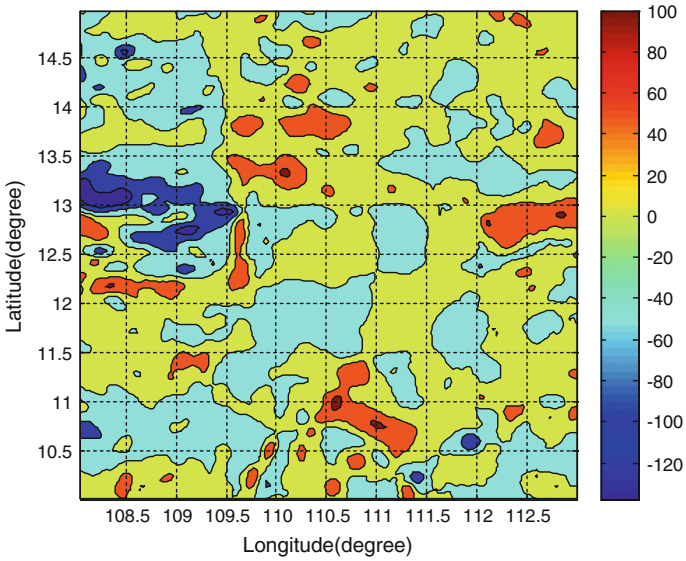


Fig. 1 Geomagnetic anomaly grid map

4.2 Simulation Experiment A

In simulation experiment A, it is assumed that gyro three axis constant drift is 0.01°/h, acceleration bias drift is 10^{-5} g, initial latitude and longitude error are all 0.135 n mile. INS initial point is (109.5°E, 10.5°N), yaw angle is 50°, AUV velocity is 9.7 n mile/h, and experiment period length is 4.12 h. Simulation track in the experimental area and the variation of geomagnetic anomaly between the inertial navigation and the reference track are illustrated in Figs. 2 and 3, respectively.

From Fig. 3, it can be seen that geomagnetic anomaly of experimental area range −30 nT from 40 nT. In the filtering, the filter step is 741 s, the solution step of inertial navigation is 1 s. The selecting principle of filter step is guaranteed that the navigation distance exceed one geomagnetic map grid in the filtering period. According to the initial state information estimation, initial state variance is 0.018 n mile^2. On the basis of the precision of magnetometer and geomagnetic anomaly map, the initial observation noise variance is 10 nT^2, and initial process observation noise variance is $10°^2$. Based on geomagnetic anomaly UKF filtering algorithm of probability weighted, filter matching experiment is carried out. Especially, UKF filtering algorithm of probability weighted denoted by improved UKF method, and filtering algorithm of same weight coefficient as the UT transformation denoted by traditional UKF method. It can be seen from Fig. 4 that INS position errors accumulate with time. And the latitude and longitude matching errors with improved UKF method is shown in Fig. 5.

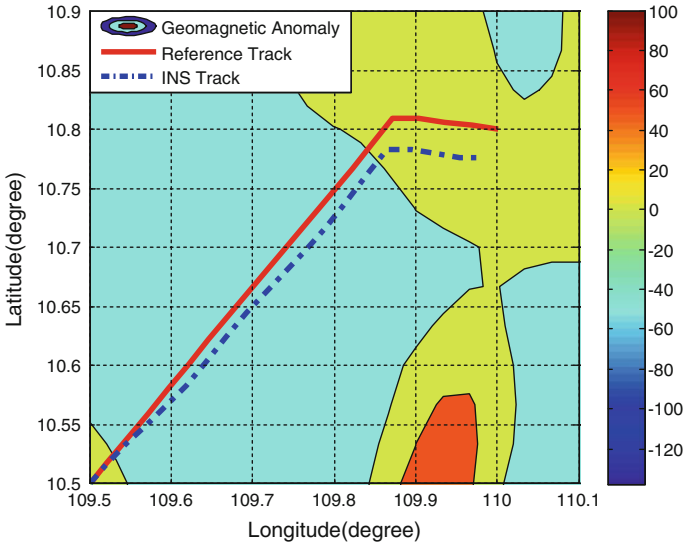


Fig. 2 The track in the simulation experiment A

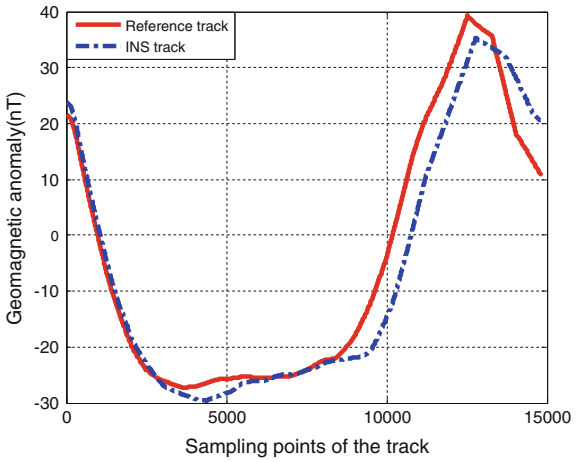


Fig. 3 Variation of geomagnetic anomaly of the inertial navigation and the reference track in the simulation experiment A

Compared with traditional UKF method, the INS position errors can be corrected more effectively by improved UKF method. For further quantitative analysis, the statistics of absolute matching errors are listed in Table 1.

It can also be seen from Table 1 that the performance of improved UKF method is much better than traditional UKF method. Moreover, serious mismatch occurred in matching results using traditional UKF method. The position accuracy is raised to 0.35 and 0.21 n mile in latitude and longitude directions respectively.

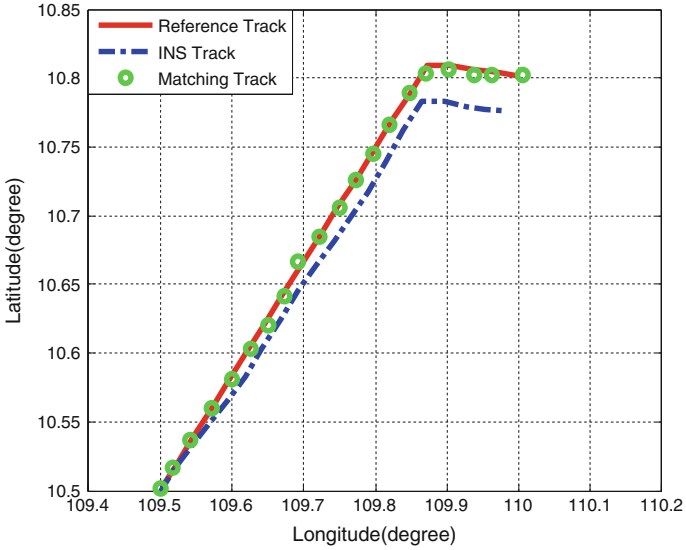
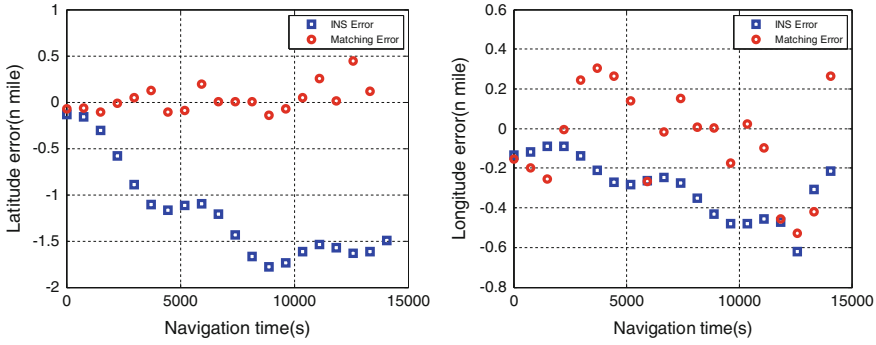


Fig. 4 The matching track in simulation experiment A



(a) The latitude matching errors **(b) The longitude matching errors**

Fig. 5 The matching errors along latitude and longitude in simulation experiment A

The position accuracy of integrated navigation system controls at 0.68 n mile. However, the position accuracy of INS reaches 1.78 n mile.

4.3 Simulation Experiment B

In order to further validate this improved algorithm in different geomagnetic feature region, the INS initial point's latitude and longitude are changed to (111.5°E, 10.5°N).

Table 1 The statistics of absolute matching errors in simulation experiment A (unit: n mile)

		Maximum	Minimum	Mean	MSE
Latitude	INS error	1.78212	0.13500	1.19305	1.33320
	Traditional matching error	98.63499	0.23887	41.09391	47.88795
	Improved matching error	0.68547	0.00202	0.25275	0.34678
Longitude	INS error	1.21560	0.08970	0.37658	0.47872
	Traditional matching error	102.28693	0.08740	69.80426	72.33497
	Improved matching error	0.52920	0.00372	0.27136	0.20869

But other simulation parameters remain unchanged and referred as simulation experiment B (Figs. 6 and 7).

From Figs. 8 and 9, we can also find that improved UKF method is more accurate and robust. And it is also shown in Table 2 that mean position accuracy is raised to 0.35 and 0.37 n mile in latitude and longitude directions through improved UKF method. It is confirmed that improved UKF method is more stable in the combined navigation system.

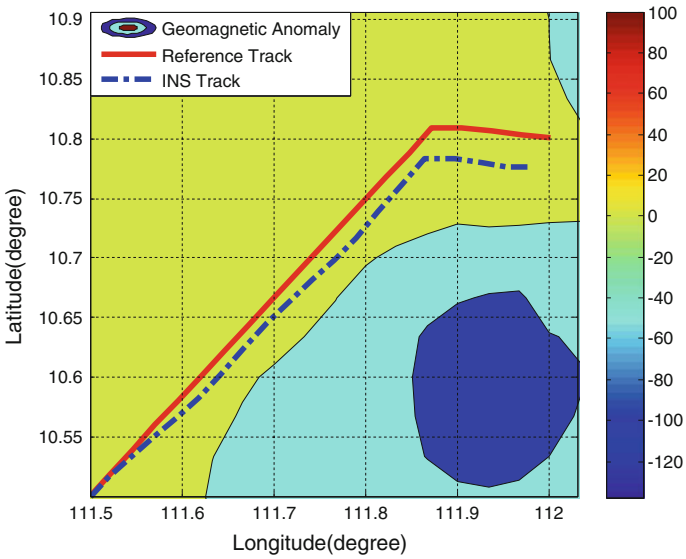


Fig. 6 The track in the simulation experiment B

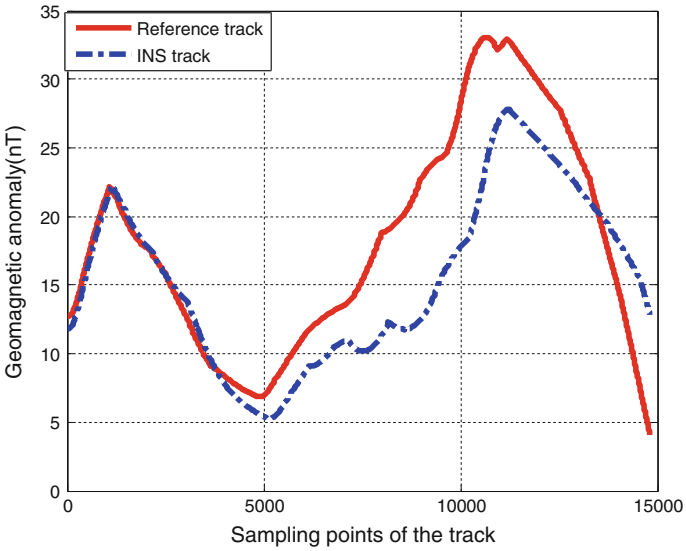


Fig. 7 Variation of geomagnetic anomaly of the inertial navigation and the reference track in the simulation experiment B

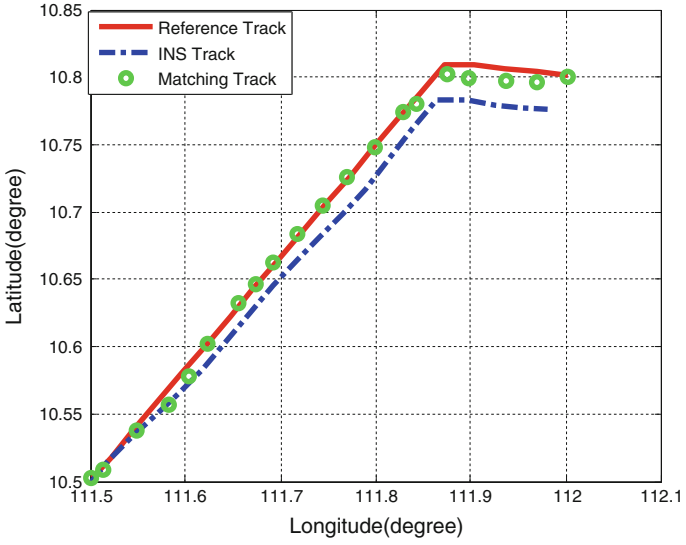
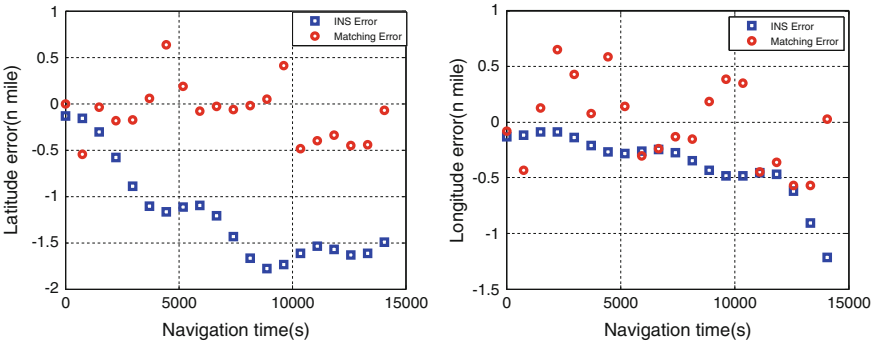


Fig. 8 The matching track in simulation experiment B



(a) The latitude matching errors **(b) The longitude matching errors**

Fig. 9 The matching errors along latitude and longitude in simulation experiment B

Table 2 The statistics of absolute matching errors in simulation experiment B (unit: n mile)

		Maximum	Minimum	Mean	MSE
Latitude	INS error	1.78212	0.13500	1.19305	1.33320
	Traditional matching error	114.57436	0.02890	69.89076	88.23783
	Improved matching error	0.63019	0.00132	0.25275	0.34608
Longitude	INS error	1.21560	0.08970	0.37658	0.47872
	Traditional matching error	231.43650	0.07652	150.68350	193.46721
	Improved matching error	0.65099	0.02395	0.31198	0.37420

5 Conclusions

When dealing with forecast observation, UKF filtering algorithm of probability weighted could give forecast observation varying degrees of "trust" in terms of probability. By replacing weight coefficient calculated through traditional algorithm, it can ensure the reliability of observation forecast in a certain extent. From two simulation results, it also can be seen that the improved UKF method can boost the accuracy of matching navigation effectively. When using weight coefficient same with status update, serious mismatch occurred in matching results using traditional UKF method. The plausible reason for this could be due to the fact that forecast observation is consist of quasi observation through traditional weight coefficient. The filter results based on quasi observation are fed back to integrated navigation system constantly, which leads to matching position not to be trusted at the filter moment.

UKF filtering algorithm has a strong dependence on the initial state value, and it is more susceptible to the uncertain system noise and perturbation errors from state model. As a consequence, further research is needed to assess the improved UKF filtering algorithm, providing for the reference of geomagnetic aided inertial navigation.

Acknowledgements Thanks to the NGDC for providing EMAG2. This work was supported by National Natural Science Foundation of China (41574010, 41476087) and State Key Laboratory of Geo-information Engineering.

References

1. Zheng H, Wang H, Wu L, Chai H, Wang Y (2013) Simulation research on gravity-geomagnetism combined aided underwater navigation. J Navig 66(1):83–98
2. Zhang X (2016). Theory and practice of geomagnetic navigation[M]. National Defense Industry Press
3. Huang C, Tian H, Zhao H (2013) Present research situation of geomagnetic filter navigation technologies. Sci Technol Eng 13(30):8976–8982
4. Psiaki M, Huang L, Fox S (1993) Ground tests of magnetometer-based autonomous navigation (MAGNAV) for low-earth-orbiting spacecraft. J Guid Control Dyn 1(1):395–407
5. Deutschmann J, Baritzhack I (1996) Attitude and trajectory estimation using earth magnetic field data. Aiaa J
6. Armstrong B, Wolbrecht E, Edwards DB (2010) AUV navigation in the presence of a magnetic disturbance with an extended Kalman filter. Oceans. IEEE, 1–6
7. Wang X, Tian Y (2010) Autonomous navigation based geomagnetic research. Chin J Geophys 53(11):2724–2731
8. Gu D, Qin Y (2006) Unscented Kalman filter for in-motion alignment of shipborne SINS. Syst Eng Electron 28(8):1218–1220
9. Hu Z, Guo C, Zhang S et al (2009) Application of unscented kalman filter in geomagnetic navigation for aerodynamic missile. J Astronaut 30(4):1443–1448
10. Liu Y, Wu M (2010) Intelligent geomagnetic matching method. In: International conference on intelligent computing and integrated systems. IEEE, pp 494–497
11. Zhao L, Gao N, Huang B et al (2015) A novel terrain-aided navigation algorithm combined with the TERCOM algorithm and particle filter. IEEE Sens J 15(2):1124–1131
12. Salmon B, Kleynhans W, Bergh F et al (2014) Meta-optimization of the extended Kalman filter's parameters through the use of the bias variance equilibrium point criterion. IEEE Trans Geosci Remote Sens 52(8):5072–5087
13. Xie W, Li Q, Xi B et al (2015) Geomagnetic contour matching algorithm based on iterative method. J Chin Inertial Technol 23(5):631–635
14. Li SS (2010) Research on the theory and method of underwater gravity-aided inertial navigation. Information Engineering University
15. Bar-Shalom Y, Tse E (1975) Tracking in a cluttered environment with probabilistic data association. Automatica 11(5):451–460
16. Dezert J (1999) Improvement of strapdown inertial navigation using PDAF. IEEE Trans Aerosp Electron Syst 35(3):835–856
17. Li SS, Wu XP, Ma B (2011) Correlative extremum matching algorithm using underwater gravity anomalies. Acta Geodaetica Cartogr Sin 40(4):464–469
18. Huang MT (2005) Determination of marine gravity field and its application. Surveying and Mapping Press

INS-Aided Single-Frequency Cycle-Slip Detection for Real-Time Kinematic GNSS

Lingxuan Wang, Yu Gan, Erhu Wei, Lifen Sui and Xuexi Liu

Abstract GNSS ambiguity fixed solution can greatly improve the accuracy of GNSS/INS integrated system. But it is difficult to achieve perfect real-time dynamic single frequency cycle-slip detection merely by GNSS observations especially in complex environment. Inertial assisted cycle-slip detection terms (DTs) based on station-satellite double-differences and satellite single-difference observations are derived. The error characteristic of the DT is analyzed emphatically. The DT error is affected by the drift of INS error. In addition, the magnitude of the influence on different satellites is related to the angle between its station-satellite vector and that that vector of the reference satellite. Thus, it is important to select reference satellite. It is proposed that two group of DTs can be used together by selecting two different reference satellites. The threshold of detection is estimated in a sliding window, where the DTs, whose INS error is submerged in GNSS error, are removed in order to reflect INS error. The method of threshold estimating has stronger self-adaptability.

Keywords GNSS/INS · Cycle-slip · Single frequency · Reference satellite Sliding window · Azimuth

1 Introduction

GNSS cycle-slip detection and ambiguity resolution are both key issues in high-precision GNSS/INS integrated position and orientation system for the using of GNSS carrier phase observation. There are many research results about the GNSS cycle-slip detection. Among them, the representative methods such as Polynomial fitting, high-order difference [1, 2], Doppler integration [3] and

L. Wang (✉) · E. Wei · X. Liu
School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
e-mail: geo_wlx@163.com

Y. Gan · L. Sui
Information Engineering University, Zhengzhou 450001, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_46

triple-difference [4], are not suitable for dynamic situations. And the combination of the MW method [5], the ionospheric residual method [6] and the combined triple-frequency method [7] cannot directly detect the single frequency cycle-slip. In addition, the pseudo-range phase combination method [8] couldn't detect small cycle-slip. Although the adoption of dual-frequency combination, but this need a joint solution to determine each frequency cycle-slip. Furthermore, the noise level of $(1, -1)$ combination is still high [9]. Some researchers estimate the current epoch ambiguity by carrier phase observations with the auxiliary of the position solution of INS [10, 11]. But when GNSS is out of lock for a long time, the accuracy of ambiguity estimation will decrease due to the influence of INS error drift. Which would lead to a decreased performance of the DTs. Colombo [12] proposed the DTs by making epoch difference with the position increment calculated from approximate formula. Obviously, this method would bring extra errors. Then scholars put forward many inertial assisted terms in the context of Precise Point Position (PPP) [13, 14]. A series of solutions have been proposed in succession [15–20]. But still do not have in-depth analysis of the error characteristics in INS-aided GNSS cycle-slip DTs, and cannot reflect the detailed impact of INS error. Thus there is no specific threshold estimation algorithm can be used.

In this paper, we deduce the INS-aided GNSS single-frequency cycle-slip DTs based on station-satellite double-difference as well as satellite single-difference. The error characteristics of DTs are analysed emphatically. It is very important to choose reference satellite. This paper presents a method of choosing two reference satellites to constitute the DTs which in order to reduce the errors caused by azimuth. According to conclusions, we recommend single-difference DTs in integrated system (even in relative-positioning applications). The method of threshold estimation by sliding window is proposed, which can fully reflect the influence of INS error and adaptively adjust the detection threshold. Finally, the calculation and analysis of experiments verify the high performance of INS-aided real-time dynamic GNSS single-frequency cycle-slip detection algorithm.

2 INS-Aided GNSS Cycle-Slip Detection Terms

Satellite single-difference could be deduced from the subtracting between two GNSS observation equations:

$$\lambda \nabla \Phi_j^{p,q} = \nabla \rho_j^{p,q} - \lambda \nabla N_j^{p,q} - \nabla \delta t^{p,q} - \nabla d_I + \nabla d_T + e_{\nabla \Phi} \tag{1}$$

where, λ is the wavelength, ρ_j^q is the distance between receiver j and satellite q, Φ_j^q denotes carrier phase observation value, N_j^q is the integer ambiguity, δt_j is the receiver clock error, δt^q is the satellite clock error, d_I and d_T are atmospheric refraction delay, $e_{\Phi,j}^q$ is the receiver noise. Record reference satellite as p, and ∇ denotes satellites difference operator.

Station-satellite double-difference observation equation could be deduced from the difference between single-difference observation equations of two stations:

$$\lambda\nabla\Delta\Phi_{i,j}^{p,q} = (\nabla\rho_j^{p,q} - \nabla\rho_i^{p,q}) - \lambda\nabla\Delta N_{i,j}^{p,q} - \nabla\Delta d_I + \nabla\Delta d_T + e_{\nabla\Delta\Phi} \quad (2)$$

where, Δ denotes stations difference operator. Take the INS position parameter into $\nabla\rho_j^{p,q}$ in formula (2) and omit the subscript symbol, which can be written as:

$$\nabla\Delta\Phi - \nabla\Delta\rho_I/\lambda + \nabla\Delta N = -\nabla\Delta d_I/\lambda + \nabla\Delta d_T/\lambda + e_{\nabla\Delta\Phi} \quad (3)$$

where, the units of $e_{\nabla\Delta\Phi}$ is cycle, which denotes double-difference noise. Suppose that it has fixed integer ambiguity $\nabla\Delta N$ in previous epoch t_m. While the INS assisted cycle-slip DTs in current epoch t_k could be written as follows:

$$DT_{dd} = \nabla\Delta\Phi - \nabla\Delta\rho_I/\lambda + \nabla\Delta N_{t_m} \quad (4)$$

Using this DT needs to restore fixed GNSS integer ambiguity resolution $\nabla\Delta N_{t_m}$ at the epoch of t_m. But, for the vulnerability of GNSS, the integer ambiguity is difficult to be accurately fixed under complicated environment which may lead to a performance reduction of the DTs. Furthermore, three-difference observation equation can be deduced by epoch difference of Eq. (2):

$$\delta\nabla\Delta\Phi = (\delta\nabla\rho_j - \delta\nabla\rho_i)/\lambda - \delta\nabla\Delta d_I/\lambda + \delta\nabla\Delta d_T/\lambda + e_{\delta\nabla\Delta\Phi} \quad (5)$$

where, δ represents epoch difference operator. However, influenced by tremendously variation of receiver coordinates, this method is no longer suitable for detecting small cycle-slip in dynamic positioning and navigation.

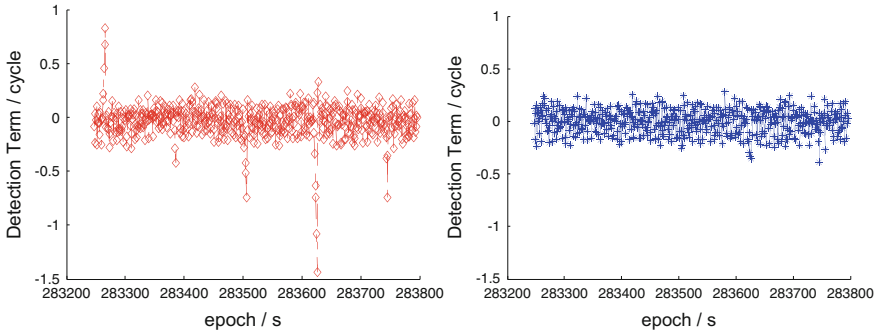
INS provides us extra position parameters. So we can substitute the INS position, $\boldsymbol{X}_{I,k-1}$ at the epoch of t_{k-1} and $\boldsymbol{X}_{I,k}$ at the epoch of t_k, into Eq. (5):

$$\delta\nabla\Delta\Phi = \delta\nabla\rho_I/\lambda - \delta\nabla\rho_i/\lambda - \delta\nabla\Delta d_I/\lambda + \delta\nabla\Delta d_T/\lambda + e_{\delta\nabla\Delta\Phi} \quad (6)$$

Thus the GNSS double-difference cycle-slip DTs with INS-aided epoch-difference could be written as:

$$DT_{dd} = \delta\nabla\Delta\Phi - \delta\nabla\rho_I/\lambda + \delta\nabla\rho_i/\lambda \quad (7)$$

The position parameter $\boldsymbol{X}_{I,k-1}$ is corrected by GNSS/INS combination at epoch of t_{k-1}. So IMU drift errors will be partially eliminated by the difference between epoch t_k and epoch t_{k-1}. We simulate 5 s partial loss of lock at each 120 s interval to evaluate the performance of each cycle-slip DT. The integer ambiguity cannot be fixed during the loss of lock. As shown in Fig. 1, the method of epoch difference still has good detection performance without the INS being timely corrected by high precision carrier phase observation.



(a) INS-aided Single-epoch DT (b) INS-aided epoch-differenced DT

Fig. 1 **a** Represents current epoch cycle DT and **b** represents epoch-differenced DT, both of them are belong to the same satellite

If satellite single-difference observations are adopted, the epoch-difference observation equation could be deduced from Eq. (1):

$$\delta\nabla\Phi = \delta\nabla\rho_j/\lambda - \delta\nabla\delta t/\lambda - \delta\nabla d_I/\lambda + \delta\nabla d_T/\lambda + e_{\delta\nabla\Phi} \tag{8}$$

INS-aided cycle-slip DT could be written as:

$$DT_{sd} = \delta\nabla\Phi - \delta\nabla\rho_I/\lambda - \delta\nabla\delta t/\lambda \tag{9}$$

Similarly, INS-aided cycle-slip DT for non-difference observations could also be deduced from epoch difference:

$$DT_{ud} = \delta\Phi - \delta\rho_I/\lambda - \delta\nabla\delta t/\lambda \tag{10}$$

However, unlike satellite single-difference and station-satellite double-difference observations, receiver clock error, which include the effects of low-cost quartz clock's speed error and drift error, could not be eliminated in non-difference observations. Furthermore, this effect typically cannot be eliminated by modeling. In other word, INS-aided non-difference DT won't be capable to effectively detect cycle-slip due to the influence of GNSS receiver clock error.

3 Error Analysis of Cycle-Slip Detection Terms

Except the GNSS observational environment and carrier phase measurement accuracy, the performance of the INS-aided GNSS cycle-slip DTs will be limited by the accuracy level of IMU. Therefore, it is necessary to analyze the specific influence of satellite observation value on INS error. The single-antenna operation

mode, such as PPP, can only use the DTs of satellite single-difference observation. On the other hand, relative positioning mood needs to pick over the two kinds of DTs between satellite single-difference and station-satellite double difference.

The error propagation process of the geometrical relationship between the receiver and the satellite is analyzed. Record $\vec{\boldsymbol{r}}_{k-1}$ and $\vec{\mathbf{r}}_k$ as the space vectors between receivers to GNSS satellite respectively at epoch t_{k-1} and epoch t_k; $\vec{\mathbf{S}}_k$ and $d\vec{\boldsymbol{X}}_k$ represent the position increment of GNSS satellite and receiver respectively. Therefore, the $\delta\rho$, distance difference between the front and back epochs, can be written as:

$$\begin{aligned}\delta\rho &= \rho_k - \rho_{k-1} = \boldsymbol{l}_k \cdot \boldsymbol{r}_k - \boldsymbol{l}_{k-1} \cdot \boldsymbol{r}_{k-1} \\ &= -\boldsymbol{l}_k \cdot d\boldsymbol{X}_k + (\boldsymbol{l}_k \cdot \boldsymbol{r}_{k-1} - \boldsymbol{l}_{k-1} \cdot \boldsymbol{r}_{k-1} + \boldsymbol{l}_k \cdot \boldsymbol{S}_k)\end{aligned} \tag{11}$$

In above formula, $\mathbf{l}_k$ represents unit vector (direction cosine vector). Set:

$$o(\boldsymbol{X}_{I,k-1}) = \boldsymbol{l}_k \cdot \boldsymbol{r}_{k-1} - \boldsymbol{l}_{k-1} \cdot \boldsymbol{r}_{k-1} + \boldsymbol{l}_k \cdot \boldsymbol{S}_k \tag{12}$$

Regardless of the influence of ephemeris error, the error of this term is mainly caused by the error of INS position $\mathbf{X}_{I,k-1}$:

$$\Lambda[o(\mathbf{X}_{I,k-1})] = (\mathbf{l}_k - \mathbf{l}_{k-1}) \cdot \Lambda[\mathbf{X}_{I,k-1}] \tag{13}$$

Here, $\Lambda[]$ indicates true error. Taking GPS satellites as an example, the average height above Earth's surface of satellites in GPS constellation is 2.0×10^7m, and the average speed of motion is 3800 m/s. According to the sampling rate of integrated system receiver, usually above 1 Hz, the angle of station-satellite vector of adjacent epochs can be estimated as:

$$\alpha \approx 3800/\left(2.0 \times 10^7\right) = 1.9 \times 10^4 \tag{14}$$

$$|\mathbf{l}_k - \mathbf{l}_{k-1}| = \sqrt{2 - 2\mathbf{l}_k \cdot \mathbf{l}_{k-1}} = \sqrt{2 - 2\cos\alpha} \approx 1.9 \times 10^{-4} \tag{15}$$

According to Eq. (15), assuming that $\Lambda[\mathbf{X}_{I,k-1}]$ reaches a hundred meters level, the influence on $\delta\rho$ is only on the order of centimeters. Namely, the absolute position drift error of IMU only has a weak influence on the distance difference. Therefore, it can be said that the error of the distance difference is mainly determined by the integral increment $d\mathbf{X}_k$ between the epochs obtained by the IMU measurement. Hence, it can be inferred that when $d\mathbf{X}_k$ reaches a certain accuracy requirement, the inertial-aided cycle-slip DT could still achieve good detection performance even if the ambiguity is not fixed for a period of time.

Then the $\delta\nabla\rho$ term, which need to be solved for both satellite single-difference and station-satellite double-difference, is analyzed. According to Eq. (11):

$$\delta\nabla\rho = \nabla(-\mathbf{l}_k \cdot d\mathbf{X}_k + o(\mathbf{X}_{I,k-1})) = -\nabla\mathbf{l}_k \cdot d\mathbf{X}_k + o(\mathbf{X}_{I,k-1}) \tag{16}$$

Therefore, the error of $\delta\nabla\rho$ can be recorded as:

$$\Lambda[\delta\nabla\rho\,] = -\nabla\mathbf{l}_k \cdot \Lambda[d\mathbf{X}_k] = -(\mathbf{l}_k^q - \mathbf{l}_k^p) \cdot \Lambda[d\mathbf{X}_k] \tag{17}$$

The model of $\nabla\mathbf{l}_k$ can be deduced from Eq. (15):

$$|\nabla\mathbf{l}_k| = \left|\mathbf{l}_k^q - \mathbf{l}_k^p\right| = \sqrt{2 - 2\cos\beta} \tag{18}$$

In above equation, the angle between the ground-satellite vectors of receiver to satellite p and receiver to reference satellite q is represented by β. It can be seen that the error of cycle-slip DT of GNSS carrier phase observation is mainly determined by the Angle of β. Compared with altitude angle, azimuth angle has a bigger influence on β. So the error of DT will be tiny if the difference of azimuth angles of receiver to satellite p and receiver to reference satellite q is small. Therefore, the selection of the reference satellite during the cycle-slip detection is crucial and requires additional consideration of the azimuth factor.

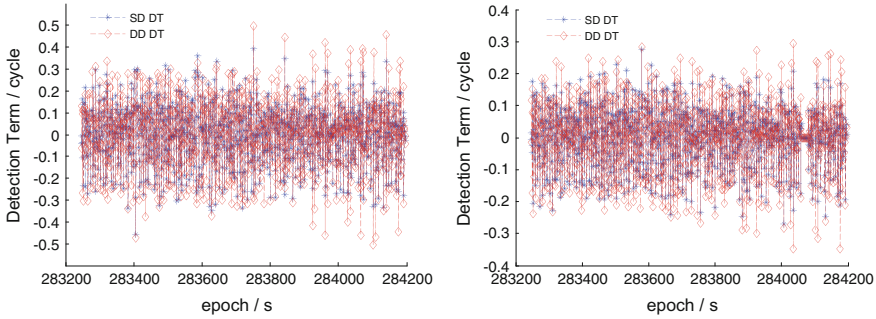
If only one reference satellite is selected, the error of the cycle-slip DTs of visible satellites which with a large azimuth angle difference from the reference satellite will be magnified. Considering use two reference satellites for cycle-slip detection, and each of the visible satellites selects a relatively small noise DT. Therefore, the selection strategy of reference satellite for cycle-slip detection can be summarized as follows: Firstly, a good quality reference satellite is determined by selected principles such as altitude angle, signal-to-noise ratio and so on; In order to weaken the influence of azimuth, another satellite, served as the second reference satellite, is determined by the azimuth which differs significantly from the first reference satellite (which difference should be almost 180°).

From the above analysis, the error of INS-aided GNSS cycle-slip DT mainly includes the carrier phase observation noise and the error of the increment of IMU position. That is, the error of INS-aided cycle-slip DT for station-satellite double-difference observation can be expressed as:

$$\sigma_{DT,dd} = \sqrt{\sigma^2(\delta\nabla\Delta\Phi) + \sigma^2(\delta\nabla\rho_I/\lambda)} = \sqrt{8\sigma_\Phi^2 + \sigma^2(\frac{\delta\nabla\Delta d_a}{\lambda}) + \sigma^2(\frac{\delta\nabla\rho_I}{\lambda})} \tag{19}$$

In the above formula, $\sigma(\delta\nabla\Delta d_a/\lambda)$ represents residual error of atmospheric refraction delay. In the same way, the error of INS-aided GNSS cycle-slip DT for satellite single-difference observation can be expressed as:

$$\sigma_{DT,sd} = \sqrt{\sigma^2(\delta\nabla\ \Phi) + \sigma^2(\delta\nabla\rho_I/\lambda)} = \sqrt{4\sigma_\Phi^2 + \sigma^2\left[(\frac{\delta\nabla d_a}{\lambda}) + (\frac{\delta\nabla\delta t}{\lambda}) + (\frac{\delta\nabla\rho_I}{\lambda})\right]} \tag{20}$$



(a) SD and DD Detection Term (PRN15) (b) SD and DD Detection Term (PRN29)

Fig. 2 SD and DD Detection Term respectively belong to PRN15 and PRN29. Here only PRN10 satellite is selected as one reference satellite according to the principle of altitude angle, signal-to-noise ratio

Similarly, $\sigma(\delta\nabla d_a/\lambda)$ represents the residual error of atmospheric refraction delay. Compared with Eqs. (19) and (20), the error level of DT_{sd} would lower if following formula is valid:

$$\sigma^2(\frac{\delta\nabla d_a}{\lambda}) + \sigma^2(\frac{\delta\nabla\delta t}{\lambda}) < 4\sigma_\Phi^2 + \sigma^2(\frac{\delta\nabla\Delta d_a}{\lambda}) \tag{21}$$

As shown in Fig. 2, the results indicates that noise level of DT_{sd} is lower than that of DT_{dd}. Therefore, GNSS/INS integrated navigation system, due to the factors such as high sampling and the high correlation errors between adjacent epochs, is recommended to use the DT of satellite single-difference, even in relative positioning conditions such as GNSS RTK/INS integration.

In addition, it is worth noting that in the experiments of Fig. 2, the average altitude angle of PRN15 during the observation period is about 49° and the average altitude angle of PRN29 is about 41°, but the noise level of PRN15 is obviously higher than that of PRN29. This is because the difference of azimuth angle of receiver to PRN29 and PRN10 is small, which leads to the small angle of the station-satellite angle β, and the DT error is not amplified by the error of INS position increment.

4 The Threshold Determination of Cycle-Slip Detection Terms

Relatively speaking, error magnitude of GNSS carrier phase observation is much smaller than that of IMU position increment. If the threshold of cycle-slip DT is determined from 4 times the mean square error of DT, the detection of small cycle-slip such as 1cycle will need $4\sigma_{DT,sd} < 1$. Thus, the error of IMU increment

should at least satisfy $\delta\nabla\rho_I/\lambda < 0.25$. The requirement of the detection for L1 frequency is higher than that of L2 frequency. So taking L1 frequency as an example, which requires:

$$\delta\nabla\rho_I < 0.0476 \tag{22}$$

Using two sets of reference satellites to form cycle-slip DTs proposed in this paper, the $|\nabla\mathbf{l}_k|$ will be less than 1 in most cases. Therefore, if the error of INS position increment is less than 4.76 cm between the two epochs, a small cycle slip can be detected successfully.

However, the error factor of IMU is complicated, which is influenced by the movement of the vehicle and the accuracy of the inertial components. And residual uncorrected errors of IMU navigation parameters will affect IMU position increment error. So it is necessary to consider the additional usage of a sliding window (t_{k-m}, t_{k-1}) with the width of m and to estimate the threshold of cycle-slip DT.

The standard deviation of the GNSS noise contained in the INS-aided single-difference detection term DT_{sd} can be recorded as $\sigma_{e_{\delta\nabla\Phi}} = 2\sigma_\Phi$, and the value mainly falls within the interval $-[4\sigma_\Phi, 4\sigma_\Phi]$ (about 95% confidence level). INS increment error is greater than GNSS noise, which is the main reference of threshold estimation. However, when the dynamic degree is tiny in some epochs or the error of the previous epoch is weakened to be more "clean" by the combination filtering, that INS errors in these epochs are submerged by GNSS noise throughout the sliding window, which make it impossible to extract valid INS information. In this case, in order to fully reflect the influence of INS errors, all the terms that satisfy $|DT_{sd}| \le 4\sigma_\Phi$ should be eliminated, and the rest should be reserved to estimate the threshold. If the value of DT in the window all falls within the scope of $-[4\sigma_\Phi, 4\sigma_\Phi]$, the DT is considered to be mainly affected by the GNSS error, and the threshold value should be estimated by all the data in the window. In conclusion, the threshold of DT can be determination by this equation:

$$T_{DT_{sd}} = \begin{cases} 4\sigma_{DT_{sd}}(DT_{sd}) & \max(|DT_{sd}|) \le 4\sigma_\Phi \\ 4\sigma_{DT_{sd}}(|DT_{sd}|4\sigma_\Phi) & else \end{cases} \tag{23}$$

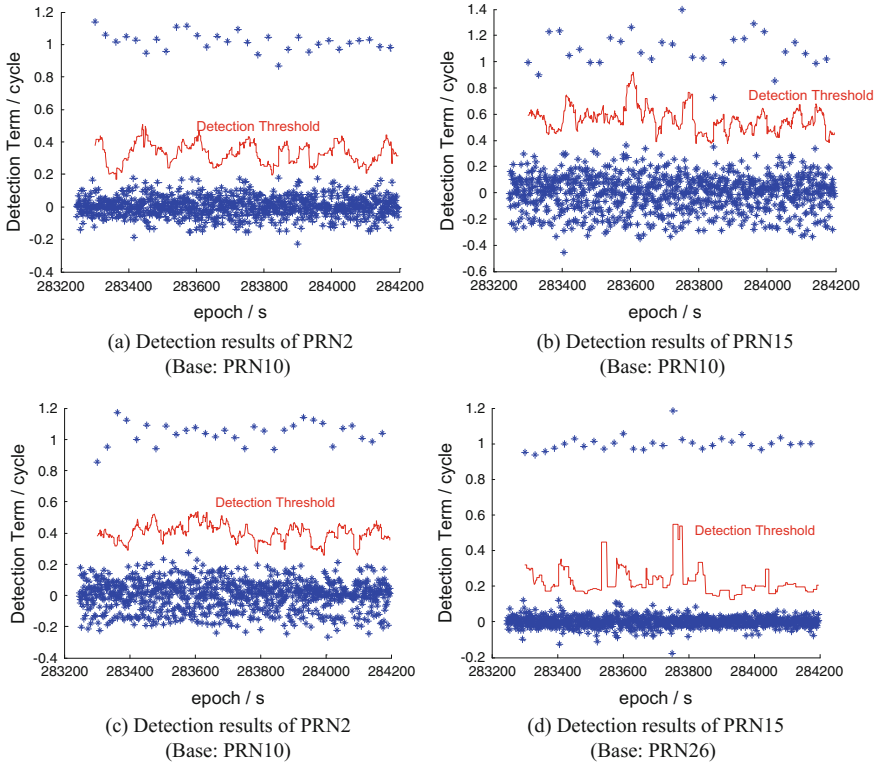
5 Calculation and Analysis

Taking GPS/INS combination system as an example. The vehicle-bone experiment with GPS/INS integrated navigation is carried out in Wuhan. The average number of visible satellites is about 6 $\sim$ 8; the receiver sampling rate is 1 Hz; and the INS sampling rate is 100 Hz. The INS gyro zero bias is 1°/h and the accelerometer bias is 0.1 mg (nominal value), which can meet the requirement of position increment error for cycle-slip DT. In order to verify the method of this paper, the original GPS data has been preprocessed to detect and fix the cycle-slip. The carrier phase

observations of the three satellites PRN2, PRN15 and PRN29, which are uniform distributed at the zenith, are respectively added small cycle-slip of 1 cycle every 30 s. The azimuth angles of the station-satellite vectors belong to the PRN2, PRN15 and PRN29 are about 99°, 199° and 298°, respectively. According to the reference star selection strategy described above, reference stars are selected as PRN10 and PRN26. And PRN10 and PRN26 have azimuth angles of about 10° and 189°, respectively. The detection was performed using the two sets of satellite single-difference DTs, and the detection threshold was estimated using the sliding window with a width of 30 epochs. Experiment results are shown in Fig. 3.

Analyzing the above results of cycle-slip detection, it can be concluded that:

1. For the three selected PRN2, PRN15 and PRN29 satellites, the performance of INS-aided GNSS cycle-slip DT presented in this paper is favorable, and all the small cycle-slips are detected. The smaller IMU error will be submerged by the



(a) Detection results of PRN2 (Base: PRN10)

(b) Detection results of PRN15 (Base: PRN10)

(c) Detection results of PRN2 (Base: PRN10)

(d) Detection results of PRN15 (Base: PRN26)

Fig. 3 **a**, **b** and **c** Shows GPS cycle-slip detection by INS-aided DT of PRN2, PRN15 and PRN29 with PRN10 as reference satellite; **d** Shows GPS cycle-slip detection by INS-aided DT of PRN15 with PRN26 as reference satellite. The blue asterisk "*" in the figure indicates the value of the cycle-slip DT, and the solid red line indicates the threshold obtained from sliding window estimation. Detection results of PRN26 and PRN29 with PRN10 as reference satellite are not as good as those of the PRN10 as the reference satellite, and they are not drawn here

GNSS error, and the part of the DT should be rejected. The sliding window is used to estimate the threshold of cycle-slip DT to suppress insignificant INS information. The threshold estimation is adaptively adjusted with the error of IMU position increment.
2. If only select a reference star, the error cycle-slip DT error will be enlarged if the difference between the azimuths of detected satellite and reference satellite is huge. This is likely to cause missed exploration and risk exploration.
3. For PRN2 and PRN29 satellites, PRN10 reference satellite can be selected to form cycle-slip DT. And PRN26 should be used as reference satellite for PRN15 to obtain the optimal cycle-slip detection performance. The method proposed in this paper can choose DT with smaller error for each satellite and enhance the reliability of cycle-slip detection.

6 Conclusions

It is difficult to achieve a perfect real-time dynamic single-frequency cycle-slip detection using GNSS alone. INS provides additional position and position incremental information which can assist in calculating the part of related geometric distance from the satellite to station. So this enable satellite single-difference or station-satellite double-difference observations can be used in cycle-slip detection.

This paper studies the INS-aided GNSS single-frequency cycle-slip DT with focusing on its error characteristics. The error of DT directly affects the performance of the cycle-slip detection, and the reference star selection is crucial to the INS-aided DT. The proposed method of utilizing two reference satellites with large difference in azimuth angle can greatly reduce the influence of azimuth factor and enhance the performance of cycle-slip DT. Combined with the sliding window usage, the threshold estimation for cycle-slip DT can be adaptively adjusted with the error of IMU position increment

Acknowledgements This study is supported by Nation Science Foundation of China (41374012; 41674016; 41274016) and Innovate Foundation of Information Engineering University (XS201504).

References

1. Li Z, Huang J (2005) GPS surveying and data processing. Wuhan University Press, Wuhan
2. Liu J (2008) The principle and method of navigation and positioning using GPS satellites, 2nd edn. Science Press, Beijing
3. Li Z, Zhang X (2009) New techniques and precise data processing methods of satellite navigation and positioning. Wuhan University Press, Wuhan
4. Yuan H, Wan W, Ning B, Li J (1998) A new cycle slip detection and correction method using triple differences solution. Acta Geod Et Cartogr Sin 27(3):189–194

5. Blewitt G (1990) An automatic editing algorithm for GPS data. Geophys Res Lett 17(3): 199–202
6. Fang R, Shi C, Wei N (2009) Real-time cycle-slip detection for quality control of GPS measurements. Geomat Inf Sci Wuhan Univ 34(9):1094–1097
7. Li J, Yang Y, Xu J, He H, Guo H (2011) Real-time cycle-slip detection and repair based on code-phase combinations for GNSS triple-frequency un-differenced observations. Acta Geod Et Cartogr Sin 40(6):716–717
8. Chen P (2010) Cycle slips detecting and repairing by use of phase reduce pseudorange law and ionized layer remnant method of difference. J Geod Geodyn 30(2):120–124 (Qingdao)
9. Tang Z (2011) Investigations on cycle slip detection and correction in GPS precise point positioning. Shandong University of Science and Technology, Qingdao
10. Altmayer C (2002) Enhancing the integrity of integrated GPS/INS systems by cycle slip detection and correction. In: Intelligent vehicles symposium. IV 2000. Proceedings of the IEEE. IEEE, pp 174–179
11. Lee HK, Wang J, Park WY, Rizos C (2003) Carrier phase processing issues for high accuracy integrated GPS/Pseudolite/INS systems. In: Proceedings of 11th IAIN World Congress, pp 252–273
12. Colombo OL, Bhapkar UV (1999) Inertial-aided cycle-slip detection/correction for precise, long-baseline kinematic GPS. In: ION GPS 1999, Nashville, USA
13. Du S, Gao Y (2012) Inertial aided cycle slip detection and identification for integrated PPP GPS and INS. Sensors 12(11):14344–14362
14. Du S (2011) An inertial aided cycle slip detection and identification method for integrated PPP GPS/ MEMS IMU system. In: Proceedings of International Technical Meeting of the Satellite Division of the Institute of Navigation, vol 28, no 2, pp 3183–3191
15. Liu S, Sun F, Zhang L et al (2015) Instantaneous re-convergence of precise point positioning by using INS-aided cycle-slip correction. J Chin Inert Technol 23(5):607–614
16. Gan Y (2015) GNSS/INS integrated system model refining and position/attitude determination using carrier phase. Information Engineering University, Zhengzhou, pp 101–115
17. Liu Q, Sui L, Xiao G et al (2015) A method of determining the weight matrix for BDS DCB resolution. J Geomat Sci Technol 32(5):473–478
18. Zhang X, Guo F, Li P et al (2012) Real-time quality control procedure for GNSS precise point positioning. Geomat Inf Sci Wuhan Univ 37(8):940–944
19. Ma G, Yu B, Jia R et al (2016) INS-aided high dynamic GNSS rapid acquisition and stable tracking. Radio Eng 46(2):23–26
20. Ye P (2011) MEMS IMU/GNSS ultra-tight integration navigation technology. Shanghai Jiao Tong University, Shanghai, pp 15–39

Research on Geomagnetic Matching Localization for Pedestrian

Peiwen Gong, Dongyan Wei, Xinchun Ji, Wen Li and Hong Yuan

Abstract Nowadays, geomagnetic localization has become a new locating method for pedestrian. In this paper, geomagnetic matching localization method is proposed, the result of localization is based on the correlation between the measured geomagnetic data and the known base. In the flat terrain, pedestrian dead reckoning (PDR) is used to obtain mileage accumulation information to assist the geomagnetic matching locating, while in the rough terrain like elevators, stairways, etc., where PDR cannot provide reliable mileage information, the dynamic time warping (DTW) algorithm is used to warp the measured geomagnetic data and the known base. The experimental results show that the proposed method can provide the results of locating reliably and continuously and meet the demands of pedestrian localization.

Keywords Geomagnetic · Pedestrian · DTW · PDR · Smart phone

1 Introduction

Nowadays, the location-based services (LBS) provide us with many navigating and locating methods. With the popularization of intelligent terminals and the development of various navigation technologies, the demand for location-based services is also on the rise. Compass navigation systems, GPS navigation systems and other satellite navigation systems can provide users with navigation and positioning services in general outdoor scenes. However, in urban canyons, tunnels and indoor scenes, satellite navigation is disturbed by signal shading, therefore, seeking new

P. Gong
University of Chinese Academy of Sciences, No 19(A) Yuquan Road, Beijing, Shijingshan District 100049, China

P. Gong · D. Wei (✉) · X. Ji · W. Li · H. Yuan
Academy of Opto-Electronics, Chinese Academy of Sciences, No 9 Deng Zhuang South Road, Beijing, Haidian District 100094, China
e-mail: weidongyan@aoe.ac.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_47

navigation method which can provide precise location information in all scenarios has become an important direction for the development of navigation.

The development of positioning technologies such as Bluetooth positioning, WLAN positioning, pedestrian dead reckoning (PDR) positioning, RFID positioning, radio network positioning and geomagnetic localization has been progressing [1–7]. At present, there are many methods and techniques available for geomagnetic localization [3–6]. Such as [3] which bases on geomagnetic beacons and magnetic dipoles, the use of low frequency alternating magnetic signals and electromagnetic induction principle is not subject to interference, but relies on the laying of magnetic beacons. In [4] and [5], geomagnetic filtering algorithm is used and geomagnetic filter technology has good real-time performance, but there are some limitations. In [4], an adaptive filtering algorithm based on orthogonalization of new information is proposed, which improves the robustness, but the noise characteristics need to be updated constantly, the model and the calculation are complicated. The literature [5] adopts the adaptive UKF filtering method, which needs to consider the environmental geomagnetic noise and has a complicated modeling. Literature [6] based on geomagnetic relativity matching localization, using geomagnetic matching and inertial navigation system method, but this method is not suitable for pedestrian geomagnetic localization.

For pedestrians, the walking environment contains flat terrain, but also includes rough terrain such as elevators, stairways, ramps and so on. In the flat terrain, magnetic localization method with PDR-assisted is proposed, while in the rugged terrain the magnetic localization method with DTW-assisted is proposed, so we can obtain stable and reliable magnetic localization results under all pedestrian scenarios.

PDR is a method by making dead reckoning for localization based on MEMS. Currently, inertial sensors required for PDR are generally integrated with most smartphones. Based on the accumulated mileage information obtained by PDR, geomagnetic matching positioning method based on correlation matching is proposed, Extended Kalman filter (EKF) was used to fuse the geomagnetic matching positioning results and PDR positioning results. Positioning equipment without the use of other equipment than phones, meanwhile enhance the geomagnetic matching positioning reliability and continuity.

Dynamic Time Warping (DTW) is an algorithm originally applied in the field of speech signal processing. It mainly solves the problem that the test speech cannot match directly with the template speech duration due to the randomness of speech signal pronunciation. The algorithm can be used to solve the problem that the matching cannot be performed without the assistance of the mileage information in the geomagnetic localization process. Therefore, under the condition of rough terrain, DTW algorithm is used in this paper to regularize the geomagnetic data collected so as to carry out the geomagnetic match positioning.

2 Magnetic Localization with PDR Assisted

2.1 System Structure

The process of magnetic matching localization is divided into two phases: offline database construction phase and online matching phase. Offline phase includes the process of magnetic data space alignment and reference coordinate system establishment. The online phase includes magnetic data spatial alignment and matching. PDR is introduced to provide space-based reference for magnetic matching. Finally, EKF is used to filter the localization results of PDR and geomagnetic. The structure of the positioning system is shown in Figs. 1 and 2.

Offline phase was divided into magnetic data space alignment and the establishment of the reference coordinate system: mobile magnetic sensors collect the magnetic field strength for the data space alignment, meanwhile the dead reckoning was carried out, mileage was estimated by the pace and step estimation. The magnetic data would be transformed to change with the spatial sequence to complete the geomagnetic data space alignment, and the reference points were divided evenly by mileage. The reference coordinate system can adopt the combination of GNSS and high-precision INS, and can also be established by the method of eigenvalue calibration and inertial recursion [2].

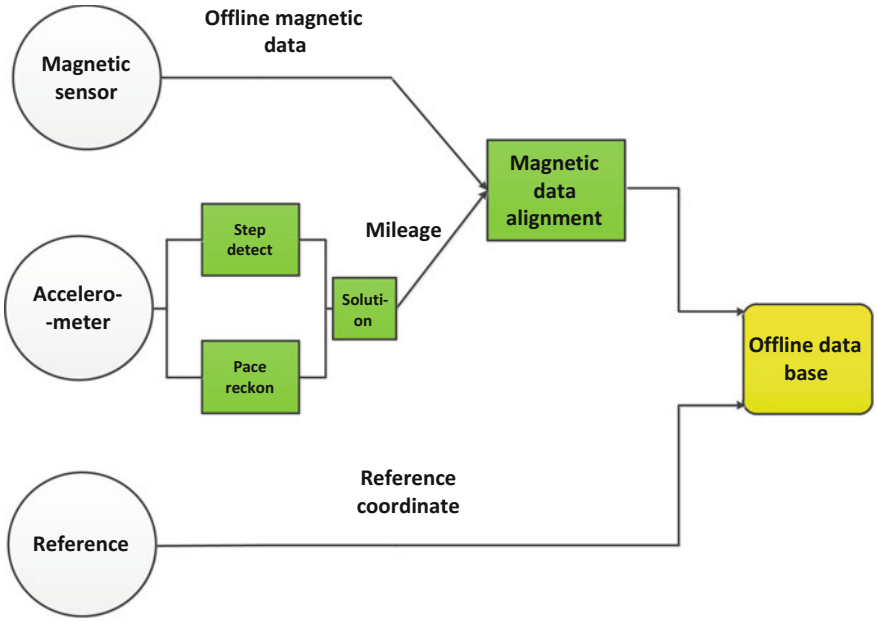


Fig. 1 The structure of offline magnetic map building

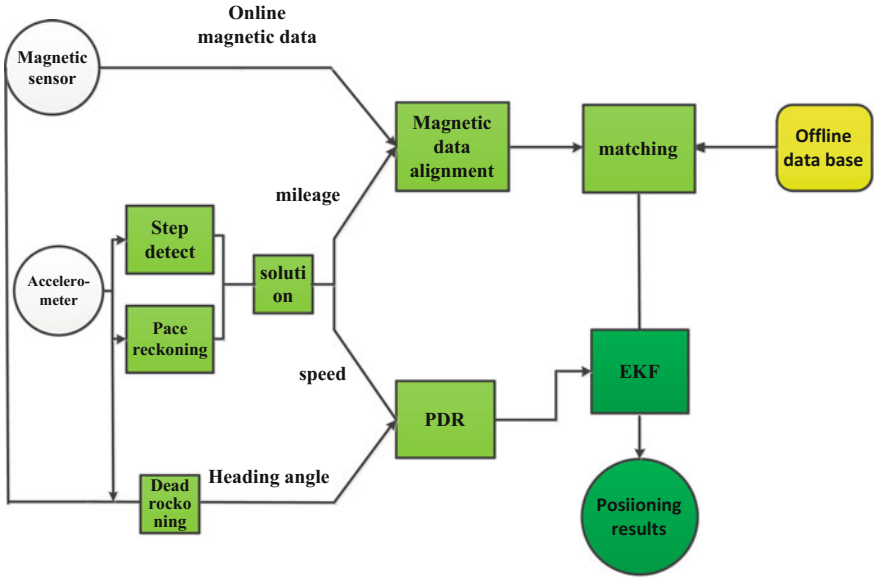


Fig. 2 The structure of online magnetic matching and assistant localization

The online phase is divided into two processes: spatial alignment of magnetic data and matching localization. Magnetic data will be transformed to change with the spatial sequence to complete the spatial alignment of geomagnetic data. In the matching localization phase, online geomagnetic data was divided into sliding window and matched with the offline geomagnetic database, at last we got geomagnetic matching position estimation results through the correlation solution.

PDR provided dead reckoning while providing mileage for magnetic matching positioning. Finally, the EKF filtering algorithm was used to filter the two positioning results to reduce the positioning error and provide stable and accurate positioning result continuously.

2.2 Magnetic Matching Localization

This chapter describes how to make spatial alignment of geomagnetic data during the process of matching and locating, and describes the principle of magnetic matching locating and offline database construction and online matching phases.

2.2.1 Spatial Alignment of Magnetic Data

In the process of collect magnetic data in the online and offline phases, the magnetic sensor carried by the mobile phone collected magnetic data with a fixed frequency. Because of the users move with different speeds during the online phase and the offline phase, as shown in Fig. 3, the magnetic data cannot be aligned by time of the two phases, so that the correlation between the magnetic data in the offline magnetic database and the online phase cannot be matched.

In order to solve the problem, it is necessary to change the geomagnetic data with time series to with the spatial sequence, align the data which were collected in the online phase and the offline database spatially. Figure 4 is the transformation of geomagnetic data with time series into with spatial sequence.

Before the transformation, the corresponding relationship of magnetic field strength and mileage and time is as follows:

$$\begin{cases} (M,T) = \{(M_1,t_1),(M_2,t_2),\ldots(M_n,t_n)\} \\ (D,T) = \{(0,t_1),(d,t_2),\ldots((n-1)d,t_n)\} \end{cases}$$

Take time as reference,combine with mileage and geomagnetic:

$$(M,D) = \{(M_1,0),(M_2,d),\ldots(M_n,(n-1)d)\}$$

2.2.2 Offline Phase

Offline stage, a smartphone integrated magnetic sensor is used to collect the magnetic field data along the way, the mileage is calculated by PDR, and the mileage is divided equally as d, take the same interval point as a reference point RP. We use eigenvalue calibration with inertial recursive [2] to complete the establishment of the reference coordinate of the entire test route. At the same time, the space alignment of the geomagnetic data was completed. The information in the offline magnetic database is shown in Table 1.

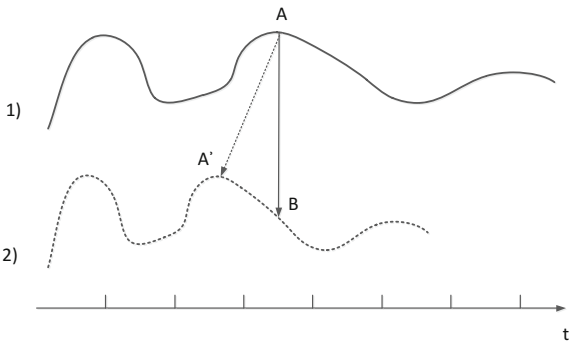


Fig. 3 Data collect: magnetic field changes with time

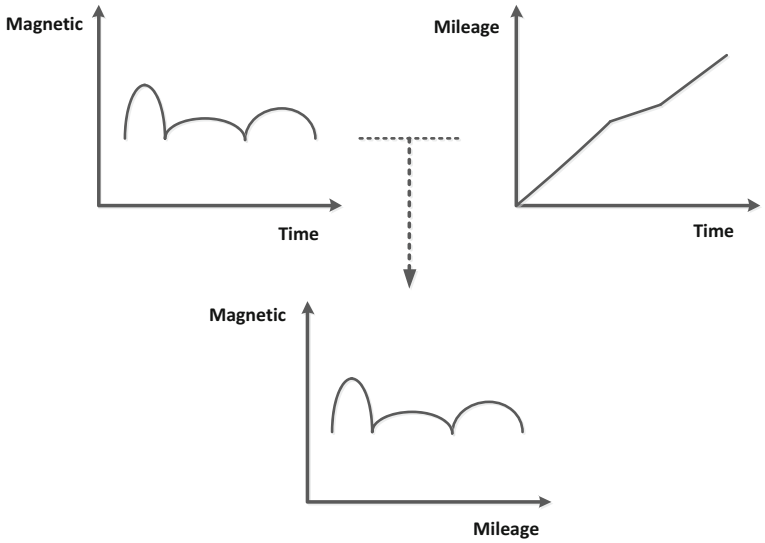


Fig. 4 Data alignment: time array to space array

Table 1 Offline magnetic map

RP	Coordinate	Mileage	Magnetic
1	$X_1 = (x_1, y_1, z_1)$	0	$M_1 = (m_{x1}, m_{y1}, z_{z1})$
2	$X_2 = (x_2, y_2, z_2)$	d	$M_2 = (m_{x2}, m_{y2}, z_{z2})$
3	$X_3 = (x_3, y_3, z_3)$	2d	$M_3 = (m_{x3}, m_{y3}, z_{z3})$
…	…	…	…
N	$Xn = (x_n, y_n, z_n)$	(n−1)d	$M_n = (m_{xn}, m_{yn}, z_{zn})$

2.2.3 Online Phase

In online matching phase, we made time-to-space sequence conversion, take w as the mileage scale, the physical meaning is to take the $D = wd$ length of the magnetic field strength vector as a sliding window, and magnetic data vector in offline geomagnetic database are matched by sliding matching. The matching starting point is where the user's online data reaches the length D, the matching process is as shown in Fig. 5.

In order to avoid the problem that the correlation of magnetic field strength cannot be directly matched by different values of magnetic field strength collected by different mobile phone sensors, the matching algorithm adopts the correlation coefficient (CC) algorithm. Correlation coefficient is used to reflect the correlation between the variables. The correlation coefficient is calculated by the method of product difference, which based on the difference between the two variables and the respective average value, the correlation between the two variables is reflected by multiplying the two differences:

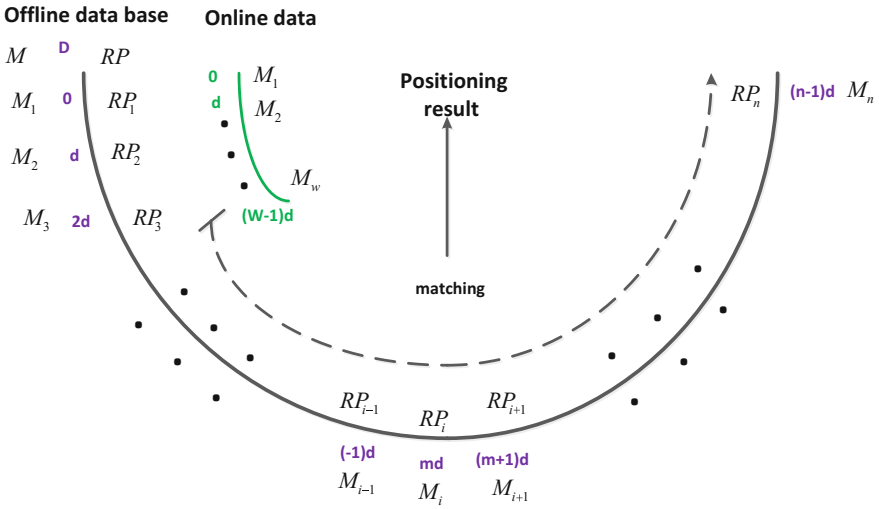


Fig. 5 Online magnetic data matching

$$R = \frac{\sum_{i=1}^{w}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{w}(x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{w}(y_i - \bar{y})^2}} \tag{1}$$

This formula reflects the correlation R between the two variables x, y, where y for the magnetic data collected in real time in the online phase, the geomagnetic data is divided in the same way, and w is the sliding window length.

As shown in Fig. 5, in the online matching process, Using the correlation coefficient method to solve the correlation coefficient R of the sliding window geomagnetic information $\boldsymbol{M_i}$= (m_{xi}, m_{yi}, m_{zi}),i = (1, 2… w − 1) and the offline data base, when R is the maximum correlation coefficient, it means that the matched magnetic data has the strongest correlation and the right location information can be obtained by reading the reference coordinates of the magnetic database which been created during offline phase.

3 Magnetic Localization with DTW Assisted

3.1 System Structure

As with the PDR-assisted geomagnetic localization method, the DTW-assisted magnetic localization method also includes two processes of offline database construction and online matching positioning. The following figure shows the construction of the magnetic localization system without mileage information (Fig. 6).

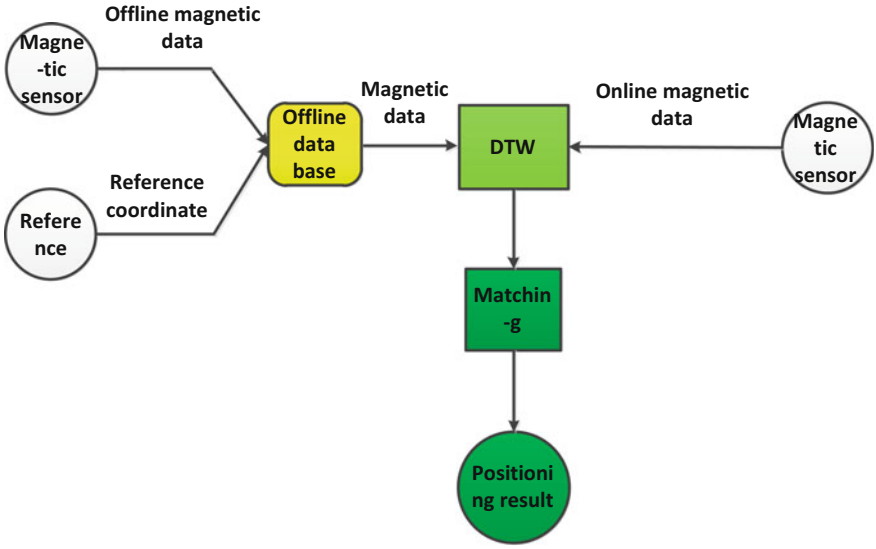


Fig. 6 The structure of magnetic localization without mileage

Offline stage is as same as the method in Sect 2.

The process of online matching positioning is divided into two phases: dynamic time warping and matching locating. DTW algorithm is used to conduct dynamic time warping of the geomagnetic data in the offline geomagnetic database and the online phase, and then the matching locating is completed to obtain the geomagnetic matching position estimation result.

In the case of no mileage information, using DTW algorithm can also complete the magnetic matching and get localization results in the pedestrian scene.

3.2 Principle of the DTW Algorithm

The DTW design was originally used in the area of speech signals. Due to the speech signal has strong randomness, the test tone and template tone cannot matched, for example, by using Euclidean distance with two time series.

Not only in the field of speech recognition, magnetic matching localization faces with similar problem. The magnetic data cannot be matched directly by the offline geomagnetic database and online geomagnetic data without processing. Different from the introduction of the mileage information in Sect. 2 as the spatial reference to align the magnetic data, this chapter uses DTW Algorithm for offline and online two data regularization.

Suppose the two magnetic signature sequences Q and P contain exactly the same geomagnetic data. However, due to the different speeds when the geomagnetism data collected in the offline phase and the online phase, the time series varies.

The distance d means the Euclidean distance of elements in sequences Q and P:

$$d(q_i, p_j) = (q_i - p_j)^2 \tag{2}$$

A path is defined as a regular path D. The physical meaning of a regular path D is that: $D_k = (i,j)_k$ is a mapping between two sequences Q and P, which includes:

$$D = D_1, D_1, \ldots, D_w, \ldots D_k \tag{3}$$

Meanwhile, the regular path D must meet the following restrictions:

(1) Boundary restrictions.
(2) Monotonous restrictions.
(3) Continuous restrictions.

With monotonic constraints and continuous limits, there are several paths to satisfy the three constraint [7].

The purpose of DTW is to find a regular path that minimizes the distance between two sequences of regular distances. In order to obtain this path, the cumulative distance $L = (i,j)$ is defined as follows: Starting from (1, 1), when the path reach to the end point (n, m), The cumulative distance is the sum of the total distance of all matching points, that is, the similarity between the sequence Q and P:

$$L(i,J) = d(q_i, p_j) + \min\{L(i-1, j-1), L(i-1, j), L(i, j-1)\} \tag{4}$$

The formula $L = (i,j)$ is the sum of the Euclidean distance of a point and the cumulative distance of the smallest neighboring element that arrives at that point. Find the sum of the minimum cumulative distance that is obtained the minimum regular cost of the regular path W, to complete the dynamic time warping. That is how to find the similarity of the two sequences.

4 Experiment and Analysis

In order to validate the measured data presented in this paper, experiments were conducted in the Academy of Opto-electronics of Chinese Academy of Sciences. The smart terminals were Huawei mate8 and the magnetic sensor is used to collect magnetic data. Meanwhile, the accelerometer is used to collect acceleration, heading angle and other data, so as to provide mileage for matching locating. The magnetic matching locating algorithm uses the correlation coefficient matching

algorithm. The PDR locating method adopted the acceleration rising detection method to make statistics, and the fusion filtering algorithm was EKF algorithm.

When matching the online magnetic data with the offline database, we choose 0.5 m as the mileage length *d*, and divided the acquired magnetic data spatially evenly. Taking into account the locating accuracy and cost requirements, the current test window length w is 35.

Figure 7 is an error results comparison chart of single magnetic matching locating source, single PDR locating source and EKF filtered fusion locating source.

As shown in the Fig. 7, the single source magnetic localization results show that the mismatch of the magnetic matching locating results affects the positioning accuracy, and the maximum locating error is 85 m. The localization error of single PDR positioning source shows that the PDR locating accuracy is positively correlated with the time divergence. PDR locating error was caused by the increase in the number of turns, and the divergence of the heading angle.

According to the existing problems of magnetic matching locating and PDR locating, EKF filtering fusion was performed on the results of two localization sources. As shown in Fig. 8, after EKF filtering fused, the mismatch problem of magnetic matching and the problem of divergence of PDR positioning is obviously improved. The localization results was continuous and stable, and the locating accuracy has been improved.

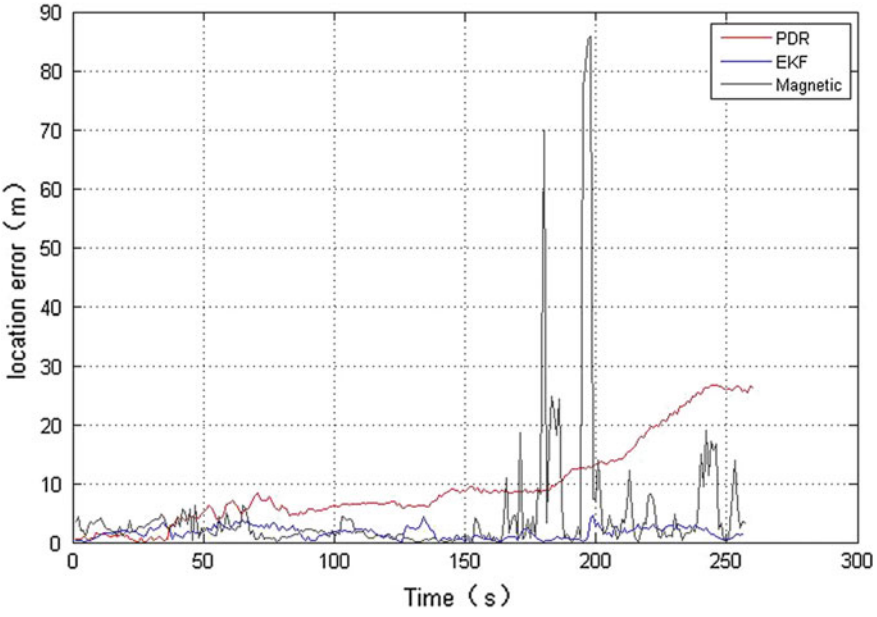


Fig. 7 Position error of three kind of localization

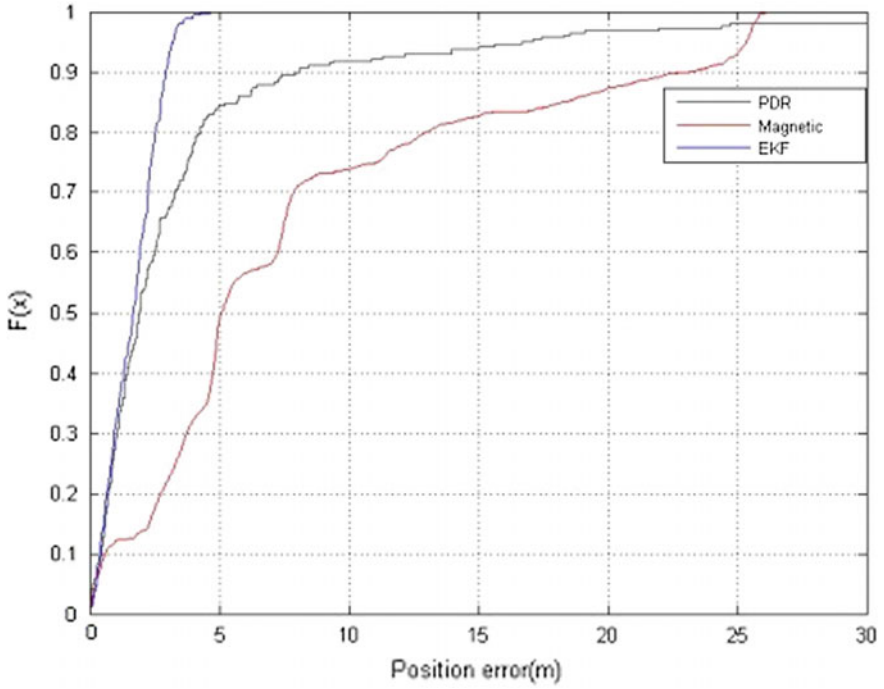


Fig. 8 Position error CDF of three kind of localization

Figure 8 is an error probability distribution statistic of magnetic matching localization, PDR localization and magnetic matching localization with PDR-aided.

As shown in Fig. 8, the single-source magnetic matching localization error distribution is 10 m (90%) and the maximum localization error is 85 m. The PDR localization error distribution is 15 m (80%) and the maximum localization error is 25.2 m (although PDR localization accuracy is high in short time, there is problem of divergence over time). After EKF filtering, the localization error is 4 m (99%) and the maximum localization error is 4.64 m. The results show that the proposed method improved the mismatch of magnetic matching method and can obtain high-precision localization results stable and continuously.

Figure 9 is the localization error of magnetic matching localization with DTW-aided.

Figure 10 is an error probability distribution statistic of magnetic matching localization with DTW-aided.

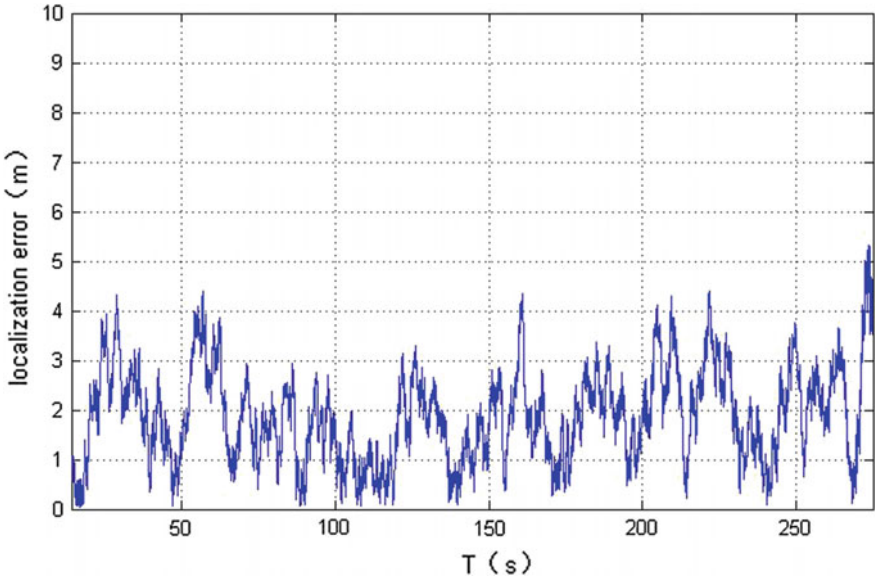


Fig. 9 Position error of magnetic localization with DTW-assisted

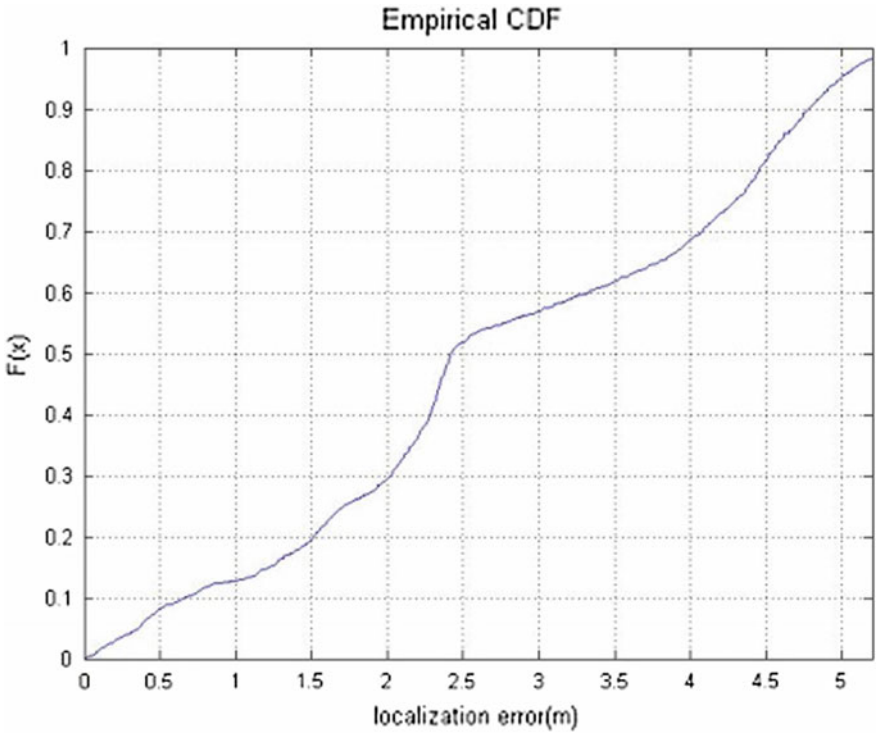


Fig. 10 CDF of position error of magnetic localization with DTW-assisted

5 Conclusion

In this paper, we proposed a magnetic matching locating method for pedestrian. The main work can be summarized as follows: (1) The PDR and DTW algorithm solved the problem that the magnetic data cannot be directly matched due to the difference of the walking speed of the user; (2) The method is low cost and designed for the majority of smartphone users, and also does not require additional equipment; (3) EKF fusion is used to remove the mismatch caused by magnetic localization error to obtain continuous and stable localization results. Experimental results show that this method can provide magnetic localization results stable and continuously.

Acknowledgements Authors greatly thank reviewers for the valuable suggestions, and thank colleagues in the same team for the thoughtful suggestions. This work was supported by National key Research Program of China "Collaborative Precision Positioning Project" (No.2016YFB0501900).

References

1. Yao TJ, Wei DY, Yuan H et al (2016) Research on the feedback correction-based fusing method for WLAN and PDR positioning. Chin J Sci Instrum 37(2):446–453
2. Li W, Wei DY, Yuan H et al (2016) A novel method of wi-fi fingerprint positioning using spatial multi-points matching. In: 2016 International conference on indoor positioning and indoor navigation, pp 1–8
3. Yeun SB, Young CK (2016) Localization based on magnetic markers for an all-wheel steering vehicle. Sensors 16(12):2015
4. Zhu ZL, Yang Y, Nian HT (2015) INS/GNS integrated method based on innovation orthogonality adaptive Kalman filter. J Chin Inert Technol (1):66–70
5. Wu M, Ma J, Tian JW et al (2011) An adaptive UKF algorithm and its application to geomagnetic navigation. Inf Control 40(4):558–562
6. Zhou YY, Zhang YS, Guo L (2011) Hardware in-the-loop simulation of GNS/INS integrated navigation system. Chin J Sci Instrum (s1):15–17
7. Pang XL, Lin CH (2010) Indoor localization through dynamic time warping. In: 2010 International conference on intelligent computation technology and automation, vol 2, pp 621–625

A Low Complexity Gaussian Parametric Message Passing Based Cooperative Localization Algorithm

Yangyang Liu and Baowang Lian

Abstract Based on the theory of factor graph and belief propagation, a low complexity cooperative localization algorithm with Gaussian parametric message passing is proposed to improve the performance where the non-cooperative localization methods failed due to the insufficient coverage of anchors. The system model is established according to the Bayesian rule. Weighted samples are used to represent the salient characteristics of the local message, and a Gaussian parametric message passing rule is designed to reduce the burden of the network traffic. By constructing a relative spatial relationship between the target and its neighbour nodes, a novel message initialization method is put forward to concentrate the samples where the messages have significant mass. In order to facilitate efficient computation of peer-to-peer messages, the nonlinear observation equation is linearized approximately by exploiting the Taylor expansion. Then the expression of the message updating is deduced and the detailed flow of the algorithm is shown. Simulation results show that the proposed algorithm leads to an excellent performance at the communication overhead and computational complexity, with losing negligible localization accuracy.

Keywords Cooperative localization · Belief propagation (BP) · Wireless sensor networks · Gaussian parametric

1 Introduction

Location based services are widely used in various applications, such as military, industrial automation, precision agriculture, search and rescue operations and the Internet of Things [1, 2]. Although global satellite navigation systems have largely solved the positioning problem in open outdoor environments, they do not function

Y. Liu (✉) · B. Lian
School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China
e-mail: liu_yang90@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_48

well in indoor, tunnel and urban canyon environments due to blockage of the GNSS signal. Cooperative method improves the positioning accuracy and coverage by establishing peer-to-peer communications and inter-measurements among the agents [3]. However, with the increase of the number of agents to be located, the computational complexity and the communication overhead have increased exponentially, which restricts the practical application of the cooperative location algorithm.

In Bayesian algorithms, the agent node is treated as a random variable and its probabilistic distribution is exchanged and updated. Recently, the message passing algorithm, also known as belief propagation based on probabilistic graphical models, has been an attractive method for the Bayesian cooperative localization due to its efficient and accuracy. Non-parametric belief propagation (NBP) combining with particle filter method is introduced into the cooperative localization in wireless sensor network (WSN), which is suitable for various models due to its flexible particle representation [4]. However, it is condemned by its large amount of computational complexity and communication overhead, which makes it difficult to apply in application. Combining BP with factor graph, a distributed cooperative algorithm for large-scale WSNs is proposed in [5], named sum-product algorithm over a wireless network (SPAWN), which plays an important role in the development of cooperative localization technology. Parametric message representative methods have attracted the attention of recent researches [6]. The message of the agent is approximated by a specific distribution cluster and only several parameters need to be exchanged among the agents. The parameters can be obtained by minimizing Kullback-Leilbler divergence (KLD), while the problems of high computational and multiple local optimal solutions are still existent. In order to solve the non-linearity of the range measurement, Taylor expansion is always adopted, but the error of the linearization is dependent on the point and the order of the expansion [7]. The censoring mechanism [8] has been proposed to reduce the system overhead and control the error propagation in cooperative localization.

In this paper, a low complexity Gaussian parametric belief propagation (GPBP) based cooperative localization algorithm is proposed to reduce the system overhead. Instead of randomly generating a large number of samples, we first estimate the prior information of the agent nodes and divided them into three categories: ring distribution, bimodal distribution and unimodal distribution by constructing the relative spatial relationship between the target and its neighbor anchor nodes, which is effective to concentrate the samples where the messages have significant mass and reduce the used samples. A Gaussian parametric message passing rule is designed to reduce the burden of the network traffic. Simulation results show that the proposed algorithm leads to an excellent performance at the communication overhead and computational complexity, with losing negligible localization accuracy.

2 System Model

We consider a two-dimensional wireless network deployed with N anchors and M agents, where the anchors and agents are defined as set $\mathbb{N}$ and $\mathbb{M}$, respectively, then $\mathbb{S} = \mathbb{N} \cup \mathbb{M}$ is the set of all the nodes. The positions of the anchors are known exactly while the agents are scattered randomly over the network. Let $\mathbf{x}_i = [x_i \; y_i]^T$ be the position of node i, and the positions of all nodes can be expressed as $\mathbb{X} = \{\mathbf{x}_i | \forall i \in \mathbb{S}\}$. Every node can range and communicate with neighbors within the maximum sense distance R. The sets of neighboring anchors and agents belonging to node i can be denoted by $\mathbb{N}_i$ and $\mathbb{M}_i$, respectively. Furthermore, we have $\mathbb{S}_i = \mathbb{N}_i \cup \mathbb{M}_i$. The measurement distance from node j to node i is perturbed by measurement noise

$$z_{j\to i} = \| \mathbf{x}_j - \mathbf{x}_i \| + v_{j\to i} \tag{1}$$

where $\|\cdot\|$ is the Euclidean norm and $v_{j\to i} \sim \mathcal{N}\left(0, \; \sigma^2_{j\to i}\right)$ is the additive white Gaussian measurement noise under the line of sight.

We define the notation $\mathbb{Z}_i = \left\{z_{j\to i} | \forall j \in \mathbb{S}_i\right\}$ as the randlge measurements set for node i, and $\mathbb{Z} = \{\mathbb{Z}_i | \forall i \in \mathbb{M}\}$ for all agents. Assuming the statistical characteristic is independent between nodes, the problem of cooperative localization in wireless networks can be converted into solving the joint probability density function (PDF) of all agents which can be obtained by message computing, passing and updating among the nodes. According to Bayesian rule, the joint PDF can be expressed as

$$\begin{aligned} p(\mathbb{X}|\mathbb{Z}) &\propto p(\mathbb{Z}|\mathbb{X}) p(\mathbb{X}) \\ &\propto \prod_{i\in\mathbb{M}} \prod_{j\in\mathbb{S}_i} p\left(z_{j\to i} | \mathbf{x}_i, \mathbf{x}_j\right) \prod_{m\in\mathbb{M}} p(\mathbf{x}_m) \prod_{n\in\mathbb{N}} p(\mathbf{x}_n) \end{aligned} \tag{2}$$

where $p\left(z_{j\to i} | \mathbf{x}_i, \mathbf{x}_j\right)$ is the likelihood function of node i and j, and $p(\mathbf{x}_m)$,$p(\mathbf{x}_n)$ are the prior distribution of anchors and agents, respectively. Then a posteriori distribution of agent i can be shown as

$$p(\mathbf{x}_i|\mathbb{Z}) = \int\limits_{\mathbb{X}\backslash\mathbf{x}_i} p(\mathbb{X}|\mathbb{Z}) \mathrm{d}\mathbb{X}\backslash\mathbf{x}_i \tag{3}$$

where $\mathbb{X}\backslash\mathbf{x}_i$ indicates all the variables in $\mathbb{X}$ except $\mathbf{x}_i$.

3 Proposed Message Passing Algorithm

3.1 Message Representation Method

In distributed cooperative localization, message refers to the node's position estimation and its uncertainty. Due to the nonlinearity of the ranging measurement and the irregular distribution of the neighbor nodes, it is intractable to compute the 2-D integrations in (3) directly. In order to solve this problem, weighted samples are always employed to approximate the target node's message for efficient peer-to-peer message computation. Generally, the message without effective prior information is considered as uniform distribution, which is represented by a large number of samples scattered in the cooperative area randomly. However, the number of the generated samples which fall close to the real position of the target node is very limited, resulting in an unexpected positioning result. Hence, the focus of this subsection is how to approximate the local message accurately with fewer samples.

In this work, for reference node $j \in \mathbb{S}_i$, we assume that $b(\mathbf{x}_j) = \mathcal{N}\left(\boldsymbol{\mu}_j, \sum_j\right)$, where $\boldsymbol{\mu}_j$ is the mean and $\sum_j$ is the covariance matrix. The range measurement from node j to node i is $z_{j\to i}$, but we have no information regarding the direction. Hence, the distribution of node i can be described as a ring distribution with the center of $\mathbf{x}_j$ and a radius of r, which can be further expressed as

$$\mathbf{x}_i = \mathbf{x}_j + r \times [\cos\theta \quad \sin\theta]^T \tag{4}$$

where $\mathbf{x}_j \sim \mathcal{N}\left(\boldsymbol{\mu}_j, \sum_j\right)$, $r \sim \mathcal{N}\left(z_{j\to i}, \ \sigma^2_{j\to i}\right)$, and $\theta \sim U(0, 2\pi]$. It should be note that if node j is an anchor node, then $\sum_j = 0$.

As shown in Fig. 1, we assume $\hat{\mathbf{x}}_i$ is the estimation of the node i, and $\hat{\theta} = \text{angle}(\hat{\mathbf{x}}_i - \boldsymbol{\mu}_j)$ is the direction between node i and node j. In order to obtain the variance $\sigma^2_{j,\hat{\theta}}$ associated with the marginal distribution of $b(\mathbf{x}_j)$ along $\hat{\theta}$, the coordinate system is rotated anticlockwise $\hat{\theta}$ angle. Then the covariance matrix of $\mathbf{x}_j$ in the new coordinate system is given as

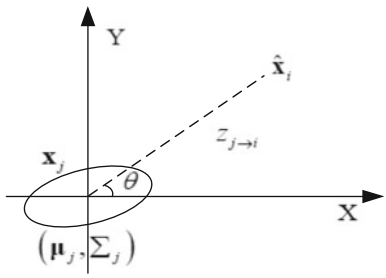


Fig. 1 Variance estimation of Gaussian marginal distribution

$$\sum_{j,\hat{\theta}} = U_{\hat{\theta}}^{T} \sum_{j} U_{\hat{\theta}} \tag{5}$$

where

$$U_{\hat{\theta}} = \begin{bmatrix} \cos\left(\hat{\theta}\right) & -\sin\left(\hat{\theta}\right) \\ \sin\left(\hat{\theta}\right) & \cos\left(\hat{\theta}\right) \end{bmatrix} \tag{6}$$

Hence, the $\sigma_{j,\hat{\theta}}^2$ associated with the (1, 1)-th element of $\sum_{j,\hat{\theta}}$ s shown as

$$\sigma_{j,\hat{\theta}}^2 = \left[\cos\left(\hat{\theta}\right), \sin\left(\hat{\theta}\right)\right] \sum_{j} \left[\cos\left(\hat{\theta}\right), \sin\left(\hat{\theta}\right)\right]^{T} \tag{7}$$

Because the marginal distribution of $\mathbf{x}_j$ along any direction is still Gaussian distribution, we can rewrite Eq. (4) as

$$\mathbf{x}_i = \boldsymbol{\mu}_j + r_\theta \times [\cos\theta \;\; \sin\theta]^{T} \tag{8}$$

where $r_\theta \sim \mathcal{N}\left(z_{j\to i},\ \sigma_{j\to i,\theta}^2\right)$ and $\sigma_{j\to i,\theta}^2 = \sigma_{j\to i}^2 + \sigma_{j,\theta}^2$.

Hence, the joint a posteriori distribution of $\mathbf{x}_i$ is related to the number of received neighbor nodes $|\mathbb{S}_i|$, which can further be considered as the overlapping areas of $|\mathbb{S}_i|$ rings. When $|\mathbb{S}_i| = 2$, the intersection area of the two rings is the target's position estimation. It is a bimodal distribution and we employ a mixture of two Gaussian distribution to approximate it. While for $|\mathbb{S}_i| \geq 3$, a steady unimodal distribution is obtained when the number of non-collinear reference nodes is no less than three and we approximate it with Gaussian distribution.

Without loss of generality, the equation set of observation between the target and its neighbor reference nodes can be formulated as follows:

$$\begin{cases} (x_1 - x_i)^2 + (y_1 - y_i)^2 = z_{1\to i}^2 \\ \vdots \\ (x_n - x_i)^2 + (y_n - y_i)^2 = z_{n\to i}^2 \end{cases} \tag{9}$$

where $n = |\mathbb{S}_i|$. We define notation $\hat{\mathbf{x}}_i^a$ and $\hat{\mathbf{x}}_i^b$ as the solutions of (6) when $|\mathbb{S}_i| = 2$, and the notation $\hat{\mathbf{x}}_i$ for $|\mathbb{S}_i| \geq 3$. Then the approximate distribution can be shown as follows

$$p(\mathbf{x}_i|\mathbb{Z}_i) = \begin{cases} C\left(\mathbf{x}_i; \mathbf{x}_j, z_{j\to i},\ \sigma_{j\to i}^2\right), & |\mathbb{S}_i| = 1 \\ \frac{1}{2} G\left(\mathbf{x}_i; \hat{\mathbf{x}}_i^a, \sigma_i^2 \mathbf{I}\right) + \frac{1}{2} G\left(\mathbf{x}_i; \hat{\mathbf{x}}_i^b, \sigma_i^2 \mathbf{I}\right), & |\mathbb{S}_i| = 2 \\ G\left(\mathbf{x}_i; \hat{\mathbf{x}}_i, \sigma_i^2 \mathbf{I}\right), & |\mathbb{S}_i| \geq 3 \end{cases} \tag{10}$$

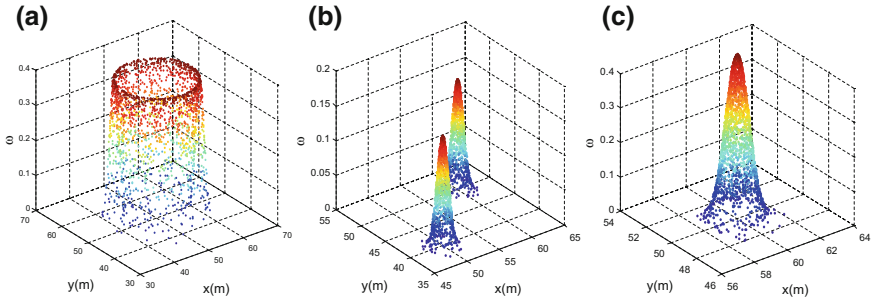


Fig. 2 The examples of generated samples from different distribution families for agent nodes. **a** $|\mathbb{S}_i| = 1$, the ring distribution; **b** $|\mathbb{S}_i| = 2$, the mixture of two Gaussian distribution; **c** $|\mathbb{S}_i| \geq 3$, Gaussian distribution

where the notations C and G refer to the ring distribution and Gaussian distribution, respectively. In the ring distribution, parameter $\mathbf{x}_j$ is the center of the ring and $z_{j\to i}$ is the radius with a variance $\sigma^2_{j\to i}$. And in Gaussian distribution, parameter $\hat{\mathbf{x}}_i$ is the mean, $\sigma_i^2\mathbf{I}$ is the variance matrix, and $\mathbf{I}$ is the identity matrix. For simplicity, let $\sigma_i^2 = \max\left\{\sigma^2_{j\to i,\theta} | \forall j \in \mathbb{S}_i\right\}$.

According to the above analysis, it can be seen that the distribution of target belongs to the different families in terms of the number of the received reference nodes $|\mathbb{S}_i|$. Therefore, for agent i, the message can be represented by randomly generating k samples $\{\mathbf{x}_i(n), w_i(n)\}_{n=1}^k$ according to (10). As an example, Fig. 2 shows the generated samples associated with different distribution families.

3.2 Message Computation

According to the sum-product algorithm update rule, the belief of each agent i is given as

$$b(\mathbf{x}_i) \propto p(\mathbf{x}_i) \prod_{j\in\mathbb{S}_i} m_{f_{i,j}\to i}(\mathbf{x}_i) \tag{11}$$

where $m_{f_{i,j}\to i}(\mathbf{x}_i)$ denotes the message from node j to node i. Considering the likelihood function associated with the range measurement between node i and j, the prior information of node j and the refined messages which are received by node j, $m_{f_{i,j}\to i}(\mathbf{x}_i)$ are further expressed as

$$\begin{aligned} m_{f_{i,j}\to i}(\mathbf{x}_i) &\propto \int p\left(z_{j\to i}|\mathbf{x}_i,\mathbf{x}_j\right)p\left(\mathbf{x}_j\right)\prod_{u\in\mathbb{S}_j} m_{f_{j,u}\to j}\left(\mathbf{x}_j\right)\mathrm{d}\mathbf{x}_j \\ &\propto \int p\left(z_{j\to i}|\mathbf{x}_i,\mathbf{x}_j\right)b\left(\mathbf{x}_j\right)\mathrm{d}\mathbf{x}_j \end{aligned} \tag{12}$$

When the node j is an anchor node, its belief is Dirac delta function. The above equation is simplified as

$$m_{f_{i,j}\to i}(\mathbf{x}_i) \propto p\left(z_{j\to i}|\mathbf{x}_i,\mathbf{x}_j\right) \tag{13}$$

Based on the message passing rule mentioned above, the belief of node j is approximated by a Gaussian distribution, e.g., $b\left(\mathbf{x}_j\right) = \mathcal{N}\left(\boldsymbol{\mu}_j, \sum_j\right)$, when the node j is an agent node. In this case, Eq. (12) is expressed as

$$m_{f_{i,j}\to i}(\mathbf{x}_i) \propto \int p\left(z_{j\to i}|\mathbf{x}_i,\mathbf{x}_j\right)\mathcal{N}\left(\boldsymbol{\mu}_j, \sum_j\right)\mathrm{d}\mathbf{x}_j \tag{14}$$

Equation (14) is intractable due to the nonlinearity of the ranging likelihood function. Hence, we approximate Eq. (14) by expanding the square root terms of the ranging model with using a first-order Taylor expansion around the position estimations of node i's and j's. We have

$$\begin{aligned} \| \mathbf{x}_j - \mathbf{x}_i \| &= \left\|\hat{\mathbf{x}}_{j\to i}\right\| - \frac{\hat{\mathbf{x}}_{j\to i}}{\left\|\hat{\mathbf{x}}_{j\to i}\right\|}\left(\mathbf{x}_j - \boldsymbol{\mu}_j\right) + \frac{\hat{\mathbf{x}}_{j\to i}}{\left\|\hat{\mathbf{x}}_{j\to i}\right\|}\left(\mathbf{x}_i - \hat{\mathbf{x}}_i\right) \\ &= \left\|\hat{\mathbf{x}}_{j\to i}\right\| - \frac{\hat{\mathbf{x}}_{j\to i}}{\left\|\hat{\mathbf{x}}_{j\to i}\right\|}\left[\left(\mathbf{x}_j - \mathbf{x}_i\right) + \hat{\mathbf{x}}_{j\to i}\right] \end{aligned} \tag{15}$$

where $\hat{\mathbf{x}}_{j\to i} = \hat{\mathbf{x}}_i - \boldsymbol{\mu}_j$. Substituting (15) into (14), we obtain

$$m_{f_{i,j}\to i}(\mathbf{x}_i) \propto \exp\left[-\frac{1}{2}\left(\widehat{\mathbf{x}}_{j\to i} - z_{j\to i}\frac{\widehat{\mathbf{x}}_{j\to i}}{\| \widehat{\mathbf{x}}_{j\to i} \|}\right)^{\mathrm{T}}\sum_{j+z}^{-1}\left(\widehat{\mathbf{x}}_{j\to i} - z_{j\to i}\frac{\widehat{\mathbf{x}}_{j\to i}}{\| \widehat{\mathbf{x}}_{j\to i} \|}\right)\right] \tag{16}$$

where $\sum_{j+z} = \sum_j + \sigma^2_{j\to i}\mathbf{I}$.

3.3 Message Updating

According to (13) and (16), the weights of the samples for agent i can be updated as follows

$$w_i^+(n) = w_i(n) \prod_{j \in \mathbb{S}_i} m_{f_{i,j} \to i}(\mathbf{x}_i(n)) \tag{17}$$

where $w_i^+(n)$ is the updated weight, and normalize it with $\sum_{n=1}^{k} w_i^+(n) = 1$.

In conventional BP algorithm, all the weighted samples need to be broadcasted to the neighbors, which leads to huge communication overhead. In order to reduce the network traffic, we approximate it by a Gaussian parametric distribution with $\left(\boldsymbol{\mu}_i, \sum_i\right)$. The parameters can be obtained as follows

$$\boldsymbol{\mu}_i = \sum_{n=1}^{k} w_i^+(n) \mathbf{x}_i(n) \tag{18}$$

$$\sum_i = \frac{\sum_{n=1}^{k} w_i^+(n)(\mathbf{x}_i(n) - \boldsymbol{\mu}_i)(\mathbf{x}_i(n) - \boldsymbol{\mu}_i)^T}{1 - \sum_{n=1}^{k} (w_i^+(n))^2} \tag{19}$$

Owing to the Gaussian approximation, we only need to broadcast five parameters (two is mean and another three is covariance) per node per iteration to complete the message delivery among the agents. According to $\mathcal{N}\left(\boldsymbol{\mu}_i, \sum_i\right)$, we regenerate k new samples randomly for agent i and assign a weight for each sample as:

$$w_i(n) = \exp\left[\frac{1}{2}(\mathbf{x}_i(n) - \boldsymbol{\mu}_i)^T \sum_i^{-1} (\mathbf{x}_i(n) - \boldsymbol{\mu}_i)\right] \tag{20}$$

Based on the analysis above, the proposed cooperative localization algorithm is described in Table 1.

Table 1 Flow of cooperative localization algorithm

1. Initialize the boundary of the cooperative network and the prior information of the nodes (e.g., Dirac delta function for anchors and random distribution for agents)
2. For each agent node i, reinitialize its distribution and randomly generate k samples according to (10), normalize the weights
3. For $l = 1$ to L\{iteration index\}, repeat 4 $\sim$ 7
4. Compute the messages using (13), (16) for anchors and agents, respectively
5. Update the weights using (17)
6. Estimate the mean and covariance using (18) and (19)
7. Regenerate k new samples randomly from $\mathcal{N}\left(\boldsymbol{\mu}_i, \sum_i\right)$, and assign a weight to each sample using (20)
8. End of the iteration, all user nodes calculate the posteriori probability distribution of their own position

4 Simulation Results

We consider a 100×100 m^2 cooperative localization area with N anchors and M agents. The positions of the anchors are known exactly and the agents are scattered randomly. The maximum sense distance between nodes is $R = 20$ m. In the scenario, the ranging measurements noise is assumed as an additive white Gaussian noise with zero mean and a standard deviation σ. The maximum iteration number is assigned as $L = 10$.

A single trial performance of the GPBP, SPAWN and H-SPAWN is illustrated in Figs. 3 and 4. The simulation configures are $N = 13$, $M = 100$, $\sigma = 1$ m and $k = 500$. In Fig. 3, a straight line is used to connect the position estimation and its real position together. It can be seen that in addition an agent on the upper right corner, the others can be located with high accuracy. Additionally, the overall performance of SPAWN is superior to the proposed GPBP, but the gap is very small. The detailed statistical results are shown in Fig. 4.

Figure 4 compares and analyzes the positioning performance of the GPBP algorithm and the SPAWN algorithm from the two aspects of the cumulative error distribution function and the MSE of the positioning mean square. In order to compare the performance of the two algorithms more intuitively, the simulation of the location performance of the H-SPAWN algorithm using 500 samples is added. From Fig. 4, we can see that SPAWN algorithm has the best location performance, which is because we need to broadcast all samples to the neighbor users in SPAWN algorithm, which brings huge communication cost and computation. The GPBP and H-SPAWN algorithms are based on Gauss parameter method of message passing, with the same order of magnitude as the system overhead, however, because the

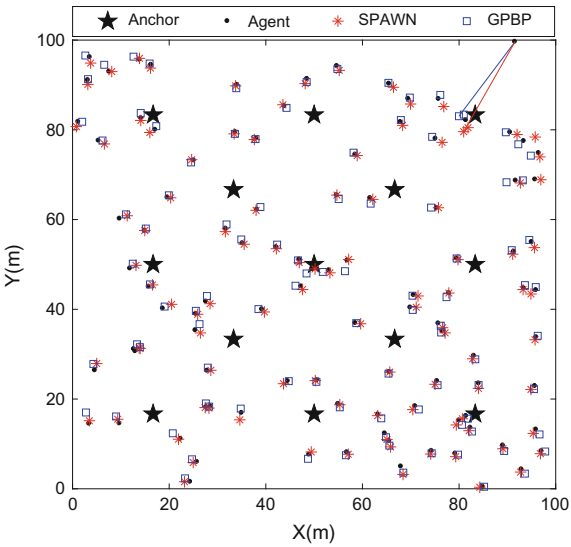


Fig. 3 Single trial performance of GPBP and SPAWN

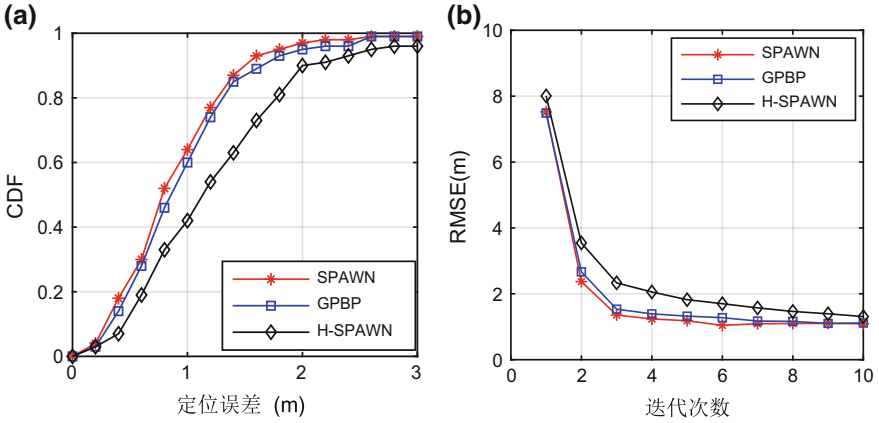


Fig. 4 Localization performance of GPBP, SPAWN and H-SPAWN: **a** CDF; **b** RMSE

message representation method, GPBP makes the sample utilization rate is higher than H-SPAWN, thus greatly enhance the positioning precision of the algorithm. In addition, from Fig. 4b, we can also see that due to the limited number of effective samples of H-SPAWN algorithm, multiple particle resampling is needed to gather particles near the target node position, so that the convergence rate is slower than GPBP.

Three algorithms of the communication overhead and computation are shown in Table 2, can be seen from Table 2, due to the adoption of the Gauss parameter message, GPBP and H-SPAWN each node in each iteration only need to pass the 5 Gauss parameters, and the communication overhead is SPAWN algorithm. At the same time, the calculation of the SPAWN algorithm is proportional to the two times of the number of samples, but the GPBP algorithm proposed in this paper is directly proportional to the number of samples. Considering the system overhead and location performance of the three localization algorithms, the GPBP algorithm proposed in this paper has the advantage of lower communication overhead and computational complexity, without losing the positioning accuracy.

Next, the Monte Carlo simulation method is used to verify the proposed algorithm, and the number of simulation is 100 times.

First, the influence of ranging error on location performance of GPBP algorithm is analyzed. With the different ranging error that $\sigma = 0.2/0.5/1.0/1.5/2.0$ m, Fig. 5a, b are the CDF curve and RMSE curve, respectively. As can be seen from the chart with the increase of the ranging error, the overall positioning performance of cooperative localization network also decreased, especially in $\sigma > 1$ m, the

Table 2 Communication overhead and computational complexity

	GPBP	SPAWN	H-SPAWN
Communication	5	$5k$	5
Complexity	$O(k\|M\|)$	$O(k^2\|M\|)$	$O(k\|M\|)$

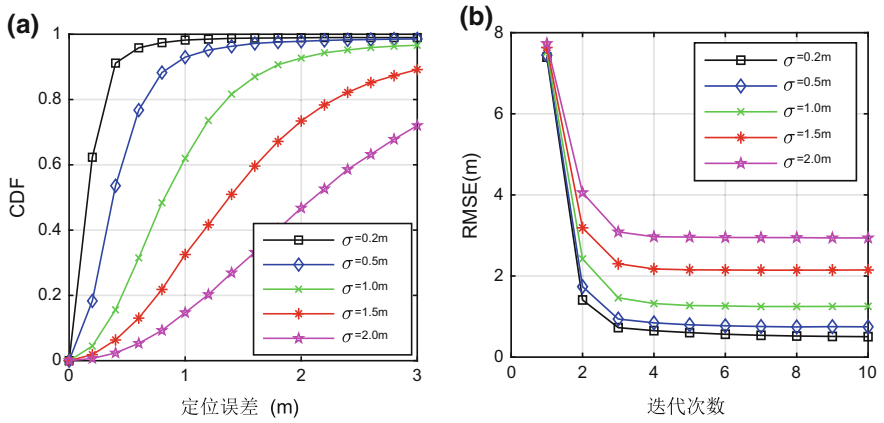


Fig. 5 Localization performance of GPBP versus different noise standard deviations: **a** CDF; **b** RMSE

overall decline in performance is relatively more, this is because in the cooperative localization algorithm in the iterative process, each user node is equivalent to a certain location ambiguity "reference node" the user, between the nodes position uncertainty will influence each other by. In addition, from Fig. 5b can be seen in the distance error under different algorithms in reaching convergence, after 4 iterations are therefore, in actual application, without loss of positioning performance, may be appropriate to reduce the number of iterations so as to further reduce the amount of calculation.

Then we analyze the influence of the density of the anchor nodes and the number of particles used on the positioning performance of the GPBP algorithm. The simulation configures are $\sigma = 1$ m, $N = 9/13$ and $k = 300/800$. From Fig. 6, we

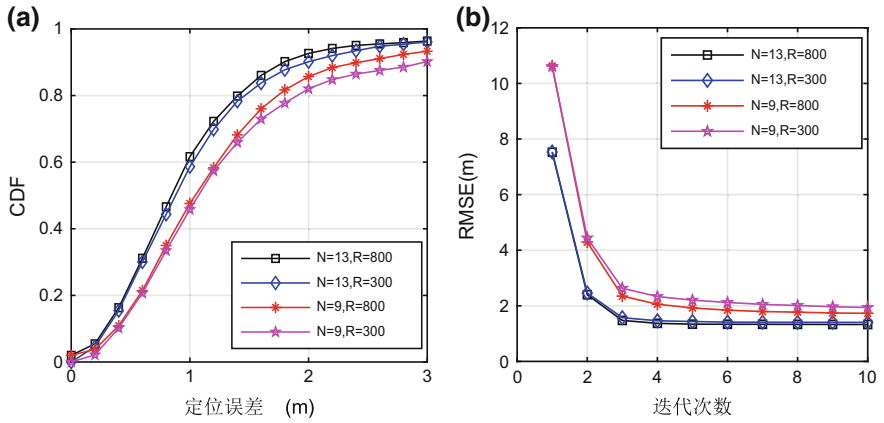


Fig. 6 Localization performances of GPBP with different samples in different anchor numbers: **a** CDF; **b** RMSE

can see the anchor node density has great influence on the accuracy of the algorithms, when the maximum communication distance measuring distance for 20 m cases, in two cases, the average number of connections a single user node and the anchor nodes were 1.07 and 1.58, more anchor nodes can give users provide more accurate location information. Increasing the number of samples can improve the performance of the algorithm to a certain extent. When the density of the reference nodes is large (for example), the performance improvement is not obvious. Therefore, we can consider the anchor node density in cooperative location area and increase or decrease the number of particles appropriately, so as to achieve the trade-off between computation and location performance.

5 Conclusion

Based on the theory of factor graph and belief propagation, a low complexity cooperative localization algorithm with Gaussian parametric message passing is proposed to improve the performance where the non-cooperative localization methods failed due to the insufficient coverage of anchors. Based on the analysis of the link information between the target user and the anchor nodes, the prior spatial information, including the position and variance, of the target is improved. This is conductive to concentrate the samples where the messages have significant mass and capture the salient characteristics of the target's distribution with fewer weighted samples. Rather than broadcast all the weighted samples to the neighbors, Gaussian parametric message passing rule is designed and only 5 parameters need to be exchanged between neighbors. Simulation results show that the proposed algorithm leads to significant reduction in the communication overhead and computational complexity with small loss in localization accuracy.

References

1. Abouzar P, Michelson DG, Hamdi M (2016) RSSI-based distributed self-localization for wireless sensor networks used in precision agriculture. IEEE Trans Wireless Commun 15 (10):6638–6650
2. Yangyang L, Baowang L, Hongwei Z (2014) Indoor pseudolite relative localization algorithm with kalman filter. Acta Physica Sinica 63(22):228402
3. Li B, Wu N, Wang H, Tseng P-H, Kuang J (2015) Gaussian message passing-based cooperative localization on factor graph in wireless networks. Sig Process 2015(111):1–12
4. Ihler AT, Fisher JW, Moses RL, Willsky AS (2005) Nonparametric belief propagation for self-localization of sensor networks. IEEE J Sel Areas Commun 23(4):809–819
5. Wymeersch H, Lien J, Win MZ (2009) Cooperative localization in wireless networks. Process IEEE 97(2):427–450

6. Cakmak B, Urup DN et al (2016) Cooperative localization for mobile networks: a distributed belief propagation—mean field message passing algorithm. IEEE Signal Process Lett 23 (6):828–832
7. Cui J, Wang Z, Zhang C et al (2016) Variational message passing based localization algorithm with Taylor expansion for wireless sensor networks. IET Commun 10(17):2396–2401
8. Velde SV, Abreu TF, Steendam H (2015) Improved censoring and NLOS avoidance for wireless localization in dense networks. IEEE J Sel Areas Commun 33(11):2302–2312

Effective Multipath Mitigation Methods for RTK in Urban Environments

Kai-qin Lin, Zhong-liang Deng and Lu Yin

Abstract RTK is one of the most precise positioning technologies, which has been widely used in many applications. However, in an urban area, GNSS receiver easily receive multipath signals caused by signal reflection and diffraction by both line-of-sight (LOS) and non-line-of-sight (NLOS) satellites. Multipath signals degrade the phase observation quality and makes it difficult to detect integer ambiguities in RTK positioning. This paper proposes two methods to mitigate multipath effect. In the first method, a least square equation is constructed by using the relationship between the classical satellite elevation angle noise model and the signal-to-noise ratio (SNR) noise model. The least square residual is used to determine NLOS and LOS satellites based on the K-means clustering algorithm. The second method mitigates LOS multipath signal based on the consistency check between prediction residuals and prediction covariance of Extended Kalman Filter. Practical experiences show that the proposed tow methods can able mitigate multipath effect and increase the availability of integer ambiguity resolution.

Keywords RTK positionings · Multipath effect · Least square residual
K-means clustering algorithm

1 Introduction

RTK is one of the most precise positioning technologies, which has been widely used in many applications. In the RTK process, integer ambiguity resolution (AR) is one of the key techniques to determine precise positioning solutions. However, in an urban area, GNSS receiver easily receive multipath signals caused by signal reflection and diffraction by both line-of-sight (LOS) and non-line-of-sight (NLOS) satellites [1]. Multipath signals decrease the quality of

K. Lin (✉) · Z. Deng · L. Yin
Beijing University of Posts and Telecommunications, Beijing 100876, China
e-mail: linkaiqin@bupt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_49

code and carrier-phase observations. The degraded observations affect the reliability of float ambiguities determined by Extended Kalman filter and make it difficult to resolve integer ambiguities.

There are many techniques to mitigate the multipath effect. For example, several types of antennas that mitigate the multipath effect, such as a choke ring antenna [2]. Recently, several methods using 3D maps and fisheye view images have recently been proposed to effectively detect NLOS satellites [3].

This paper proposes two methods to mitigate multipath effect. In the first method, a least square equation is constructed by using the linear correlativity between the satellite elevation angle noise model and the SNR noise model. The SNR residuals given by least square are used to determine NLOS and LOS satellites based on the K-means clustering algorithm. However, the multipath signal from a LOS satellite is difficult to detect by the strength of SNR. For those signals, the second method mitigates LOS multipath signal based on the consistency check between prediction residuals and prediction covariance of Extended Kalman Filter. Practical experiences show that the proposed tow methods are able to mitigate multipath errors and increase the availability of integer ambiguity resolution.

2 NLOS Satellites Elimination Method Based on the SNR Residuals of Least Squares

In the RTK process, there are two classical noise models for the carrier phase measurement errors. One of the noise model is based on the satellite elevation angle which can be written as follows [4]:

$$\sigma^2 = a^2 + b^2 / \sin^2 E \tag{1}$$

where σ^2 is the measurement errors of carrier-phase; a and b are coefficients. Another noise model based on the strength of SNR is expressed as [5]:

$$\sigma^2 = B\left(\frac{\lambda}{2\pi}\right) \cdot 10^{-\frac{SNR}{10}} \tag{2}$$

where B is tracking loop bandwidth of carrier-phase; λ is carrier wavelength. E is satellite elevation angle.

Figure 1 shows the strength of the SNR is highly dependent on satellite elevation angle. The higher the elevation angle, the better quality of SNR.

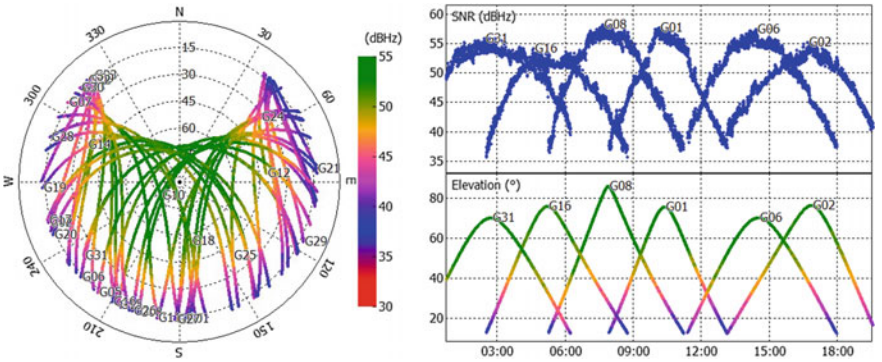


Fig. 1 Skyplot with SNR observations (left); SNR observations and satellites elevation over time in an open sky environment (right)

2.1 *Least Squares Model Based on the SNR and Elevation Angle*

Figure 2 shows the linear correlativity between the variables $1/\sin^2 E$ and $10^{\wedge}(-SNR/10)$ for six satellites. When the satellite elevation angle is high enough, the linear correlativity is very prominent. However, there is a large measurement error in SNR at low satellite elevation angles and there is no longer a prominent linear correlation between the two variables. In this study, elevation angle greater than 28° is considered as high elevation angle and the following equation can be determined by the Eqs. 1 and 2.

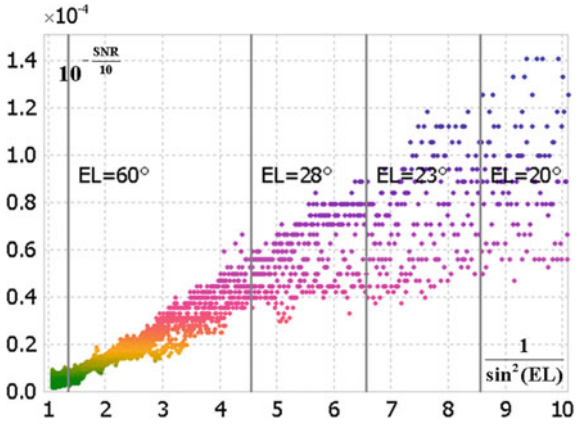


Fig. 2 The linear correlativity between $1/\sin^2(E)$ and $10^{\wedge}(-SNR/10)$

$$\begin{cases} 10^{-\frac{SNR(1)}{10}} = \dfrac{1}{B}\left(\dfrac{2\pi}{\lambda}\right)\left(a^2 + b^2/\sin^2 E^{(1)}\right) + \varepsilon^{(1)} \\ 10^{-\frac{SNR(2)}{10}} = \dfrac{1}{B}\left(\dfrac{2\pi}{\lambda}\right)\left(a^2 + b^2/\sin^2 E^{(2)}\right) + \varepsilon^{(2)} \\ \vdots \\ 10^{-\frac{SNR(n)}{10}} = \dfrac{1}{B}\left(\dfrac{2\pi}{\lambda}\right)\left(a^2 + b^2/\sin^2 E^{(n)}\right) + \varepsilon^{(n)} \end{cases} \tag{3}$$

where, n is the number of satellites tracked by receiver. Suppose there are n satellites tracked by receiver.

Where $\varepsilon^{(i)}$ is the SNR measurement error of the i-th satellite. The least squares equation of Eq. 3 can be expressed as:

$$\Delta \boldsymbol{y} = \boldsymbol{H} \cdot \Delta \boldsymbol{x} + \boldsymbol{\varepsilon} \tag{4}$$

where $\boldsymbol{y}$ is n dimensional SNR measurement vector; $\boldsymbol{x}$ is 3 dimensional coefficient vector; $\boldsymbol{H}$ is $n \times 3$ dimensional observation matrix and $\boldsymbol{\varepsilon}$ is n dimensional SNR measurement noise vector. They are expressed as follows:

$$\boldsymbol{x} = \begin{bmatrix} a & b & B \end{bmatrix}^T \tag{5}$$

$$\boldsymbol{y} = \begin{bmatrix} 10^{-\frac{SNR(1)}{10}} & 10^{-\frac{SNR(2)}{10}} & \cdots & 10^{-\frac{SNR(n)}{10}} \end{bmatrix}^T \tag{6}$$

$$\boldsymbol{H} = \begin{bmatrix} \frac{\partial y^{(1)}}{\partial a} & \frac{\partial y^{(1)}}{\partial b} & \frac{\partial y^{(1)}}{\partial B} \\ \vdots & \vdots & \vdots \\ \frac{\partial y^{(n)}}{\partial a} & \frac{\partial y^{(n)}}{\partial b} & \frac{\partial y^{(n)}}{\partial B} \end{bmatrix} \tag{7}$$

where the i-th element of $\boldsymbol{H}$ are expressed as: $\frac{\partial y^{(i)}}{\partial a} = \frac{2\pi}{\lambda B}\left(2a + b^2/\sin^2 E^{(i)}\right)$, $\frac{\partial y^{(i)}}{\partial b} = \frac{2\pi}{\lambda B}\left(a^2 + 2b/\sin^2 E^{(i)}\right)$, $\frac{\partial y^{(i)}}{\partial B} = -\frac{2\pi}{\lambda B^2}\left(a^2 + b^2/\sin^2 E^{(i)}\right)$

The solution of least squares equation is defined as:

$$\hat{\boldsymbol{x}}_{LS} = \left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\boldsymbol{y} \tag{8}$$

The predicted SNR measurement vector $\hat{\boldsymbol{y}}_{LS}$ determined by $\hat{\boldsymbol{x}}_{LS}$ can be written as:

$$\hat{\boldsymbol{y}}_{LS} = \boldsymbol{H}\hat{\boldsymbol{x}}_{LS} \tag{9}$$

The difference between $\hat{\boldsymbol{y}}_{LS}$ and $\boldsymbol{y}$ is called the SNR residual vector, $\boldsymbol{w}$. In mathematical terms:

$$\boldsymbol{w} = \boldsymbol{y} - \hat{\boldsymbol{y}}_{LS} = \left[\boldsymbol{I} - \boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\right]\boldsymbol{y} = \left[\boldsymbol{I} - \boldsymbol{H}\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\right]\boldsymbol{\varepsilon} \tag{10}$$

The sum of squared errors (SSE) of vector $\hat{\boldsymbol{y}}_{LS}$ is defined as:

$$\boldsymbol{SSE} = \boldsymbol{w}^T\boldsymbol{w} \tag{11}$$

Define the detection threshold T^2 and compare it with $\boldsymbol{SSE}/\sigma_0^2$. If $\boldsymbol{SSE}/\sigma_0^2 > T^2$, it means that there are one or more NLOS multipath signals received by receiver and needed to be eliminated by some NLOS multipath signal elimination methods.

2.2 NLOS Satellites Elimination Method Based on the K-Means Clustering Algorithm

NLOS multipath signals need to be eliminated after detection. NLOS multipath signals not only increase SSE, but also affect the SNR residuals of other normal signals. In this paper, K-means clustering algorithm is used to cluster SNR residuals to separate the NLOS multipath signal from the normal signal. K-means clustering aims to partition n SNR residuals into k clusters in which each residual belongs to the cluster with the nearest mean, serving as a prototype of the cluster. The input parameters of the K-means clustering algorithm are n absolute components of SNR residual vector $\boldsymbol{w}$, and the output parameters are k residual clustering subsets $\boldsymbol{S} = \{S_1, S_2, \ldots, S_k\}$ as to minimize the within-cluster sum of squares $\sum_{i=1}^{k}\sum_{p\in X_i}\|p - m_i\|^2$. Randomly select k SNR residuals as the initial center of k residual clustering subsets. The algorithm proceeds by alternating between two steps:

1. The n residual components are assigned to the nearest subset S_i according to the principle of least squared euclidean distance.

$$S_i = \{p : \|p - m_i\|^2 \le \|p - m_j\|^2 \ \forall j,\ 1 \le j \le k\} \tag{12}$$

2. Calculate the means of the SNR residual of each cluster subset to be the new cluster center.

$$m_i = \frac{1}{|S_i|}\sum_{p\in X_i} p \tag{13}$$

The algorithm has converged when the center no longer change. The clustering subset that has the smallest mean is regarded as the normal satellites clustering subset, while the remaining clustering subsets are considered as the NLOS satellites clustering subset.

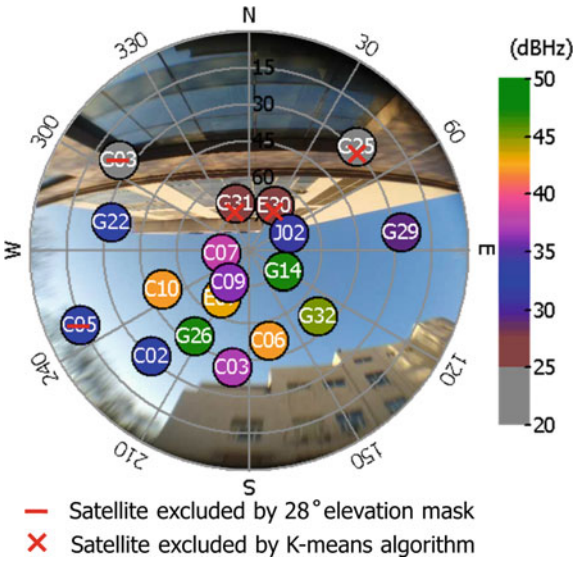


Fig. 3 Skyplot in a severe multipath environment

Figure 3 shows Skyplot in a severe multipath environment. Obviously, G03, G31, E30 and G25 were NLOS satellites. The signals transmitted by them were NLOS signals and whose strength of SNR was decreased by 5–10 dB Hz. SNR observations were used to calculate least squares solution except the observations of G03 and C05 whose elevation angles were less than 28°. Figure 4 shows the clustering results of residuals by using K-means clustering algorithm. Satellites in NLOS satellites cluster subset will not used to calculate RTK positioning.

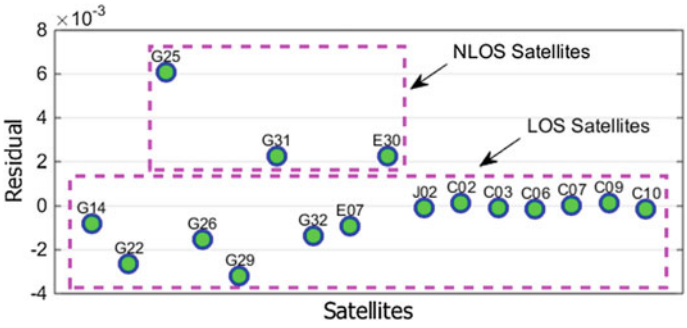


Fig. 4 Clustering results based on the K-means clustering algorithm

3 LOS Multipath Signals Mitigation Method Based on the Prediction Residuals of EKF

In the RTK process, double difference (DD) code and carrier-phase observations are used to determine float ambiguities by Extended Kalman filter. Then, LAMBDA method is used to resolve the integer ambiguities. However, multipath signals decrease the quality of code and carrier-phase observations. The degraded observations increase the prediction residuals of Extended Kalman filter and make it difficult to resolve integer ambiguities. This section proposed a consistency check method of EKF prediction residuals to mitigate multipath signals. The DD observation models can be written as follows [6]:

$$\boldsymbol{h} = \begin{cases} P_{ub}^{s_2 s_1} = \rho_{ub}^{s_2 s_1} \\ \phi_{ub}^{s_2 s_1} = \lambda^{-1}\rho_{ub}^{s_2 s_1} + N_{ub}^{s_2 s_1} \\ \vdots \\ P_{ub}^{s_n s_1} = \rho_{ub}^{s_n s_1} \\ \phi_{ub}^{s_n s_1} = \lambda^{-1}\rho_{ub}^{s_n s_1} + N_{ub}^{s_n s_1} \end{cases} \tag{14}$$

where $P_{k,rb}^{s_2 s_1}$ is DD pseudorange measurement; $\phi_{k,rb}^{s_2 s_1}$ is DD carrier-phase measurement and $N_{ub}^{s_2 s_1}$ is DD integer ambiguity. The Kalman filter state vector are position, velocity, acceleration and DD integer ambiguity which is defined as:

$$\boldsymbol{x}_u = \begin{bmatrix} \boldsymbol{r}_u & \boldsymbol{v}_u & \boldsymbol{a}_u & \mathrm{N}_{ub}^{s_2 s_1} & \cdots & \mathrm{N}_{ub}^{s_n s_1} \end{bmatrix}^T \tag{15}$$

Kalman filter update phase is expressed as:

$$\tilde{\boldsymbol{x}}_k = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{w}_k \tag{16}$$

$$\tilde{\boldsymbol{P}}_k = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{P}}_{k-1}\boldsymbol{F}_{k,k-1}^T + \boldsymbol{Q}_k \tag{17}$$

where $\boldsymbol{F}_{k,k-1}$ is state transition matrix and $\tilde{\boldsymbol{x}}_k$ is the predicted state vector. The Kalman filter update phase uses the following equation:

$$\begin{aligned} \boldsymbol{K}_k &= \tilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T(\boldsymbol{H}_k\tilde{\boldsymbol{P}}_k\boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1} \\ \hat{\boldsymbol{x}}_k &= \tilde{\boldsymbol{x}}_k + \boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k) \\ \hat{\boldsymbol{P}}_k &= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\tilde{\boldsymbol{P}}_k \end{aligned} \tag{18}$$

where $\boldsymbol{R}_k$ is the covariance matrix of observation noise vector $\boldsymbol{v}_k$. $\boldsymbol{K}_k$ is the Kalman optimal gain matrix. $\hat{\boldsymbol{x}}_k$ is a state vector estimated by Kalman filter and $\hat{\boldsymbol{P}}_k$ is the covariance estimate of $\hat{\boldsymbol{x}}_k$. The prediction residual and prediction covariance of Kalman filter can be written as:

$$\boldsymbol{v} = \boldsymbol{y} - \boldsymbol{H}\tilde{\boldsymbol{x}}_k \tag{19}$$

$$\boldsymbol{D} = \boldsymbol{R} + \boldsymbol{H}\tilde{\boldsymbol{P}}_k\boldsymbol{H}^T \tag{20}$$

NLOS multipath signals will increase the prediction residuals $\boldsymbol{v}$. By checking the consistency of $\boldsymbol{v}$ and $\boldsymbol{D}$, we can determine whether the residual v_i is not an outlier:

$$|v_i| > C \cdot \sqrt{d_{i,i}} \tag{21}$$

where $d_{i,i}$ is the ith diagonal element of the covariance matrix $\boldsymbol{D}$ and C is the balance factor, usually take 2 or 3. If the above formula is satisfied, it is mean that v_i is an outlier and the observation of i-th satellite will be eliminated.

4 Experiment and Results

To evaluate the proposed two methods in this paper. Static RTK experiments were performed in two environments with severe multipath signals. The rover and base stations were both ublox-M8T multi constellation single frequency receiver with 1 Hz observations. The baseline length between them was within 500 m. Figure 7 shows the Skyplot at two different test points. These were Skyplot with SNR observations at each point for GPS, QZSS, Beidou and Galileo. Point 1 was next to a mid-rise building and point 2 was surrounded by one high-rise and one low-rise building. The two multipath mitigation methods proposed in this paper are as follows:

Method 1: NLOS satellites elimination method based on the SNR residuals of least squares.

Method 2: LOS multipath signals mitigation method based on the prediction residuals of EKF.

The experiment results are compared by the following three schemes:

Scheme 1: Normal RTK
Scheme 2: Normal RTK+Method 1
Scheme 3: Normal RTK+Method 1+Method 2

In this study, the parameters of Normal RTK are set to 20° mask angle and SNR above 30 dB Hz. The experimental evaluation indicators are given bellow.

1. Number of valid satellites: The average number of valid satellites that are used for RTK solution.
2. NLOS satellites exclusion rate: That is the ratio of the number of NLOS satellites excluded by RTK to the number of NLOS satellites determined by fisheye view.

3. Ambiguity availability: That is the ratio of the number of fix solutions and the number of epochs. If the ambiguity ratio value was above the ratio test threshold, the solution was considered as fix solution.
4. Ambiguity reliability: That is the ratio of the number of reliable solutions and the number of fix solutions. If the horizontal error of the fix solutions was within 0.1 5 m, these solutions were considered as reliable solutions (Fig. 5).

Tables 1 and 2 shows the experimental results at each point. The proposed two methods can increase both the NLOS satellites exclusion rate and ambiguity reliability. The Ambiguity availability at point 2 was significantly lower than point 1, it means that there were more severe multipath signals at point 2. The visible area in the zenith direction at point 2 was smaller due to the blockage of the two buildings. The smaller visible area decreased the number of valid satellites and make it difficult to resolve integer ambiguity.

For the NLOS satellites exclusion rate, point 2 has lower performance improvement. In this situation, receiver was easy to receive more NLOS multipath signals. A large number of NLOS multipath signals decreased the solution of least squares and make it difficult to cluster least squares residuals by using K-means algorithm.

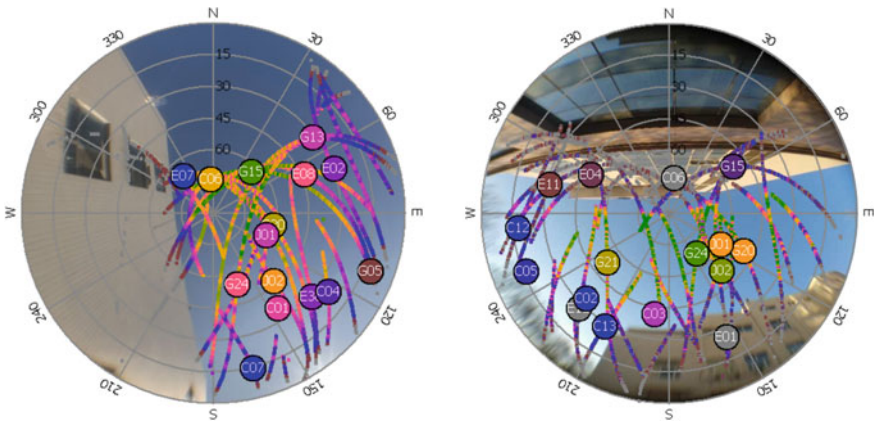


Fig. 5 Skyplot with SNR observations at each point

Table 1 Experiment results at point 1 with 6 h observation

	Number of valid satellites	NLOS satellites exclusion rate	Ambiguity availability	Ambiguity reliability
Scheme 1	11.51	56.33% (17257/30635)	87.27% (19167/21962)	99.96% (19161/19167)
Scheme 2	10.89	73.07% (22385/30635)	94.17% (20683/21962)	100% (20683/20683)
Scheme 3	10.77	74.83% (22924/30635)	94.65% (20788/21962)	100% (21962/21962)

Table 2 Experiment results at point 2 with 3 h observation

	Number of valid satellites	NLOS satellites exclusion rate	Ambiguity availability	Ambiguity reliability
Scheme 1	9.07	69.88% (72329/103502)	9.54% (1100/11528)	52.72% (580/1100)
Scheme 2	6.53	74.78% (77399/103502)	8.44% (973/11528)	57.34% (558/973)
Scheme 3	6.32	78.61% (81363/103502)	8.98% (1035/11528)	61.96% (641/1035)

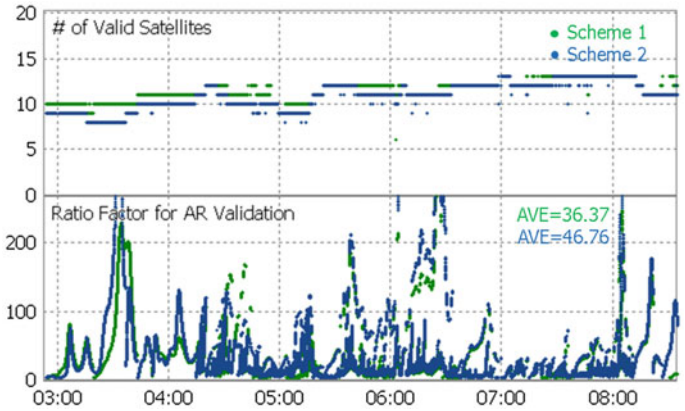


Fig. 6 Number of valid satellites and ratio factor at point 1 for scheme 1 and scheme 2

Figure 6 shows the results of the number of valid satellites and ambiguity ratio factor at point 1 for scheme 1 and scheme 2. Based on the results, applying the NLOS multipath signals elimination method improved the results, as can be seen from the number of valid satellites and the ratio factor. Generally, a higher ratio factor means that a higher ambiguity reliability can obtained. In addition, from the results in Table 1, the proposed LOS multipath signals mitigation method can future improve the ambiguity availability and reliability. Unfortunately, the proposed two methods can not improve the positioning accuracy of reliable solution as shown in Fig. 7.

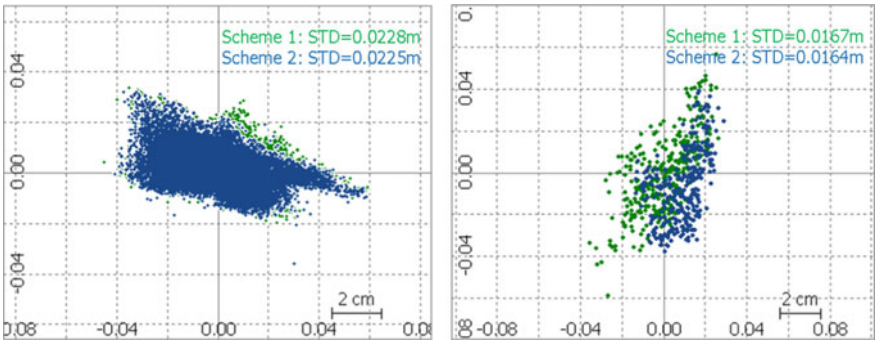


Fig. 7 Positioning error of reliable solutions at each point for scheme 1 and 2

5 Conclusions

This paper evaluated proposed two multipath mitigation methods for improving the RTK performance in urban environment. For NLOS multipath signals, a NLOS satellites elimination method based on the SNR residuals of least squares to select good-condition satellites. For LOS multipath signals that is difficult to detect by SNR residuals, a LOS multipath mitigation method based on the consistency check between prediction residuals and prediction covariance of EKF was proposed.

Static RTK experiments were performed in two environments with severe multipath signals. Based on the test results, the proposed NLOS satellites elimination method can effectively increase the ambiguity reliability of RTK. Unfortunately, the LOS multipath signals mitigation method has a limited ability to maintain high availability and reliability, this may be due to the lack of visible satellites and use only single-frequency signal. Thus, we would use multi-constellation and multi-frequency receiver for further experiments.

Acknowledgements Foundation project: National Science and Technology Major Project (2016ZX03001023-002).

References

1. Groves PD, Jiang Z, Wang L, et al (2012) Intelligent urban positioning using multi-constellation GNSS with 3D mapping and NLOS signal detection
2. Rost C, Wanninger L (2009) Carrier phase multipath mitigation based on GNSS signal quality measurements. J Appl Geodesy 3(2):81–87
3. Tokura H, Kubo N (2016) Effective satellite selection methods for RTK-GNSS NLOS exclusion in dense urban environments. In: Proceedings of the 29th international technical meeting of the satellite division of the institute of navigation (ION GNSS+2016), Oregon, September 2016

4. RTKLIB (2013) RTKLIB ver. 2.4.2 Manual, April 29
5. Hartinger H, Brunner FK (1999) Variances of GPS phase observations: the SIGMA-ε model. GPS Solutions 2(4):35–43
6. Takasu T, Yasuda A (2009) Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. In: International symposium on GPS/GNSS. International Convention Centre, Jeju, Korea, 2009 pp 4–6

A Frequency Estimation Method for Improving Acquisition Accuracy of TC-OFDM

Zhongliang Deng, Yin Yang, Jun Mo, Wenxuan Song, Xinmei Bian and Zongshan Tang

Abstract Aiming at the shortcomings that existing traditional receivers have a low frequency estimation accuracy in terrestrial fading channels, this paper presents a high-precision frequency estimation method. Using multiple parallel down converters and integral unit to down convert and integrate the stored RF signals with multiple local frequencies, then using cubic spline interpolation algorithm to fit the integral result to finally obtain the residual carrier frequency estimation value. The simulation results show that in the terrestrial fading channels, compared with the traditional linear frequency search method, the proposed algorithm can improve the capture frequency estimation accuracy of TC-OFDM receiver, the mean value of the frequency estimation error is less than 1 Hz reaching the request of the receiver phase-locked loop.

Keywords TC-OFDM · Capture · Parallel frequency search · Frequency estimation

1 Introduction

The traditional linear frequency search method is widely used in GNSS receivers because of the existence line-of-sight channels between the satellite and the receiver. However, in the Time & Code Division-Orthogonal Frequency Division Multiplexing (TC-OFDM) system, the channel from the base station to the receiver is a typical random access channel [1]. Compared with satellite channels, the variation of terrestrial channel parameters is larger, which leads to the failure of linear frequency search. Therefore, how to improve the success rate and efficiency of TC-OFDM receiver in frequency-capture on the ground-weak channel is an important research direction.

Z. Deng · Y. Yang (✉) · J. Mo · W. Song · X. Bian · Z. Tang
School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China
e-mail: yangyin@bupt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_50

Compared with the GNSS, the residual carrier of the TC-OFDM receiver is mainly caused by the crystal error [2] and needs to be peeled off by the digital controlled oscillator and the digital mixer in the receiver baseband processor. The receiver must make the residual carrier less than a certain threshold before it enters the tracking loop. The process of acquiring residual carriers is called frequency capture or frequency search [1]. In the GNSS receivers, FFT search method and linear frequency search method are commonly used for frequency search [3–5]. Among them, the FFT search method requires a dedicated FFT module and amount of computation is large. And the linear frequency search method is greatly affected by the channel and is not suitable for TC-OFDM systems working on the terrestrial weak channel.

Aiming at the problem of TC-OFDM frequency acquisition on the weak ground channel, this paper presents a high-precision frequency estimation method. By using a plurality of parallel down-converters and integrator cleaners, the stored radio frequency signals are down converted and integrated with a plurality of local frequencies, then using the cubic spline interpolation algorithm to fit the integral result and finally the residual carrier frequency estimate is obtained. The simulation results show that in the terrestrial fading channels, compared with the traditional linear frequency search method, the proposed algorithm can improve the acquisition accuracy of TC-OFDM receiver, the mean value of the frequency estimation error is less than 1 Hz reaching the request of the receiver phase-locked loop.

2 Parallel Frequency Search Model

2.1 Ground Weak Channel

The traditional linear frequency search method is widely used in GNSS receivers because of the line-of-sight channels between the satellite and the receiver. According to literature [3], in TC-OFDM system, the channel from base station to receiver is a typical random channel, and the amplitude A_{IF} of the signal is a Rayleigh distribution random variable. Since A_{IF} fluctuates randomly and violently many times in one second, the non-related integral value P obtained from Eq. (10) also fluctuates violently, which resulting the failure of traditional linear frequency search.

2.2 Parallel Frequency Search Method

In order to adapt to the terrestrial fading channel and solve the influence of the rapid and random change of the signal amplitude on the frequency search, the TC-OFDM system adopts the parallel frequency search method, and the algorithm structure diagram is shown in Fig. 1.

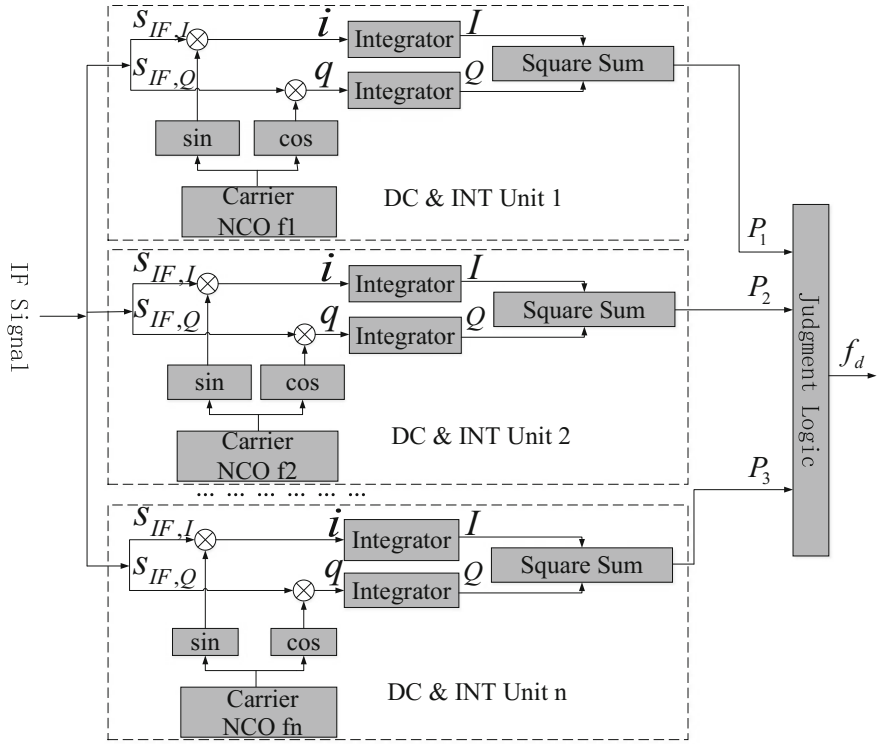


Fig. 1 Structure of parallel residual carrier frequency searching

In this architecture, there are multiple parallel down-converter & integration units that can down-convert and integrate with multiple local frequencies simultaneously. Each of these units is similar to the traditional linear search method.

IF signal can be obtained from the RF signal that received by the antenna through down-conversion at RF front-end and filter by low-pass filter. The IF signal is split into two channels I, Q into the baseband processor, namely:

$$s_{IF,I}^{(i)}(t) = A_{IF}^{(i)} m^{(i)}(t-\tau^{(i)}) c^{(i)}(t-\tau^{(i)}) \cos[2\pi(f_{IF}+f_d^{(i)})t + \theta_{IF}^{(i)}] + n_I(t) \quad (1)$$

$$s_{IF,Q}^{(i)}(t) = A_{IF}^{(i)} m^{(i)}(t-\tau^{(i)}) c^{(i)}(t-\tau^{(i)}) \sin[2\pi(f_{IF}+f_d^{(i)})t + \theta_{IF}^{(i)}] + n_Q(t) \quad (2)$$

In the formula, $m^{(i)}$ is the navigation message, $A_{IF}^{(i)}$ is the IF signal amplitude, $c^{(i)}$ is the spreading code, $\tau^{(i)}$ is the propagation delay of signal from base station to receiver antenna, $f_d^{(i)}$ is the residual frequency caused by Doppler frequency and receiver crystal deviation, $\theta_{IF}^{(i)}$ is the first phase of the IF, f_{IF} is the IF signal, n_I,n_Q are the noise signal. Superscript (i) represents signals from different base stations.

The signals of each base station are initially mixed at the receiver antenna and then down-converted via the RF front end. The signal received by the baseband processor is

$$s_{IF}(t) = \sum_i (s_{IF,I}^{(i)} + j s_{IF,Q}^{(i)}) + n(t) \tag{3}$$

where $n(t)$ the noise signal.

The digitally controlled oscillator in the baseband processor produces mutually orthogonal sine and cosine signals:

$$u_{os}(t) = \sin(2\pi f_{NCO} t + \theta_{NCO}) \tag{4}$$

$$u_{oc}(t) = \cos(2\pi f_{NCO} t + \theta_{NCO}) \tag{5}$$

Among them, f_{NCO} and θ_{NCO} are respectively the numerical control oscillator frequency and initial phase. In order to further strip the carrier, the carrier striping module in the baseband processor performs the following operation:

$$\begin{aligned} i &= s_{IF,I} \cdot u_{oc} + s_{IF,Q} \cdot u_{os} \\ &= A_{IF} m(t-\tau) c(t-\tau) \cos[2\pi(f_d - f_{NCO})t + (\theta_{IF} - \theta_{NCO})] \end{aligned} \tag{6}$$

$$\begin{aligned} q &= s_{IF,Q} \cdot u_{oc} - s_{IF,I} \cdot u_{os} \\ &= A_{IF} m(t-\tau) c(t-\tau) \sin[2\pi(f_d - f_{NCO})t + (\theta_{IF} - \theta_{NCO})] \end{aligned} \tag{7}$$

In a zero IF receiver, f_d is zero. In order to facilitate the analysis, the noise and superscript (i) are ignored here.

In the capture process of the TC-OFDM system, the capture of the TXID portion has completed the pseudo-code synchronization. Therefore, the following content is carried out under the condition which the pseudo-code synchronization has been completed. The frequency capture is to find the frequency f_L of local numerically controlled oscillator, and make f_L approximately equal to f_d. Since the period of the navigation message is much longer than the integration period, then assumed that $m(t)$ is a constant value in one integration time. The integrator in the down-conversion & integral unit of Fig. 1 performs the integral operation as follows:

$$\begin{aligned} I &= \int_0^{T_s} A_{IF} c^2(t-\tau) \cos[2\pi(f_d - f_{NCO})t + (\theta_{IF} - \theta_{NCO})] dt \\ &= A_{IF} T_s \mathrm{sinc}[(f_d - f_{NCO})T_s] \cos[\pi(f_d - f_{NCO})T_s + (\theta_{IF} - \theta_{NCO})] \end{aligned} \tag{8}$$

$$\begin{aligned} Q &= \int_0^{T_s} A_{IF} c^2(t-\tau)\sin[2\pi(f_d - f_{NCO})t + (\theta_{IF} - \theta_{NCO})]dt \\ &= A_{IF} T_s \mathrm{sinc}[(f_d - f_{NCO})T_s]\sin[\pi(f_d - f_{NCO})T_s + (\theta_{IF} - \theta_{NCO})] \end{aligned} \tag{9}$$

Then the final non-coherent integration value is:

$$P = I^2 + Q^2 = A_{IF}^2 T_s^2 \mathrm{sinc}^2[(f_d - f_{NCO})T_s] \tag{10}$$

As can be seen from the above equation, when the integral time T_s and the intermediate frequency amplitude A_{IF} is a fixed value, the output of the non-coherent integration is a function of the difference between the numerically controlled oscillator frequency f_{NCO} and the residual carrier frequency f_d, the image of which is shown in Fig. 2

According to the nature of the sinc function, the maximum of the integral output P appears where the frequency offset is zero, the value of P is symmetric about $f_d - f_{NCO} = 0$, and the width of the main lobe is $2/T_s$. In the process of parallel frequency search, Multiple Carrier NCOs in different parallel down convertor and integrator (DC&INT) units generate multiple local frequencies $f_{NCO1}, f_{NCO2}, \ldots, f_{NCOn}$ at the frequency interval Δf, After the integration of interval T_s, n results $P_1, P_2, \ldots, P_n$ are obtained. Compared all P values to get the maximum value and the corresponding frequency is the closest to the real. In order to ensure that there is at least one search frequency within the width of the main lobe, the maximum frequency interval Δf can not exceed $1/T_s$.

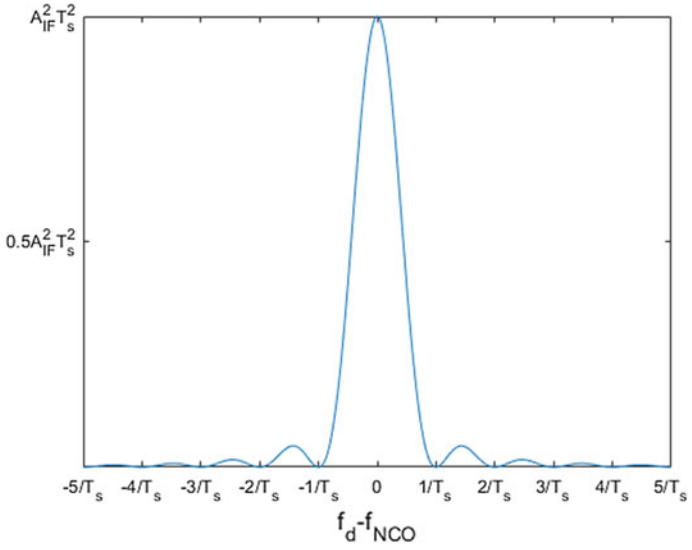


Fig. 2 Relationship between integral output and frequency difference

In the parallel frequency search method, the frequency search accuracy depends on the size of the preset frequency interval Δf, and the maximum error is $\Delta f/2$. In a certain frequency band, the smaller the frequency interval Δf is, the more accurate the frequency search result is. However, the more the preset frequency points are, the more the number of cells is and the more computation is required, which is hard to implement in hardware. The residual carrier estimation method can greatly improve the accuracy of frequency search with only a small amount of computation added.

3 Non-coherent Early Mix Late Method

3.1 Least Squares Fitting Algorithm

According to the nature of the function, in the field of the main lobe width, chose multiple frequency points integral results to perform quadratic fitting algorithm. The mathematical model is

$$Y = ax^2 + bx + c \tag{11}$$

Select the three largest n results of the integration of the three results p_j, p_k, p_l, then make up the array p_j, p_k, p_l, which corresponding local frequency f_{NCOj}^2, f_{NCOk}, f_{NCOl}. According to the above Eq. (11), three equations are established. Write it as a matrix:

$$Hu = Y \tag{12}$$

Among them:

$$H = \begin{bmatrix} f_{NCOj}^2 & f_{NCOj} & 1 \\ f_{NCOk}^2 & f_{NCOk} & 1 \\ f_{NCOl}^2 & f_{NCOl} & 1 \end{bmatrix}, \quad u = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad Y = \begin{bmatrix} P_j \\ P_k \\ P_l \end{bmatrix}$$

where $u = [\,a \quad b \quad c\,]^T$ is the coefficient of the quadratic curve and can be found from the above equation:

$$u = (H^T H)^{-1} H^T Y = H^{-1} Y \tag{13}$$

Finally, according to the vertex formula of the quadratic curve, the corresponding frequency of the vertex fitting is taken as the result of the parallel frequency acquisition:

$$f_d = -\frac{b}{2a} \tag{14}$$

3.2 Cubic Spline Interpolation

The cubic spline interpolation uses the special piecewise polynomial interpolation, which not only guarantees the sub-order low-order interpolation polynomial, but also improves the smoothness of the interpolation function. Using cubic spline interpolation can not only achieve smaller interpolation error but also avoid the Runge phenomenon which occurs when using higher order polynomials.

Let $[a, b]$ be an interpolated node, $a = x_0 < x_1 < \cdots < x_n < b$, given the function $y_i = f(x_i)$, $i = 0, 1, \ldots, n$ at each node x_i. If function $S(x)$ satisfies the conditions that $S(x) = y_i (i = 0, 1, \ldots, n)$, $S(x)$ are all the polynomial which order is three times and below three times. When $S(x)$ has a second-order continuous derivative at $[a, b]$. $S(x)$ is called cubic spline interpolation function.

In a cubic spline interpolation function, it is required that $S(x)$ only need to determine one polynomial in each subinterval $[x_i, x_{i+1}]$.

$$S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i, (i = 0, 2, \ldots, n-1) \tag{15}$$

where a_i, b_i, c_i, d_i is to be determined and the following continuity conditions are to be satisfied:

$$S(x_i - 0) = S(x_i + 0),$$

$$S'(x_i - 0) = S'(x_i + 0),$$

$$S''(x_i - 0) = S''(x_i + 0).$$

In this case, there are $3n - 3$ conditions in total. When $S(x)$ satisfies the interpolation condition $S(x) = y_i (i = 0, 1, \ldots, n)$, there are $4n - 2$ conditions.$S(x)$ needs to determine 4 undetermined coefficient on each small inter-cell $[x_i, x_{i+1}]$ and there has a total of n small inter-cells. Therefore, there are $4n$ parameter need to be determined. In order to uniquely identify cubic spline interpolation functions, two additional boundary conditions are added. Common boundary conditions are: first-order boundary conditions, second-order boundary conditions, periodic boundary conditions. The actual problem is usually given by the state requirements of the cubic spline interpolation at the endpoints.

First-order boundary conditions: the first derivative value at the given endpoint, $S'(x_1) = y'_1$, $S'(x_n) = y'_n$.

Second-order boundary conditions: the second derivative value at the given endpoint, $S''(x_1) = y''_1$, $S''(x_n) = y''_n$.

Periodic boundary conditions: If y $= f(x)$ it is a periodic function with period of $b - a$, then $S(x)$ satisfies the conditions at the endpoints that

$$S'(x_1 + 0) = S'(x_n - 0), \quad S''(x_1 + 0) = S''(x_n - 0)$$

In this paper, using the first-order boundary conditions, the second-order derivative at the node is used to represent the cubic spline interpolation function.

Let $h_i = x_i - x_{i-1}$, $S''(x_i) = M_i (i = 0, 1, \ldots, n)$, obtained by cubic Hermite interpolation, n-1 equations with n pending parameters M_i:

$$\mu_j M_{j-1} + 2M_j + \lambda_j M_{j+1} = g_j (j = 2, 3, \ldots, n-1) \tag{16}$$

where, $\begin{cases} \mu_j = h_{j-1}/(h_{j-1} + h_j) \\ \lambda_j = 1 - \mu_j \\ g_j = 6\left(\frac{y_{j+1} - y_j}{h_j} - \frac{y_j - y_{j-1}}{h_{j-1}}\right)\frac{1}{h_{j-1} + h_j} \end{cases}$

Bring the boundary conditions $S'(x_1) = y'_1$, $S'(x_n) = y'_n$, Write the system of equations in matrix form:

$$\begin{bmatrix} 2 & \alpha_1 & & & \\ \lambda_1 & 2 & \alpha_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \lambda_{n-1} & 2 & \alpha_{n-1} \\ & & & \lambda_n & 2 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_{n-1} \\ M_n \end{bmatrix} = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_{n-1} \\ g_n \end{bmatrix} \tag{17}$$

where $\alpha_1 = 1$, $\lambda_n = 1$, $g_1 = \frac{6}{h_1}\left(\frac{y_2 - y_1}{h_1} - y'_1\right)$, $g_n = \frac{6}{h_{n-1}}\left(y'_n - \frac{y_n - y_{n-1}}{h_{n-1}}\right)$.

The coefficient matrix of the system of equation set (17) is tridiagonal and diagonally dominant. Therefore, there is a unique solution and then take the M into Eq. (16). Finally, we can construct the interpolation function $S(x)$ in the interval of $[a, b]$.

4 Algorithm Simulation and Analysis

By establishing the parallel frequency search model, this experiment simulates and analysis not only the cubic spline interpolation performance but also the performance of least squares fitting as compared. In MATLAB platform, this paper uses TC-OFDM system 8191 Code as a pseudo-random code, the integration time is 1.6382 ms, the signal center frequency is 754 MHz, the signal to noise ratio is −20 dB, the number of simulation is 100 and each simulation have 500–599 Hz

different residual carrier RF signal data. The following uses the RF signal data with residual carrier at 502 Hz as an example for simulation analysis.

4.1 Algorithm Steps

(1) The signal data is multiplied with the locally generated local code to align the spreading code with the local code.
(2) Using 17 down-conversion & integration units to search the parallel frequency in the upper and lower 1600 Hz bands at 200 Hz intervals and then calculate seventeen related results. The f_{NCO} corresponding to the maximum correlation value is the residual carrier value obtained by the parallel frequency search.
(3) Select three largest related results from the maximum of the 17 largest correlation values P_i, $i = 1, 2, \ldots, 15$ of parallel search to calculate the least squares fit curve and the maximum value of the curve is the residual carrier frequency value of the fine parallel frequency search.
(4) Select the maximum correlation value of the maximum correlation value P_i, $i = 1, 2, \ldots, 15$ and the two sides of each of the two correlation results, a total of five, and calculate the cubic spline interpolation curve. The maximum value of the curve is the residual carrier frequency value of the fine parallel frequency search.

4.2 Experimental Results

In Fig. 3, using an integration time of 1.6382 ms, 1 Hz for the frequency interval, and the results obtained are shown in solid lines in the figure. As can be seen from the figure, the maximum of the solid line appears at 502 Hz. Therefore, we can see that the residual frequency of the collected IF data is 502 Hz.

In actual receivers, the hardware resources are not sufficient to support parallel frequency search at 1 Hz intervals. The experiment uses 17 down-conversion & integral units to calculate the correlation results at intervals of 200 Hz. These results are marked with open squares in Fig. 3. It can be seen that 600 Hz frequency corresponding to the most relevant value and the result is shown in Fig. 3 as a solid square. In this case, the residual carrier frequency is considered as 600 Hz which the actual situation is quite different. In contrast, the result of the conventional linear frequency search is indicated by an asterisk in Fig. 3. It can be seen from the figure that the traditional linear frequency search has the largest correlation value at −200 Hz and does not match with the result of the solid line, so that the correct frequency search result can not be obtained.

In Fig. 4, under the same conditions, take three of the maximum correlation results of parallel frequency search, after using the least squares fitting method,

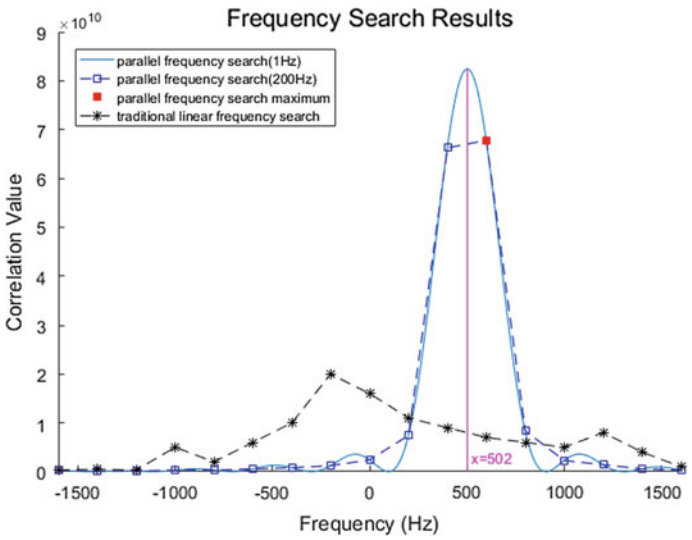


Fig. 3 Results of parallel frequency search

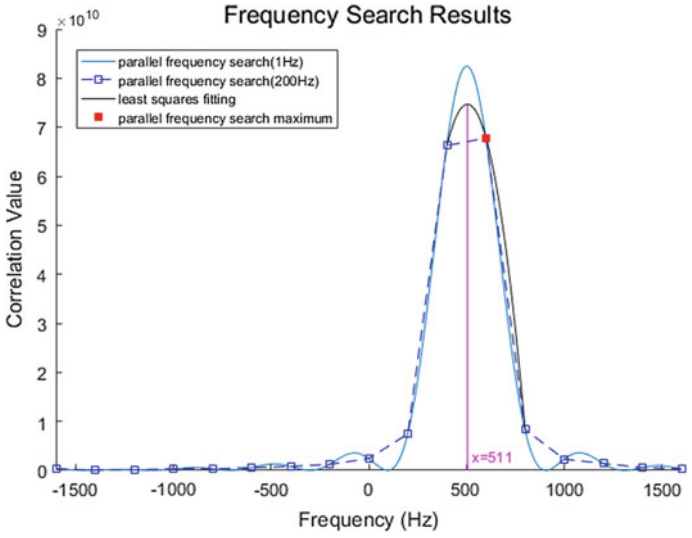


Fig. 4 Least squares fitting curve figure

using the maximum value of the curve at the frequency of 511 Hz as the parallel frequency search residual carrier frequency values. It can be seen that the error is 9 Hz compared with the actual residual frequency, which greatly improving the frequency search accuracy.

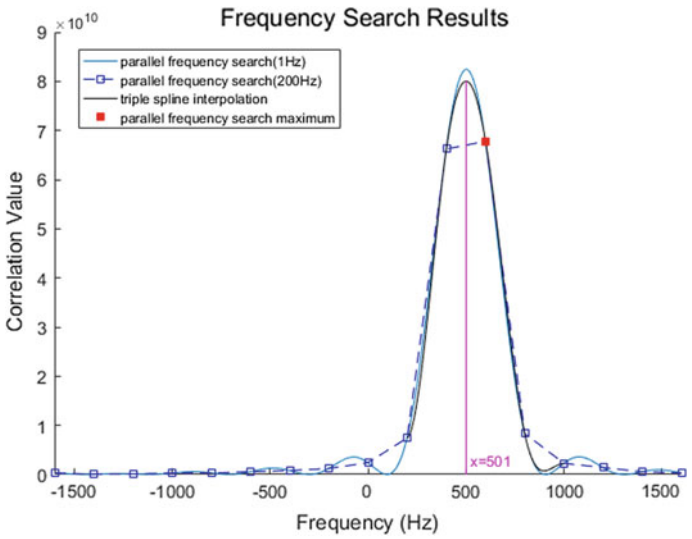


Fig. 5 Cubic spline interpolation curve

Table 1 Statistical property of outcome under different acquisition methods

Statistical characteristics	Residual carrier measurement method		
	Parallel frequency search	Least squares fitting	Triple spline interpolation
Error mean	50.5	8.15	0.73
Mean square error	833.25	3.33	0.28

In Fig. 5, under the same conditions, the correlation results of the maximum parallel frequency search and two correlation results of both sides are selected, for a total of five. Then use the cubic spline interpolation and the adopt the value of the curve at the maximum value of 501 Hz as the parallel frequency search residual carrier frequency value. As can be seen, the error from the actual residual frequency is 1 Hz, which is higher than that of the least squares method.

In Table 1, the accuracy of parallel frequency search results depends on the set frequency interval Δf, the maximum error of $\Delta f/2$. After least-squares fitting, the error is less than the parallel frequency search, and the average error is about 1/6 of the parallel frequency search. After cubic spline interpolation, the mean error is about 1/11 of the least squares fit. The results show that cubic spline interpolation has higher fitting accuracy. Meanwhile, the variance of the capture error of the cubic spline interpolation is also smaller than that of the least square fitting. This shows that the error caused by the cubic spline interpolation on the variation of the residual carrier changes little when capturing signals containing different residual carriers.

5 Conclusions

Based on the analysis of the traditional linear frequency search method, a high-precision frequency estimation method is proposed and simulated. Using multiple parallel downconverters and integral unit cleaner to down convert and integrate the stored RF signals and multiple local frequencies, using cubic spline interpolation algorithm to fit the integral result to finally obtain the residual carrier frequency estimation value. By comparing and analyzing the parallel frequency estimation algorithm of cubic spline interpolation with the traditional linear frequency search algorithm and the least square fitting algorithm, we can draw the conclusion that first of all the parallel frequency search method has the advantages of high frequency acquisition success rate and high frequency estimation accuracy under the weak terrestrial channel compared with the traditional linear frequency search method, secondly the fitting frequency estimation accuracy of cubic spline interpolation is higher than that of least square fitting frequency estimation, thirdly, the frequency estimation error of the parallel frequency estimation method including cubic spline interpolation is less than 1 Hz and the error variance is less than 0.5 Hz^2 which fully meets the request of loop pull-in requirement of the phase-locked loop.

Acknowledgements Foundation project: The National Key Project of Research and Development Plan of China (2016YFB0502003).

References

1. Deng Z, Yuan X, Yu Y, et al (2012) A method and realization of parallel residual carrier frequency searching for base station positioning system. In: International conference on wireless communications, networking and mobile computing, pp 1–4
2. Deng Z, Mo J, Jia B et al (2017) An acquisition scheme based on a matched filter for novel communication and navigation fusion signals. Sensors 17(8):1766
3. Wei J, Fan S, Zhou C, Liu J (2014) Research on the tracking algorithm of GPS weak signal under multipath interference. J Yunnan Nationalities Univ Nat Sci Ed 05:373–378
4. Rinder P, Bertelsen N (2004) Design of a single frequency GPS software receiver. Master Thesis, Aalborg University, Denmark
5. Janos G, Andrei C (2012) A simplified adaptive Kalman filter algorithm for carrier recovery of M-QAM signals. In: Proceedings of the 17th international conference on methods and models in automation and robotics (MMAR), Szczecin, Poland, pp 303–307
6. Xu X, Zhong T (2013) Construction and realization of cubic spline interpolation function. Autom Meas Contr 25(11):76–78
7. Deng Z, Yu Y, Yuan X, Wan N, Yang L (2013) Situation and development tendency of indoor positioning. China Commun 10(3):P42–P55
8. Shen F, Li W, Li Q (2015) GPS receiver signal tracking method based on I/Q branch coherent integration measurements filter. J Electron Inf Technol 01:37–42

Ray-Tracing Aided Indoor Bayesian Positioning Algorithm

Hanhua Wang, Zhongliang Deng, Xiao Fu and Jing Li

Abstract Wireless signal based localization method has become the most widespread approach to determine receiver position in indoor and complex urban environment, due to the limited coverage of Global Navigation Satellite System signal. Bayesian estimation based Received Signal Strength (RSS) indoor positioning algorithm is applied in vast wireless signal localization systems. However, it suffers from burdensome workload on data collecting and building fingerprint database. To solve this problem, a ray-tracing aided Bayesian positioning algorithm is proposed in this paper. Unlike traditional Bayesian algorithm, the new algorithm does not need collecting fingerprint database but uses ray-tracing method to predict RSS probability distribution in space. Meanwhile, using RSS probability distribution as prior probability, positioning result could be achieved by calculating maximum posterior probability. Experiments proved that predicted RSS probability distribution model is very close to the actual statistical result. In addition, the algorithm reduces 99.7% time-consuming. As the cost of time, ray-tracing aided Bayesian algorithm only sacrifice 9.43% positioning accuracy.

Keywords RSS · Ray-tracing · Bayesian estimation · Bluetooth

1 Introduction

Global Navigation Satellite System (GNSS) could provide a reliable positioning accuracy in outdoor environment. However, in indoor and complex urban environment, because of blockage of buildings and humans, GNSS often suffers from multipath and signal attenuation.

Currently, wireless sensor network (WSN) based positioning technology is widely used in indoor environment. Yang et al. proposed a visible light

H. Wang (✉) · Z. Deng · X. Fu · J. Li
School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China
e-mail: whh0710@bupt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_51

communication and positioning system with positioning accuracy about 30 cm in 90%. However, the system need particular device and could not work in non-line-of-sight (NLOS) environments [1]. Kempke et al. proposed a UWB based positioning system with 40 cm positioning accuracy in line-of-sight (LOS) environment [2]. Zafari et al. presents an iBeacon based localization system with 0.947 m positioning accuracy [3]. Compared to other WSN based positioning technologies, Bluetooth based Bayesian positioning technology could provide an acceptable positioning accuracy in both NLOS and (LOS) environments with low node deployment cost. However, traditional Bayesian positioning algorithm has a shortcoming on waste of time and human resources for data collecting [4].

To reduce load of data collection, this paper proposes a ray-tracing aided Bayesian positioning algorithm. Unlike traditional Bayesian algorithm, the new algorithm could directly predict RSS probability distribution, approximated by Gaussian Mixture Models (GMM), without data collecting.

GMM has be proved applicable for approximating probability distribution of RSS [5]. Meanwhile, ray-tracing algorithm is a widely used algorithm for simulating indoor radio wave propagation and penetration [6]. Integrated dielectric constant of building material with building map information, ray-tracing method could calculate relatively precise radio wave propagation model [7]. Ray-tracing aided Bayesian positioning algorithm uses radio wave propagation model to predict parameters in GMMs at reference points (RP). Meanwhile, using RSS probability distribution as prior probability, positioning result could be achieved by calculating maximum posterior probability.

In experiment, we use Bluetooth access points (AP) as wireless signal sources. The predicted RSS probability distribution models are close to actual statistical result. Meanwhile, the positioning results show that the algorithm could provide a reliable positioning accuracy.

Ray-tracing aided Bayesian positioning algorithm are described in Sect. 2. Experiments and results are presented in Sect. 3. Section 4 shows conclusion and future work.

2 Ray-Tracing Aided Bayesian Positioning Algorithm

2.1 Ray-Tracing Algorithm

Ray-tracing algorithm is widely used in wireless signal propagation simulation. The algorithm could provide a relatively precise simulation result. Ray-tracing theory is based on that the waves from a transmitting antenna can be modeled as many ray tubes shooting from the antenna [6]. The ray-tracing algorithm uses geometric optics and mirror image theories to describe the reflection and refraction of radio waves. The ray-tracing algorithm, used in this paper, is a point-to-point method based on five steps as following.

1. Analysis relationships of transmitter, receiver and obstacles.
2. Use path loss model to calculate RSS of line-of-sight path.
3. Calculate attenuation rate and transmission coefficients based on permittivity of obstacle material, and then calculate RSS of transmission path.
4. Use mirror image theory to calculate mirror point of transmitter.
5. Calculate reflection coefficients and further achieve RSS of reflection path.

Path loss can be expressed as [8]

$$PL(d) = PL(d_0) + 10n \log\left(d/d_0\right) + X_\sigma \tag{1}$$

where d is distance between transmitter and receiver, d_0 = 1 m, n is a parameter, which equals to 2 in free space, X_σ is a zero mean Gaussian random variable with σ standard deviation that reflects the variation in average received power.

According to ITU-R Recommendation P.1238-6 [9], the attenuation rate, A, of an electromagnetic wave propagating through obstacles is given by:

$$\mathrm{A} = 1636 \frac{\sigma}{\varepsilon_r} \tag{2}$$

where σ is the conductivity of obstacle material, ε_r is the relative permittivity of material. Meanwhile, from the complex permittivity η, the reflection coefficient is given by:

$$R_N = \frac{\cos\theta - \sqrt{\eta - \sin^2\theta}}{\cos\theta + \sqrt{\eta - \sin^2\theta}} \tag{3}$$

$$R_P = \frac{\cos\theta - \sqrt{\frac{\left(\eta - \sin^2\theta\right)}{\eta^2}}}{\cos\theta + \sqrt{\frac{\eta - \sin^2\theta}{\eta^2}}} \tag{4}$$

where R_N is reflection coefficient when E-field component normal to reflection plane, R_P is reflection coefficient when E-field component parallel to reflection plane, θ is the angel between the incidence ray and the normal line of the reflecting surface. When incident E-field is circularly polarized, the reflection coefficient R_C could be calculated as

$$R_C = \frac{R_N + R_P}{2} \tag{5}$$

2.2 Gaussian Mixture Mode

RSS probability distribution could be approximated by Gaussian Mixture Model [10]. The Gaussian Mixture Model (GMM) is a probabilistic approach, which approximating a PDF by a mixture of Gaussian Models with different parameter [11]. In this paper, we use GMM to approximate RSS probability distribution in Bayesian algorithm. PDF of GMM is expressed as,

$$\mathrm{P}\left(\mathrm{y}|\theta_q\right) = \sum_{q=1}^{Q} \frac{\alpha_q}{\sqrt{2\pi}\sigma_q} \exp\left(-\frac{\left(y-\mu_q\right)^2}{2\sigma_q^2}\right) \tag{6}$$

$$\theta_q = \left(\mu_q, \sigma_q^2\right) \tag{7}$$

where μ_q, σ_q^2 and α_q are expectation, variance and weight of GMM's each Gaussian model.

2.3 Positioning Algorithm

2.3.1 Traditional Bayesian Positioning Algorithm

In traditional Bayesian positioning algorithm, there are two steps, learning step and localization step [6]. In learning step, a prior probability database should be build. The database stores RSS probability distribution, $\mathrm{P}\left(RSS_i|Point_j\right)$, at j reference points, $\left\{Point_1, Point_2, Point_3, \ldots, Point_j\right\}$, for i wireless signal access point, $\{RSS_1, RSS_2, RSS_3, \ldots, RSS_i\}$. $\mathrm{P}\left(RSS_i|Point_j\right)$ could be estimated by collected RSS data.

In traditional Bayesian algorithm's localization step, algorithm uses probability of RSS in fingerprint database as prior probability to calculate posterior probability, $\mathrm{P}(Point_i|RSS_i)$, by Bayesian estimation. Positioning result could be obtained by calculating maximum product of posterior probability. Using Bayesian estimation, in N access points, RSS of each access point could provide a probability of positioning result. Bayesian estimation could be expressed as

$$\mathrm{P}\left(Point_j|RSS_i\right) = \frac{\mathrm{P}\left(RSS_i|Point_j\right) \cdot P\left(Point_j\right)}{\sum_{j=1}^{N} \mathrm{P}\left(RSS_i|Point_j\right) \cdot P\left(Point_j\right)} \tag{8}$$

where $Point_j$ is the reference point j, RSS_i is the RSS of access point i, $P\left(Point_j\right)$ is the probability of reference point j, which is 1/N. Positioning result is calculated by

$$\arg\max \mathrm{P}\left(Point_j|\mathrm{RSS}\right) = \prod_{i=1}^{K} \mathrm{P}\left(Point_j|RSS_i\right) \tag{9}$$

where K is number of reference points. The reference point, which has the highest posterior probability, is the positioning result.

Plenty of RSS data should be collected for estimating RSS probability model. Therefore, the algorithm has a shortcoming of time and human resources cost.

2.3.2 Ray-Tracing Aided Bayesian Positioning Algorithm

In this paper, a ray-tracing aided Bayesian positioning algorithm is proposed for time-consuming problem. Ray-tracing aided Bayesian positioning algorithm does not has learning step and fingerprint database. The algorithm uses ray-tracing to predict RSS probability distribution function directly. RSS probability distribution function, $\mathrm{P}(RSS_i|(\mathrm{x},\mathrm{y}))$, could be predicted with less data collected. When environment information is precise, ray-tracing algorithm could provide an accurate RSS distribution. Therefore, in same environment, using trained path loss model, ray-tracing algorithm could directly calculate simulated RSS at any points, $RSS_{RT}(x,y)$. To present predicted RSS probability distribution function by GMM, $RSS_{RT}(x,y)$ could be used as an expectation for RSS probability distribution function. In general, human body shadowing could cause an attenuation about 6 dB on RSS for 900 MHz –100 GHz electromagnetic wave [9]. Therefore, considering influence of human body shadowing, attenuated RSS, $(RSS_{RT}(x,y)-6)$, could be used as another expectation. Path loss model's X_σ could provide variances. The whole RSS probability distribution function could be achieved by ray-tracing algorithm directly. The RSS probability distribution function could be expressed as

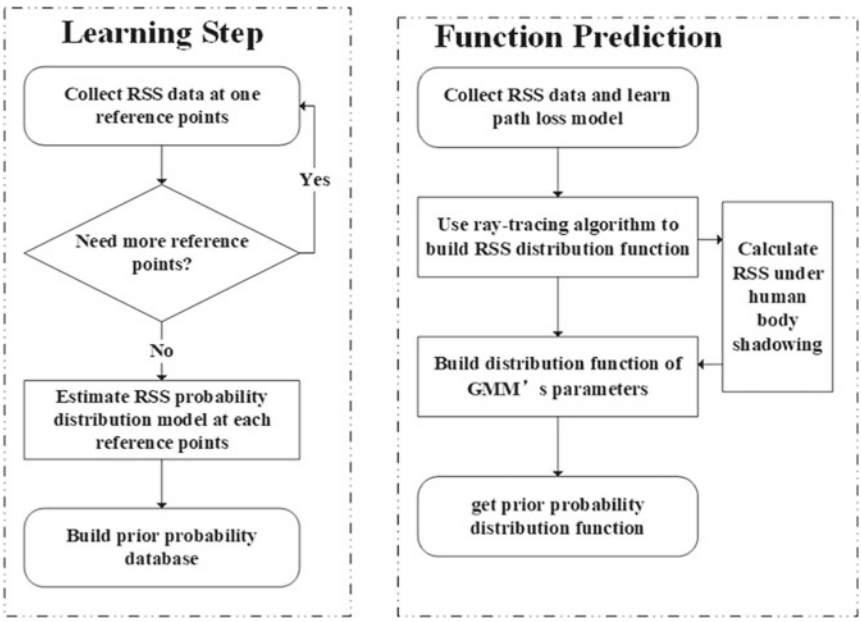
$$\begin{aligned}\mathrm{P}(RSS_i|(\mathrm{x},\mathrm{y})) = &\frac{\alpha}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\left(RSS_i - RSS_{RT}(x,y)\right)^2}{2\sigma^2}\right)\\ &+\frac{1-\alpha}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\left(RSS_i - (RSS_{RT}(x,y)-6)\right)^2}{2\sigma^2}\right)\end{aligned} \tag{10}$$

where $RSS_{RT}(x,y)$ is RSS calculated by ray-tracing, $(RSS_{RT}(x,y)-6)$ is RSS under human body shadowing, σ is variance of X_σ in path loss model.

The processing of RSS GMM prediction step is showed in Fig. 1b. Figure 1a shows processing of learning step in traditional algorithm.

Compared with traditional Bayesian positioning algorithm, ray-tracing aided algorithm does not need collect RSS data at every reference points. Therefore, the method could save a large amount of time and human resources.

Because ray-tracing aided Bayesian positioning algorithm doesn't has fingerprint database but RSS probability distribution function in space, positioning result could



(a) Learning step in traditional method (b) Predict step in ray-tracing aided method

Fig. 1 Flow chart of different method

get by calculating position, (x, y), where Eq. (11) has maximum value. Using RSS probability distribution as prior probability, positioning result could be achieved by calculating maximum posterior probability.

$$\arg\ \max \mathrm{P}((\mathrm{x},\mathrm{y})|\mathrm{RSS}) = \prod_{i=1}^{K} \mathrm{P}((\mathrm{x},\mathrm{y})|RSS_i) \tag{11}$$

$$\mathrm{P}((\mathrm{x},\mathrm{y})|RSS_i) = \frac{P(RSS_i|(x,y))}{\int_{x0}^{x1}\int_{y0}^{y1} P(RSS_i|(x,y))dxdy} \tag{12}$$

To simplify the solve processing, we could use logarithmic likelihood to handle Eq. (11) as

$$\ln\ \mathrm{P}((\mathrm{x},\mathrm{y})|\mathrm{RSS}) = \sum_{i=1}^{K} \mathrm{P}((\mathrm{x},\mathrm{y})|RSS_i) \tag{13}$$

3 Experiment and Result

3.1 Experiment Environment

Experiments are carried out in room 908 and 906, Research Building, Beijing University of Posts and Telecommunication. Map of experiment area is showed as Fig. 2. In experiment, receiver could receive signal from both LOS and NLOS channel. We use 360 beacon APs with Bluetooth 4.0 technology as signal source, which broadcast signal once per seconds. Huawei P9 is used as signal receiver.

3.2 Experiment Result

According to Recommendation ITU-R P.1238-6 [9], for 2.4 GHz radio wave, concrete material's complex permittivity, relative permittivity and conductivity are $7 - j0.85$, 5.31 and $0.066\,\mathrm{S/m}$ separately. For glass material, they are $6.76 - j0.09$, 6.27 and $0.023\,\mathrm{S/m}$.

We collected about 150 pieces of RSS data at 6 points separately in room 908 and 906. Using least square to fit those data with Eq. (1), we got $n = 1.7$ and $PL(1m) = -54\,\mathrm{dBm}$ in path loss model. Using maximum likelihood method to approximate error probability distribution of those data with zero mean Gaussian model, we got $\sigma = 3.12$ in equation.

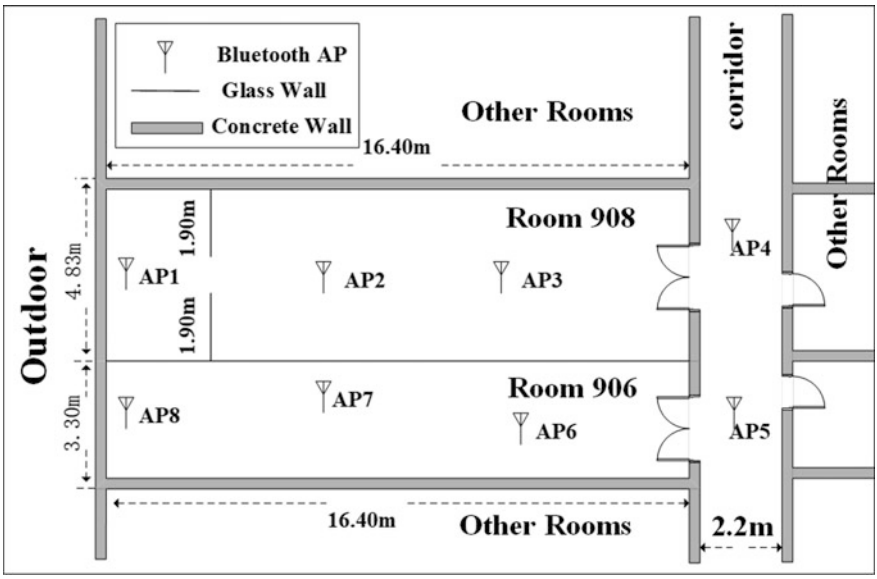


Fig. 2 Map of experiment site

Approximating accumulated 117 sets of RSS data in previous experiments with GMM, we got 522 couples of weights and expectations. After analyzing those data, we found out the average of max expectation model's weight is 0.787.

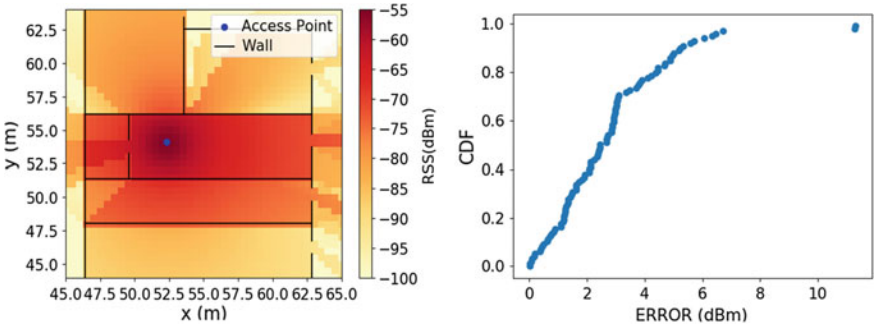
Integrated path loss model and parameters of material, we calculated RSS distribution in room 908 and 906 using ray-tracing algorithm. Figure 3a shows calculated result of AP2. We collected RSS data at 18 points. After using Maximum Expectation (ME) algorithm to approximate collected data with GMM, we got expectations of max weight Gaussian model in GMMs. Comparing predicted max RSSs and those expectations, when RSS > −80 dBm, 80% of ray-tracing predicted result has an error less than 4 dBm. CDF of difference between predicted max RSS and expectations with max weight is showed in Fig. 3b.

After getting expectations from ray-tracing and human body shadowing, weights from Eq. (11) and variances from path loss model, we predict RSS GMM at the 8 points. Part of CDFs are showed in Fig. 4.

We tested positioning algorithm at 16 points in room 908 and 906. Figure 5a shows distribution of $P((x, y)|RSS_{AP1} = -60\,\text{dBm})$. Figure 5b shows distribution of $P((x, y)|\text{RSS})$, where $\text{RSS} = \{RSS_{AP1} = -63\,\text{dBm}, RSS_{AP2} = -63\,\text{dBm}, RSS_{AP3} = -71\,\text{dBm}, RSS_{AP4} = -60\,\text{dBm}, RSS_{AP5} = -51\,\text{dBm}, RSS_{AP6} = -63\,\text{dBm}, RSS_{AP7} = -82\,\text{dBm}, RSS_{AP8} = -89\,\text{dBm}\}$.

We have tested the algorithm's performance of static positioning. We did test at 19 test points and repeated 10 times. In test, we collected 10 s RSS data at each test points. In NLOS environment, positioning result shows that 50% of positioning error are under 1.4157 m. We also tested the performance of traditional Bayesian positioning algorithm in same environment with 128 reference points, as Fig. 6.

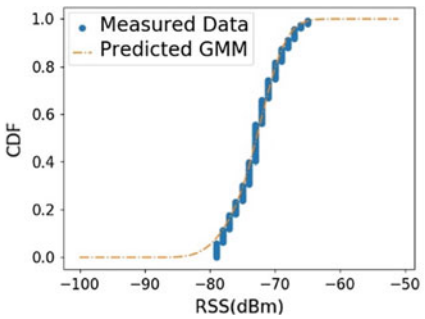
Compared with traditional Bayesian positioning algorithm, ray-tracing aided algorithm could improve efficiency by 99.25% (119.0954/120) in database building step and only lower positioning accuracy by 9.44% (0.1588/1.6829). Compared data is showed in Table 1.



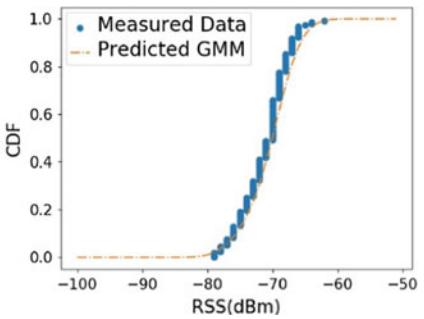
(a) Calculated max RSS distribution for AP2

(b) CDF of error

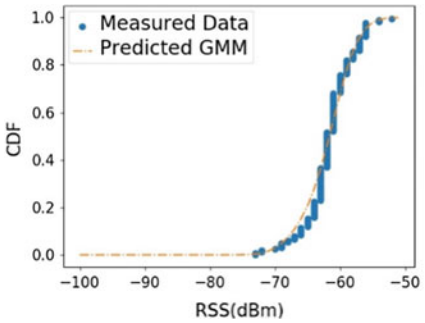
Fig. 3 Result of ray-tracing algorithm: **a** Ray-tracing calculated max RSS distribution of AP2, **b** CDF of error between ray-tracing calculated max RSS and real collected



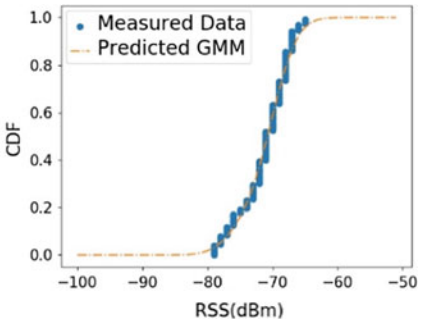
(a) AP4's RSS CDF at point (52.4,52.9)



(b) AP2's RSS CDF at point (59.9,55.2)

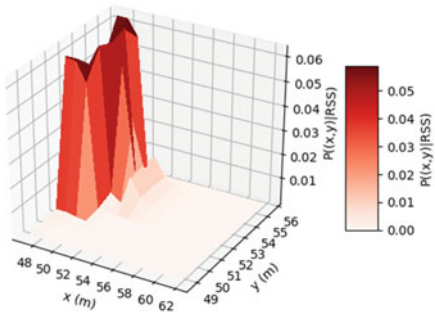


(c) AP3's RSS CDF at point (54.8, 52.9)

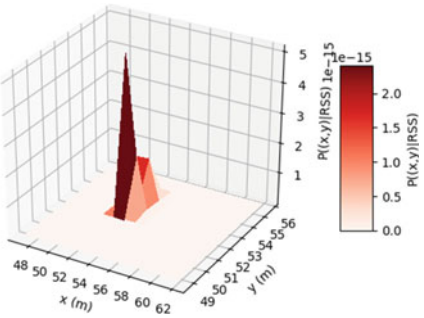


(d) AP5's RSS CDF at point (57.5, 52.9)

Fig. 4 CDF of collected data and predicted model at different points: **a** AP4's RSS CDF at point (52.4, 52.9), **b** AP2's RSS CDF at point (59.9, 55.2), **c** AP3's RSS CDF at point (54.8, 52.9), **d** AP5's RSS CDF at point (57.5, 52.9)



(a) $\mathrm{P}((\mathrm{x},\mathrm{y})|RSS_{AP1} = -60\ dBm)$ distribution



(b) P $((\mathrm{x},\mathrm{y})|\mathrm{RSS})$ distribution

Fig. 5 Posterior probability of Bayesian positioning algorithm: **a** $\mathrm{P}((\mathrm{x},\mathrm{y})|RSS_{AP1} = -60\,\mathrm{dBm})$ distribution, **b** $\mathrm{P}((\mathrm{x},\mathrm{y})|\mathrm{RSS})$ distribution

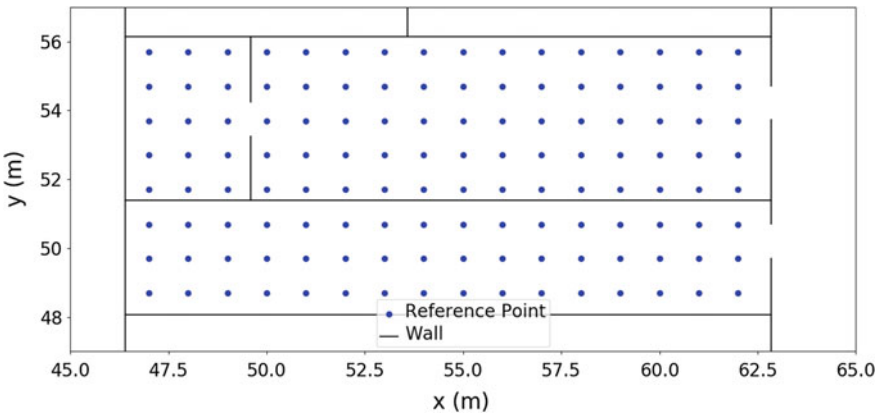


Fig. 6 Reference points localization

Table 1 Compared data between ray-tracing aided Bayesian algorithm and traditional algorithm

	Ray-tracing aided	Traditional
Mean (m)	1.8417	1.6829
Max (m)	3.8319	2.8771
50% (m)	1.4157	1.4010
Time cost per RP (s)	112.196/124 = 0.9048	120

4 Conclusion and Future Work

This paper proposes a ray-tracing based Bayesian positioning algorithm. Compared to traditional algorithm, this algorithm does not need to collect RSS data at reference points before localization step. This feature enables the algorithm that can save much time and human resources. In addition, the algorithm provides a positioning accuracy near to traditional Bayesian positioning algorithm.

There are still some further studies could be do in future. To improve the accuracy of predicted model, other probability distribution models could be studied and tested for approximating RSS probability distribution. Nakagami-m model is widely used to approximating amplitude probability distribution for communication channel model [12]. A positioning result filter could be used to solve instability of positioning result and improves the positioning accuracy [13]. Meanwhile, performance of the algorithm with other signal sources could be tested.

Acknowledgements This work is financially supported by National Key R&D Program of China (No. 2016YFB0502001 and No. 2016YFB0502003).

References

1. Yang Z, Wang Z, Zhang J, Huang C, Zhang Q (2015) Wearables can afford: light-weight indoor positioning with visible light. In: Proceedings of the 13th annual international conference on mobile systems, applications, and services, ACM, pp 317–330
2. Kempke B, Pannuto P, Dutta P (2015) PolyPoint: guiding indoor quadrotors with ultra-wideband localization. In: Proceedings of the 2nd international workshop on hot topics in wireless, ACM, pp 16–20
3. Zafari F, Papapanagiotou I, Devetsikiotis M, Hacker T (2017) An iBeacon based proximity and indoor localization system. arXiv preprint arXiv:1703.07876
4. Liu H, Darabi H, Banerjee P, Liu J (2007) Survey of wireless indoor positioning techniques and systems. IEEE Trans Syst Man Cybern Part C Appl Rev 37(6):1067–1080
5. Sheng X, Hu YH, Ramanathan P (2005) Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network. In: IEEE fourth international symposium on information processing in sensor networks, 2005. IPSN 2005, pp 181–188
6. Kim H, Ling H (1992) Electromagnetic scattering from an inhomogeneous object by ray tracing. IEEE Trans Antennas Propag 40(5):517–525
7. Glassner AS (ed) (1989) An introduction to ray tracing. Elsevier, The Netherlands
8. Andersen JB, Rappaport TS, Yoshida S (1995) Propagation measurements and models for wireless communications channels. IEEE Commun Mag 33(1):42–49
9. Recommendation ITU-R P.1238-6 (2009) Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900 MHz to 100 GHz. Available online: https://www.itu.int/rec/R-REC-P.1238-6-200910-S/en
10. Alfakih M, Keche M, Benoudnine H (2015) Gaussian mixture modeling for indoor positioning wifi systems. In: IEEE 2015 3rd international conference on control, engineering & information technology (CEIT), pp 1–5
11. Koshizen T, Bartlett P, Zelinsky A (1999) Sensor fusion of odometry and sonar sensors by the Gaussian mixture Bayes' technique in mobile robot position estimation. In: 1999 IEEE international conference on systems, man, and cybernetics, IEEE SMC '99 conference proceedings, vol 4, pp 742–747
12. Nakagami M (1960) The m-distribution—a general formula of intensity distribution of rapid fading. In Statistical methods in radio wave propagation, pp 3–36
13. Gustafsson F, Gunnarsson F, Bergman N, Forssell U, Jansson J, Karlsson R, Nordlund PJ (2002) Particle filters for positioning, navigation, and tracking. IEEE Trans Signal Process 50 (2):425–437
14. Madigan D, Einahrawy E, Martin RP, Ju WH, Krishnan P, Krishnakumar AS (2005) Bayesian indoor positioning systems. In: INFOCOM 2005, 24th annual joint conference of the IEEE computer and communications societies, Proceedings IEEE, vol 2, pp 1217–1227

Sensor Synchronization for Android Phone Tightly-Coupled Visual-Inertial SLAM

Zheyu Feng, Jianwen Li and Taogao Dai

Abstract At present, the majority of Android phones only support satellite positioning and cellular localization. Both of them are of poor indoor performance, which limits the development of the relevant indoor location based services. In this paper, we attempt to achieve positioning with raw image and Inertial Measurement Unit (IMU) data from Android phone. We first introduce a state-of-the-art framework for tightly-coupled monocular visual-inertial Simultaneous Localization and Mapping (SLAM) using image and IMU data. Then we focus on the unsynchronization problem between camera and IMU of Android phone, and propose a grid search algorithm based on spherical quaternion interpolation for delay estimation. The results of indoor and outdoor experiments show that the algorithm can estimate the delay of image timestamp effectively, and the percentage of positioning plane error is 0.79% indoors and 8.09% outdoors respectively.

Keywords Android phone · Tightly-coupled · Synchronization
SLAM · VINS

1 Introduction

With the popularization of smartphones and the increasing demand for location information, location based services, especially on smart phone, have become an indispensable part of people's daily life. The foundation of location based service is positioning. Although Google has launched Tango project in June 2014 to achieve

Z. Feng · J. Li (✉)
Information Engineering University, Zhengzhou 450001, China
e-mail: zzljw@126.com

Z. Feng
e-mail: von9604@gmail.com

T. Dai
63883 Troops, Luoyang 471000, China
e-mail: 18538881619@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_52

visual navigation using infrared sensor and image sensor, the main positioning methods for Android phone are still satellite positioning and cellular localization. Satellite positioning is of high accuracy with poor anti-interference capability, and cannot be used indoors. Meanwhile, although cellular localization is available indoors, the positioning accuracy can only reach several hundred meters [1], which cannot meet the requirement of indoor navigation. It's obvious that the technology bottleneck of positioning is one of the main factors that limit the development of indoor location based services.

SLAM, which aims to achieve localization in unknown environment, is one of the research hotspots of autonomous navigation for mobile robot at present. Visual SLAM is attracting more and more attention. Because image sensor is usually of wide detection range, and visual image has large amount of information and rich features easy to extract [2]. Whereas, all feature-based visual SLAM solutions do not have robust dynamic performance. They are easy to lose tracking with severe motion.

IMU has a high sampling frequency and is highly sensitive to the motion of the carrier. Additionally, accurate motion measured from IMU can be used for monocular SLAM to recover scale. Therefore, the Integration of a monocular camera and an IMU becomes a popular low-cost SLAM solution. A robust and versatile monocular visual-inertial state estimator based on sliding window graph-based optimization framework called VINS is proposed in [3], which has been well applied in Apple's smart devices [4]. Thus we attempt to achieve positioning with this framework with data from Android phone.

The visual-inertial System sensors usually synchronize directly on hardware level, so sensors are assumed synchronized before data processing [3–6]. However, at present, the image and the IMU are not synchronized in Android system, which will result in failure for the visual-inertial SLAM. To our best knowledge, the main research about sensor synchronization focuses on correcting rolling shutter effect [7] of an image with gyroscope data to achieve image stabilization [8, 9] or more accurate feature extraction [10]. In [11–13], synchronization between camera and IMU is considered in the visual-inertial SLAM and online calibration of temporal parameter is proposed. However, the algorithm is hard to implement for the VINS graph-based optimization framework. Therefore, in this paper, aiming at solving the unsynchronization problem, a simple algorithm for image delay estimation is proposed. Experimental results show that, the framework proposed in [3] can be improved for Android phone with this algorithm.

2 VINS Framework

In this section, we will briefly introduce the VINS framework applied [3]. The framework can be divided into four parts: sensor data pre-processing, system initialization, nonlinear optimization and loop closure. The overall framework is shown in Fig. 1.

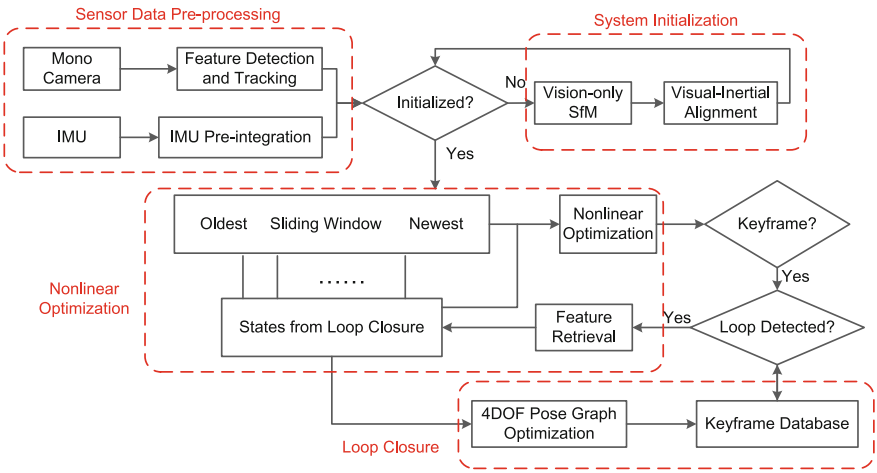


Fig. 1 VINS overall framework

The specific process of the framework is as follows:

(1) **Sensor Data Pre-processing** For the camera output images, features are tracked by KLT sparse optical flow algorithm. Simultaneously, new feature points are detected to ensure that the number of features of each image is fixed. After that, RANSAC fundamental matrix test is applied to remove outliers. In this process, the keyframes are also selected. If the average parallax of a frame is higher than the set threshold or the number of feature points tracked is below the set threshold, the frame is regarded as a keyframe. On the other hand, the gyroscope and accelerometer measurements are pre-integrated. The image feature points and IMU pre-integrals obtained in this step are saved in a sliding window for subsequent system initialization and non-linear optimization.

(2) **System Initialization** First of all, by setting IMU average acceleration threshold, the system ensures the observability of the scale. Then, the structure of the visual frames is restored by triangulating the image feature points in the sliding window. The relative rotations and the up-to-scale translations between adjacent image frames are obtained. Thus initial values for system scale, gravity, position, velocity, attitude, and bias can be yielded from these rotations and translations with the pre-integrals of IMU measurements. After that, the initial gravity value is optimized to the standard value.

(3) **Nonlinear Optimization** After initialization, velocity, position, attitude, IMU-camera extrinsic parameters, bias, and feature point depths are nonlinearly optimized in the tightly-coupled state estimator. After optimization, the states in the sliding window are required to be added and deleted. If the second frame in the sliding window is a keyframe, it will be retained and the last frame will be deleted; if the second frame in the sliding window is not a keyframe, its visual

observations will be removed while its IMU measurements will be retained. The states are marginalized out to preserve a priori information when they are deleted. If the number of feature points tracked from the latest frame is less than 5 or the two consecutive state values differ greatly, the system considers fault happened and reinitializes, with previously recorded keyframes kept.

(4) **Loop Closure** FAST feature points are extracted and identified with BRIEF descriptor. Then these points are classified with bag-of-words technique. When a loop closure is recognized for a new image frame, the states in the sliding window are optimized together with the feature points of the looped keyframe. And the relative pose between the new image frame and the looped frame. Since the yaws and coordinates (x, y, z) of the entire system contain accumulated drifts, it is also necessary to finally optimize all poses in 4 degrees of freedom to obtain accurate states in the sliding window.

In summary, the VINS Framework algorithm incorporates the latest research achievements in the field of monocular visual-inertial SLAM. It has excellent implementation on measurement data pre-processing, system initialization, nonlinear optimization and loop-closure detection. However, because the monocular camera and IMU in Android phone are not synchronized, the visual-inertial SLAM cannot be applied directly for Android phone.

3 Delay Estimation

In the Android system, sensor data is obtained through the system-provided callback function. Camera images are acquired using the function onPreviewFrame (byte [] data, Camera camera) call, while IMU data can be obtained with the function onSensorChanged (SensorEvent event). SensorEvent contains sensor data and accurate measurement time [14, 15]. However, the callback function for camera image does not provide a timestamp. As a result, the time when the function is called is taken as image taken time for timestamping. Thus, there is a delay t_d between the timestamp t' of the image and the image taken time t, namely $t' = t + t_d$. The delay is the time interval between the moment of taking an image and the moment when the image is passed to the callback function. In addition, the delay is related to the specific system operation situation (Figs. 2 and 3).

3.1 System Configuration

In order to make this delay as constant as possible, the system is required to do the following settings [16].

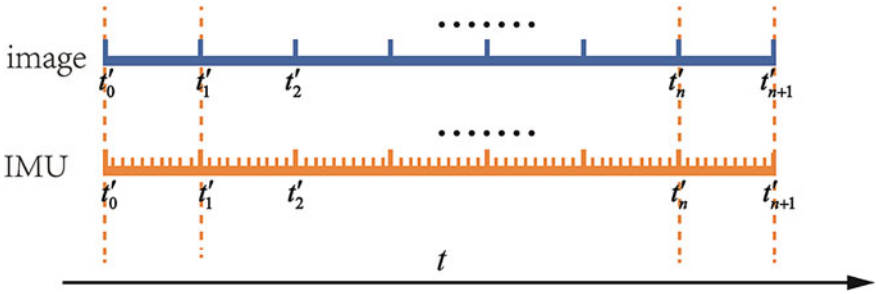


Fig. 2 Timestamps when camera and IMU are synchronized $t_d = 0$

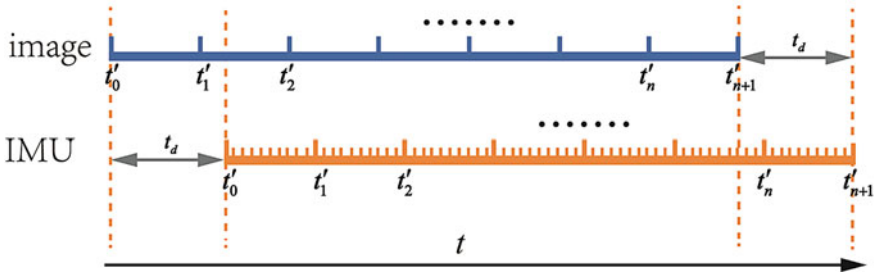


Fig. 3 Timestamps when camera image timestamps contain a delay $t_d > 0$

- Disable the camera auto exposure, lock the exposure time;
- Put the IMU data and camera image callback function into a separate thread respectively, to avoid the user interface (UI) thread blocking leading to delay instability;
- Avoid too many operations in the camera image callback function, to ensure that the callback function has been completed before called next time.

3.2 *Algorithm for Delay Estimation*

Assuming the IMU attitude change between t_i and t_{i+1} is $\gamma_{b_{i+1}}^{b_i} = \gamma_{b_{i+1}}^{b_i}(t_i, t_{i+1})$, and the camera attitude change is $\gamma_{c_{i+1}}^{c_i} = \gamma_{c_{i+1}}^{c_i}(t_i, t_{i+1})$, assuming the extrinsic rotation between the IMU and the camera q_c^b is known, we can get the following relationship.

$$\begin{aligned}\gamma_{b_{i+1}}^{b_i} &= q_{b_i}^{w^{-1}} \otimes q_{b_{i+1}}^{w} \\ &= (q_{c_i}^{w} \otimes q_b^c)^{-1} \otimes q_{c_{i+1}}^{w} \otimes q_b^c \\ &= q_c^b \otimes \gamma_{c_{i+1}}^{c_i} \otimes q_b^c\end{aligned} \tag{1}$$

Then we have

$$\gamma_{b_{i+1}}^{b_i} \otimes (q_c^b \otimes \gamma_{c_{i+1}}^{c_i} \otimes q_b^c)^{-1} = \mathbf{0} \tag{2}$$

where $\otimes$ is quaternion multiplication, $(\cdot)^{-1}$ is quaternion inversion. Select two frames with adjacent timestamps t_i' and t_{i+1}' respectively. And a fundamental matrix is calculated with the feature pairs are matched for calculation. Then fundamental matrix is decomposed to obtain the camera's attitude change.

Due to the delay of image timestamp, the camera attitude change $\hat{\gamma}_{c_{i+1}}^{c_i}$ form t_i to t_{i+1} is actually obtained. And the IMU angular velocity measurements from t_i' to t_{i+1}' are pre-integrated [17] to get the IMU attitude change $\hat{\gamma}_{b_{i+1}'}^{b_i'}$.

In order to obtain the delay, we first use a sliding window of size n to store the IMU attitude changes and the camera attitude changes. Then the IMU attitude changes are taken as reference. And the camera attitude changes at each aligned time are obtained by spherical interpolation [18] when the delay parameter changes. Thus residuals can be calculated, and we can find the delay that results in the minimum residual.

$$e = \sum_{i=1}^{k} \frac{\hat{\gamma}_{b_{i+1}'}^{b_i'} \otimes (q_c^b \otimes \hat{\gamma}_{c_{i+1}'}^{c_i'} \otimes q_b^c)^{-1}}{k} \tag{3}$$

In fact, in the window, the number of the camera attitude changes could vary with the different delay parameter. In order to ensure that the number of the camera attitude changes in the window is constant, we need to remove the non-public parts of the referenced IMU attitude changes according to the search range. The specific algorithm is shown in Fig. 4.

We determined the search range as [−0.5s, 0.5s], and found that the result was stable when the window size $n = 50$. The variation of residual with delay is shown in Fig. 5.

According to the analysis above, the image timestamps in Android phone must be greater than the image taken time. Thus the delay must be above 0, and only the delay greater than 0 need to be considered. It can be seen that the residual has a minimum value within the search range. We add this algorithm into the VINS system initialization part to ensure sensor synchronization for subsequent procedures.

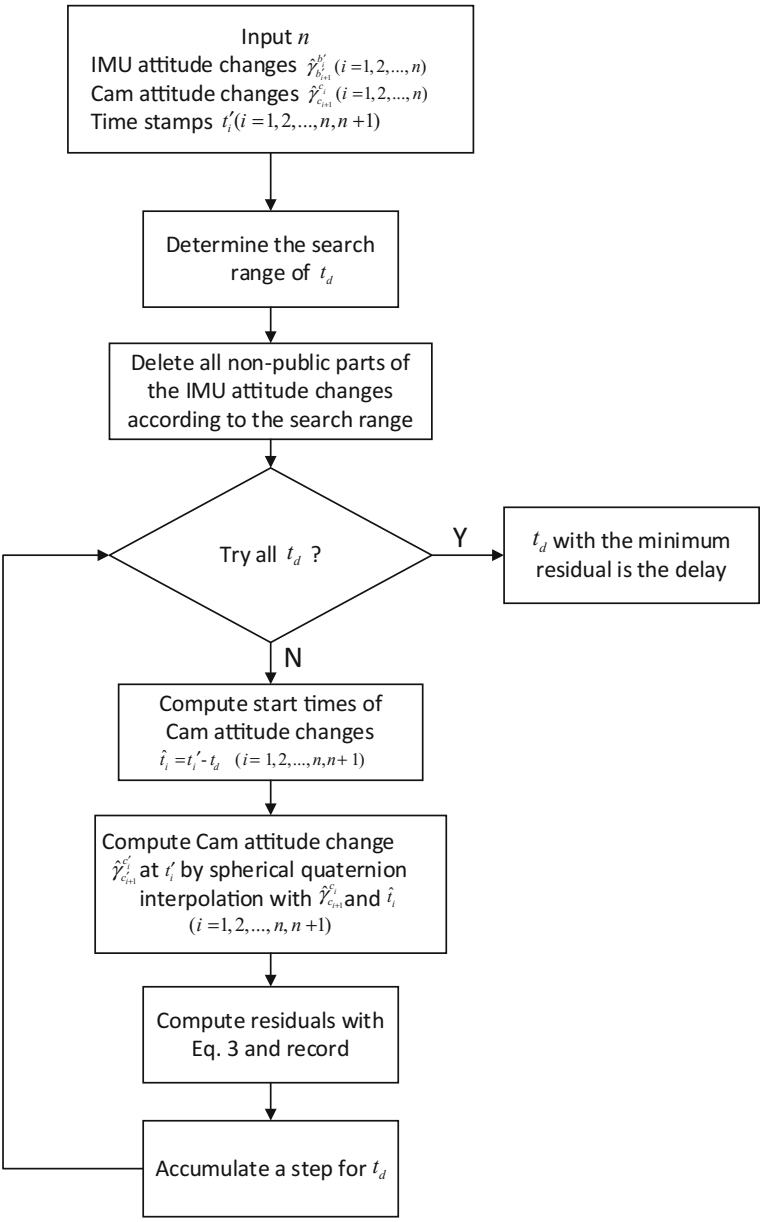


Fig. 4 The algorithm searches in the determined range of t_d, and obtains the camera attitude changes at each aligned time by spherical interpolation. It calculates the residual between the attitude changes of the camera and the IMU in the window. The delay with the minimum residual is the delay between image timestamp and image taken time

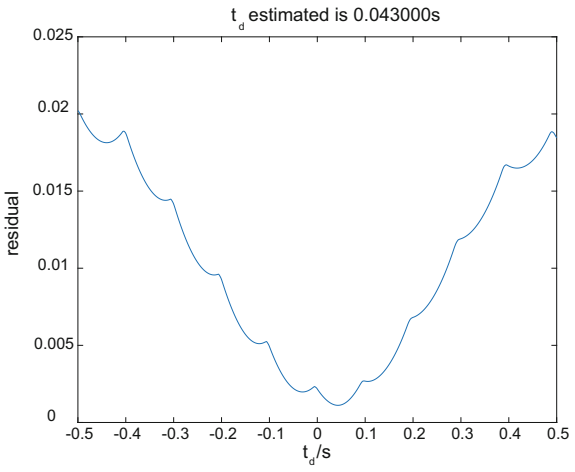


Fig. 5 Variation of residual with delay. It's apparent residuals are approximately symmetrical about the estimated delay

4 Experiments and Analysis

In order to verify the validity of the algorithm, we use a Huawei P9 mobile phone to obtain images and IMU data indoors and outdoors respectively. The improved VINS algorithm is used to locate the phone with loop closure detection disabled. During the experiments, mobile phone image sampling rate is set to 20 Hz, IMU sampling rate is 100 Hz. Experimenter holds Android mobile phone on hand and walk along a closed experimental line with the speed of about 1.5 m/s. The start and end point of the closed experimental lines are coincident. The accuracy is verified by calculating the coordinate difference between the start and end point.

4.1 Indoor Experiment

Indoor experiment is performed in an indoor library, the experimental route is about a square. In the process of system initialization, the variation of the residual with different delay is shown in Fig. 6, so we can get the delay of image timestamp $t_d = 0.0240$ s.

Positioning results after synchronization is shown in Fig. 7. The blue points in Fig. 7a are the map points constructed by the visual-inertial SLAM, and the black line is the trajectory of the Android phone. The orange points are the map point constructed with current visual frame. The final positioning plane accuracy is shown in Table 1.

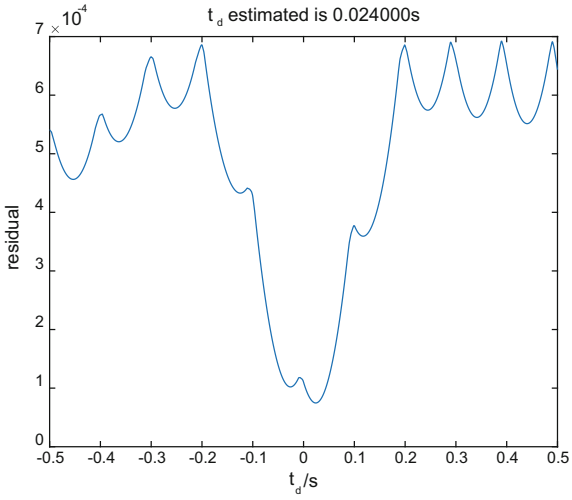


Fig. 6 Variation of residual with different delay in the indoor experiment

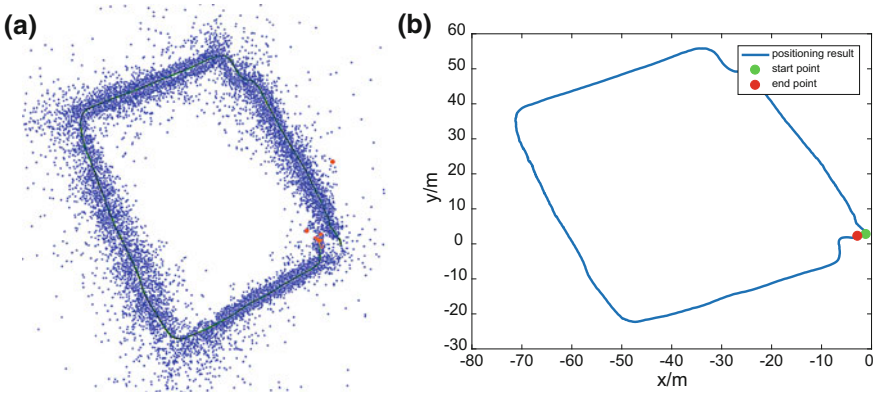


Fig. 7 Positioning results of the indoor experiment. **a** Carrier trajectory and map points during data processing. **b** Final positioning result

Table 1 Plane accuracy of indoor and outdoor experiments

	Duration (s)	Distance (m)	Plane accuracy (m)	Error percentage (%)
Indoor	164.216	230.704	1.829	0.79
Outdoor	271.594	400.000	32.366	8.09

4.2 Outdoor Experiment

Outdoor experiment is performed in an athletic field, the experimental route is the runway there. In the process of system initialization, the variation of residual with different delay is shown in Fig. 8, we can get the delay of image timestamp $t_d = 0.0162$ s. Positioning results after synchronization is shown in Fig. 9. The final positioning plane accuracy is shown in Table 1.

4.3 Result Analysis

It can be seen from above results that the delay search algorithm can effectively estimate the delay of image timestamp so that the VINS algorithm can be carried out smoothly. The indoor positioning accuracy is much better than the outdoor positioning accuracy, because in the large-scale outdoor environment, most of map points are far away from the camera. In this case, the same pixel error of the image can result in significantly larger triangulation error of map points, so the accuracy of poses is much worse.

5 Summary

In this paper, aiming to solve the unsynchronization problem between Android camera and IMU, an algorithm for delay estimation is proposed. With this algorithm, the VINS algorithm could be applied to the Android raw data successfully.

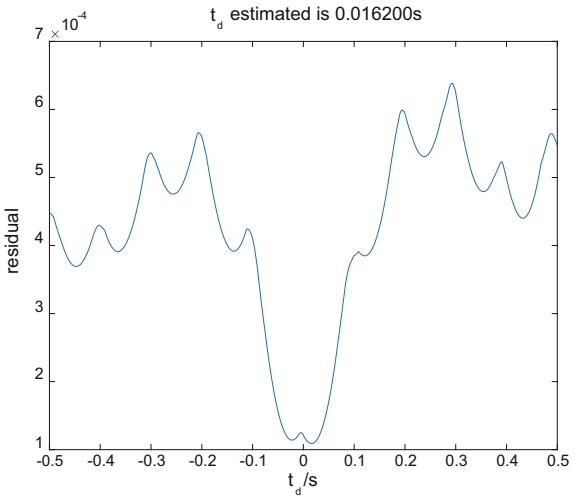


Fig. 8 Variation of residual with different delay in outdoor experiment

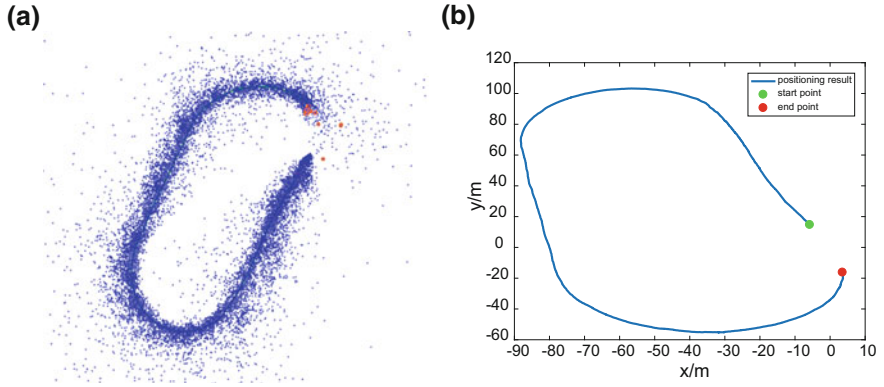


Fig. 9 Positioning results of outdoor experiment. **a** Carrier trajectory and map points during data processing. **b** Final positioning result

The experimental results show that the positioning accuracy indoor and outdoor after the synchronization is 0.79 and 8.09%, respectively. The validity of the synchronization algorithm is verified, and the good performance of the VINS algorithm for indoor localization is also proved. It is foreseeable that in the near future, visual-inertial SLAM will be an effective means for Android users to achieve indoor positioning and navigation.

References

1. Xu Y (2016) The research and implementation of positioning technology based on base station of mobile terminal. Dalian Maritime University, China
2. Quan M, Piao S, Li G (2016) An overview of visual SLAM. CAAI Trans Intell Syst 11 (06):768–776
3. Qin T, Li P, Shen S (2017) VINS-mono: a robust and versatile monocular visual-inertial state estimator. arXiv preprint arXiv:1708,03852
4. Li P, Qin T, Botao H et al (2017) Monocular visual-inertial state estimation for mobile augmented reality. In: Proceedings of the IEEE international symposium on mixed and augmented reality, Nantes, France
5. Corke P, Lobo J, Dias J (2007) An introduction to inertial and visual sensin. Int J Rob Res 26 (6):519–535
6. Kelly J, Sukhatme G (2011) Visual-inertial sensor fusion: localization, mapping and sensor-to-sensor self-calibration. Int J Rob Res 30(1):56–79
7. Geyer C, Meingast M, Sastry S (2005) Geometric models of rolling shutter cameras. In: Proceedings workshop omnidirectional vision
8. Karpenko A, Jacobs D, Baek J, Levoy M (2011) Digital video stabilization and rolling shutter correction using gyroscopes. Technical Report, Stanford University
9. Hanning G, Forslöw N, Forssén PE, Ringaby E, Törnqvist D, Callmer J (2011) Stabilizing cell phone video using inertial measurement sensors. In: Proceedings of IEEE international workshop on mobile vision

10. Hwangbo M, Kim J-S, Kanade T (2011) Gyro-aided feature tracking for a moving camera: fusion, auto-calibration and GPU implementation. Int J Rob Res 30(14):1755–1774
11. Jia C, Evans BL (2012) Probabilistic 3-D motion estimation for rolling shutter video rectification from visual and inertial measurements. In: Proceedings of IEEE international workshop multimedia signal processing
12. Li M, Kim B, Mourikis A (2013) Real-time motion tracking on a cellphone using inertial sensing and a rolling shutter camera. In: Proceedings of IEEE international robotics and automation
13. Li M, Mourikis A (2013) 3-D motion estimation and online temporal calibration for camera-IMU systems. In: Proceedings of IEEE international robotics and automation
14. Google. Camera API. https://developer.android.com/
15. Google. Motion sensors. https://developer.android.com/
16. Latimer R, Holloway J, Veeraraghavan A et al (2014) SocialSync: sub-frame synchronization in a smartphone camera network. Computer vision—ECCV 2014 workshops. Springer International Publishing, Basel, pp 561–575
17. Forster C, Carlone L, Dellaert F et al (2015) IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation. Robotics: science and systems
18. Shoemake K (1985) Animating rotation with quaternion curves. In: International conference on CGIT, pp 245–254

A Low-Cost GNSS/INS Attitude and Position Determination System with Ultra-Short Baseline for SATCOM On-The-Move Terminals

Peirong Fan, Wenyi Li, Xiaowei Cui and Mingquan Lu

Abstract In SATCOM On-The-Move (SOTM) systems, an accurate pointing from moving antennas towards the communication satellite is the key to a stable signal tracking. To realize a reliable SOTM operation under GNSS-challenged environments with limited size, weight and power (SWaP), a low-cost two-antenna GNSS/INS attitude and position determination system is proposed in this paper, with ultra-short baseline ranging from 0.2 to 0.5 m, and consumer-grade navigation components. Specifically, INS dead reckoning aiding is introduced into the GNSS single-epoch baseline search, reducing the original spherical traversal to a one-dimension searching. A new ambiguity resolution acceptance test method is also developed to increase the success rate of ambiguity resolution. In the process of sensor fusion, an Innovation Pre-filtered Extended Kalman Filter (IPEKF) is proposed to deal with GNSS outliers resulting from noisy observations and multipath interference. At last, static and kinematic experiments were performed on a prototype SOTM terminal with 0.267 m baseline. Corresponding results verified a heading estimation accuracy of 0.3°, and in GNSS outage situations, the system was able to output usable attitude determination results within approximately 40 s. Applications of the proposed system could be further extended to attitude estimations of unmanned aerial vehicles (UAV) and ocean-going ships.

Keywords SATCOM On-The-Move · Ultra-short baseline · Two-antenna GNSS attitude determination · Ambiguity resolution · GNSS/INS integrated navigation Innovation pre-filtering

P. Fan · W. Li · X. Cui (✉) · M. Lu
Department of Electronic Engineering, Tsinghua University,
Beijing 100084, China
e-mail: cxw2005@mail.tsinghua.edu.cn

P. Fan
Beijing Satellite Navigation Center, Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_53

1 Introduction

In recent years, burgeoning private and military space-based satellite constellations have brought great attention on the Satellite Communication On-the-move (SOTM) technologies. With its ability to support near real-time communication [1] and high speed OTM data transferring, SOTM have brought great benefits for both military and civil users.

An accurate pointing from moving SOTM antennas towards the communication satellite is one of the most critical technical challenges confronted by researchers. Within limited size, weight and power (SWaP), ground communications are required in harsh environments under leaf cover, through tunnels or in steep terrains, driving rigorous pointing requirements. Moreover, for Ka-band SOTM vehicles, which features a higher data throughput compared to other frequencies, yet is much harder for moving terminals to pick up [2], such requirements become more demanding.

To improve the pointing accuracy with restricted SWaP, much attention has been paid on the fusion of low-cost sensors, including Micro-Electro-Mechanical System (MEMS) based inertial navigation systems (INS), and consumer-grade Global Navigation Satellite System (GNSS) receivers. In principle, inertial sensors estimate SOTM terminal states by integrating angular velocities and accelerometers measured by its inertial measurement units (IMU), which are accurate in short terms but suffers from long-term drifting errors. By contrast, multi-antenna GNSS receivers are drift free, yet prone to be affected by signal blockage and noisy observations. As a result, integration of the two systems becomes a well-known strategy to reduce the impact of GNSS measurement errors and to estimate INS states which are not directly-observable.

Much work has been carried out in integrating INS and GNSS receivers. In [3], GNSS-measured velocities and the sideslip angle are combined to correct SOTM terminal accelerations, so that faulty observations of GNSS receivers are detected and eliminated. In [4], a general sensor fusion framework is proposed under the strategy of direct collocation, therefore the number of iterations is decreased. Unfortunately, these methods, while useful, still can not satisfy the ever-increasing requirements of SOTM for both high robustness and limited budget.

In this paper, a low-cost attitude and position determination system for SOTM terminals is proposed with consumer grade GNSS receivers (navigation type) and MEMS INS module. Firstly, an INS aided ambiguity resolution algorithm is examined, and this will be followed by an Innovation Pre-filtered Extended Kalman Filter (IPEKF). The next section will present experimental results of a prototype SOTM terminal. At last, conclusions are drawn in Sect. 5.

2 INS-Aided GNSS Attitude Determination

2.1 GNSS Measurement Models

Typically, GNSS antennas of the attitude and position determination system are rigidly mounted on the SOTM vehicle. Therefore, attitudes of the terminal can be equivalently represented by the relative position between the base and rover antennas, namely the baseline vector.

Model the GNSS carrier phase measurements as following [5]:

$$\Phi_r^i = \frac{1}{\lambda}\left(\rho_r^i + cdt_r - cdt^i + T_r^i - I_r^i + E_r^i\right) - N_r^i + MP_r^i + \eta_r^i \tag{1}$$

where Φ_r^i is the carrier phase (in carrier cycle) between satellite i and rover r, λ is the carrier wavelength, ρ_r^i is the real geometric range, c is the speed of light, E_r^i is the ephemeris error, dt_r and dt^i are the clock biases of the receiver r and satellite i, respectively, I_r^i and T_r^i are the ionospheric and tropospheric delay, respectively, N_r^i is the integer ambiguity, MP_r^i is carrier multipath errors, and η_r^i presents all other phase tracking errors.

Firstly, to eliminate the clock bias of satellites in Eq. (1), measurements between rover and base receivers are differenced, therefore yields the single differenced (SD) carrier phase $\Delta\Phi_{rb}^i$. Due to the ultra-short baseline constraint in our system, common-mode errors, including I_r^i, T_r^i and E_r^i are approximately equal between both receivers, thus can also be removed from the SD expression. Secondly, to eliminate the clock bias of receivers, we subtract the SD carrier phase of satellite i from all other satellites, and get the following double differenced (DD) measurement:

$$\nabla\Delta\Phi_{rb}^{ij} = \frac{1}{\lambda}\nabla\Delta\rho_{rb}^{ij} - \nabla\Delta N_{rb}^{ij} + \nabla\Delta MP_{rb}^{ij} + \nabla\Delta\eta_{rb}^{ij} \tag{2}$$

where $\Delta(\cdot)_{rb}^i = \left((\cdot)_r^i - (\cdot)_b^i\right)$ and $\nabla\Delta(\cdot)_{rb}^{ij} = \left(\Delta(\cdot)_{rb}^i - \Delta(\cdot)_{rb}^j\right)$ represent the carrier SD and DD calculations, respectively.

In GNSS-friendly environments, the multi-path error $\nabla\Delta MP_{rb}^{ij}$ is small and can further be removed from the DD model.

2.2 Integer Ambiguity Resolution

In Eq. (2), the baseline vector of interest is buried in the DD range $\nabla\Delta\rho_{rb}^{ij}$ and the term $\nabla\Delta\Phi_{rb}^{ij}$ is obtained from GNSS observations, therefore, the key question of solving the equation becomes identification of the DD integer ambiguity $\nabla\Delta N_{rb}^{ij}$.

A commonly adopted method for the ambiguity resolution (AR) problem is the LAMBDA algorithm in ambiguity domain, in which whether the ambiguity search can succeed or not greatly depends on the precision of GNSS code pseudorange measurements [6]. However, for most SOTM terminals, consumer grade receivers are employed and usually only single frequency measurements are available, resulting in much larger measurement noises in the code pseudorange.

Conversely, an angle domain AR algorithm is described in [6], which abandons the noisy GNSS code pseudorange measurements and searches for the true baseline vector with a two-dimension traversal. The ascendancy of this method over the above ones in ambiguity domain was a significantly higher AR success rate for terminals with ultra-short baseline (ranging from 0.2 to 0.5 m) and low-cost single frequency GNSS receivers.

Main steps of the method is outlined as following:

Firstly, assuming the multi-path errors and other noises in Eq. (2) are negligible, the equation can be rearranged as

$$\frac{1}{\lambda}\nabla\Delta\rho_{rb}^{ij} = \nabla\Delta\tilde{\Phi}_{rb}^{ij} + \nabla\Delta\bar{N}_{rb}^{ij} \tag{3}$$

where $\nabla\Delta\tilde{\Phi}_{rb}^{ij}$ is the fractional part of the DD carrier phase ranging from −0.5 to 0.5 cycle, $\nabla\Delta\bar{N}_{rb}^{ij}$ is the corresponding DD ambiguity. The term $\nabla\Delta\tilde{\Phi}_{rb}^{ij}$ can be calculated from GNSS observations.

Secondly, generate a series of candidate baseline vectors with a two dimensional traversal along pitch and heading angles of the system. Typically, these vectors are expressed in spherical coordinates, as shown in Eq. (4).

$$\mathbf{r}_{rb,mn} = \begin{bmatrix} l\sin(\varphi_m)\cos(\theta_n) \\ l\cos(\varphi_m)\cos(\theta_n) \\ l\sin(\theta_n) \end{bmatrix} \tag{4}$$

where $\mathbf{r}_{rb,mn}$ is the candidate baseline vector, l is the priori baseline length, θ_n is the nth candidate pitching, and φ_m is the mth candidate heading angle.

The third step is to calculate the Mean Square Residuals (MSR) variable for each $\mathbf{r}_{rb,mn}$ according to Eqs. (5) and (6).

$$MSR\left(\mathbf{r}_{rb,k}\right) = \|\mathbf{v}_k\|_2 \tag{5}$$

where $\|\cdot\|_2$ is the mean square function, $\mathbf{v}_k = \left(v_k^{21}, v_k^{31}, \ldots, v_k^{n1}\right)^{\mathrm{T}}$, with

$$v_k^{ij} = rem\left(\frac{1}{\lambda}\nabla\Delta\rho_{rb,k}^{ij} - \nabla\Delta\tilde{\Phi}_{rb}^{ij}\right) \tag{6}$$

where $rem(\cdot)$ is the remainder function (ranging from −0.5 to 0.5).

At last, validate the smallest MSR valley with an acceptance test method defined in Eq. (7). Theoretically, the minimum MSR value corresponds to the true baseline of the vehicle. However, as the existence of carrier phase multi-path noises, wrong AR results may appear and a confidence measure to access is required, which is express by the quotient of the smallest MSR valley MSR_{1st} and the second-smallest valley MSR_{2nd}.

$$\kappa_{MSR} = \frac{MSR_{2nd}}{MSR_{1st}} \tag{7}$$

If the MSR condition number κ_{MSR} exceeds an empirical threshold, which is recommended to be 1.3–1.5, the corresponding AR solution obtained from Eq. (8) below will be accepted and marked as a 'fix' result, while the others are marked as 'float' results.

$$\nabla\Delta\bar{N}_{rb}^{ij} = round\left(\frac{1}{\lambda}\nabla\Delta\rho_{rb,m}^{ij} - \nabla\Delta\tilde{\Phi}_{rb}^{ij}\right) \tag{8}$$

2.3 INS-Aided Attitude Determination

There are some inherent shortcomings of the AR algorithm above, which can be summarized as follows:

(1) Computation load of the method is heavy. As the baseline gets longer slightly, the number of candidate vectors increases quadratically.
(2) Low AR success rate under GNSS challenged environments. Since the AR is solved instantaneously at every single epoch and the 'float' ambiguities are not estimated or maintained, the algorithm is very likely to fail the acceptance test in situations where GNSS is blocked or experiences high dynamic motions.
(3) Insufficient self-adaptive ability to remain stable for different strength levels of GNSS signals. It was based on several fairly limited number of static experiments that the empirical MSR threshold range was decided, however, as for SOTM terminals, the actual complicated urban environment is unavoidable, as well as high dynamic motions, thus resulting in a considerable decline of the MSR condition number κ_{MSR}. Consequently, a growing probability of which the interested baseline vector appears at the second-smallest MSR valley can be seen. In addition, inaccurate prior knowledge, such as the baseline length error due to misalignment between the geometric centre and phase centre of low-cost GNSS antennas, may also cause failure of the AR solution.

In this section, an INS-aided ambiguity resolution algorithm is presented, which is subjected to various SOTM road conditions and requires much smaller computation amount.

There are two main reasons to introduce INS dead reckoning (DR) attitudes into the AR process. First, taking the pitch angle from INS as a priori knowledge, the afore-mentioned spherical traversal space in the AR problem can be compressed to a one-dimension searching; Second, by providing a new MSR acceptance test method, the INS heading which features high accuracy in short terms is well suited to improve system robustness and surge the AR success rate.

1. INS Pitch Angle Aiding

In general, the INS system integrates the angular velocity measured by its gyroscopes to obtain the vehicle attitude angles. If the relationship between INS body coordinate and the local navigation coordinate is well-known, we can get a set of high short-term accuracy pitch and heading predictions.

According to [6], the DD geometric range can be approximated as:

$$\frac{1}{\lambda}\nabla\Delta\rho_{rb}^{ij} \approx \left(\mathbf{e}^{j}-\mathbf{e}^{i}\right)^{\mathrm{T}}\cdot\frac{1}{\lambda}\mathbf{r}_{rb} \tag{9}$$

where $\mathbf{e}^{k}, k=i,j$ represents the line-of-sight (LOS) unit vector pointing from the GNSS receiver to the satellite k.

Similar to Eq. (4), the vector $(\mathbf{e}^{j}-\mathbf{e}^{i})$ can be expressed in spherical coordinates as:

$$\mathbf{e}^{j}-\mathbf{e}^{i}=\begin{bmatrix} e^{ij}\sin\left(\varphi^{ij}\right)\cos\left(\theta^{ij}\right) \\ e^{ij}\cos\left(\varphi^{ij}\right)\cos\left(\theta^{ij}\right) \\ e^{ij}\sin\left(\theta^{ij}\right) \end{bmatrix} \tag{10}$$

where $e^{ij}=\|\mathbf{e}^{j}-\mathbf{e}^{i}\|$ is the length of vector $(\mathbf{e}^{j}-\mathbf{e}^{i})$, θ^{ij} and φ^{ij} are its pitch and heading angles, respectively.

Here we substitute Eq. (9) with Eqs. (3), (4) and (10), resulting a scalar form approximation of the relationship between carrier phase measurements and the integer ambiguity:

$$\frac{\left(\nabla\Delta\tilde{\Phi}_{rb}^{ij}-\frac{1}{\lambda}\cdot l\sin(\theta)\cdot\|\mathbf{e}^{j}-\mathbf{e}^{i}\|\sin\left(\theta^{ij}\right)\right)+\nabla\Delta\bar{N}_{rb}^{ij}}{\|\mathbf{e}^{j}-\mathbf{e}^{i}\|\cos\left(\theta^{ij}\right)} \approx \frac{l\cos(\theta)}{\lambda}\cos\left(\varphi-\varphi^{ij}\right) \tag{11}$$

Define heading projections of $(\mathbf{e}^{j}-\mathbf{e}^{i})$ and $\mathbf{r}_{rb}$ as following.

$$l_{\theta}=l\cos(\theta), \quad e_{\theta}^{ij}=e^{ij}\cos\left(\theta^{ij}\right) \tag{12}$$

And take the term $\hat{\Phi}^{ij}=\frac{1}{\lambda}\cdot l\sin(\theta)\cdot e^{ij}\sin\left(\theta^{ij}\right)$ as a DD carrier phase component of $\nabla\Delta\tilde{\Phi}_{rb}^{ij}$ perpendicular to the horizontal plane, we can get:

$$\frac{\left(\nabla\Delta\tilde{\Phi}_{rb}^{ij} - \hat{\Phi}^{ij}\right) + \nabla\Delta\bar{N}_{rb}^{ij}}{e_{\theta}^{ij}} \approx \frac{l_{\theta}}{\lambda}\cos\left(\varphi - \varphi^{ij}\right) \tag{13}$$

As can be seen in Eq. (13), the theoretical two dimensional DD measurement model in Eq. (3) is now approximated by a scalar expression. The interested DD integer ambiguity $\nabla\Delta\bar{N}_{rb}^{ij}$ can be determined by a given heading φ.

The geometric relationship between the circular baseline vector search space and the combined fractional parts of the projected DD carrier phase measurements is depicted in Fig. 1. The plane in yellow represents the projected heading plane. A circle determined by l_{θ} on it is the baseline vector search space. According to Eq. (13), for each fractional DD carrier phase, we can draw a set of candidate points which are symmetric about φ^{ij} on the circle. Different symmetrical points correspond to different ambiguities and the true baseline vector must be at one of the listed points. For error-free situations, if there are several DD pairs of available forming multiple sets of symmetric points, the true baseline vector could be determined as the point shared by all these candidate sets, which is marked by the green point in Fig. 1.

2. INS Yaw Angle Aiding

An INS-aided acceptance test method is introduced in Fig. 2. Instead of labelling candidate AR solutions as 'fix' or 'float' according to the MSR ratio κ_{MSR}, we evaluate the candidate baseline vectors in three steps.

Firstly, validate both the smallest and the second-smallest MSR valleys according to constraint formulas of Eqs. (14) and (15).

$$MSR1: \left|\tilde{r}_{rb1} - \hat{r}_{rb}\right| < \delta r_{thresh}, \left|\tilde{\theta}_1 - \theta_{INS}\right| < \delta\theta_{thresh}, \frac{MSR2}{MSR1} > \kappa_{thresh} \tag{14}$$

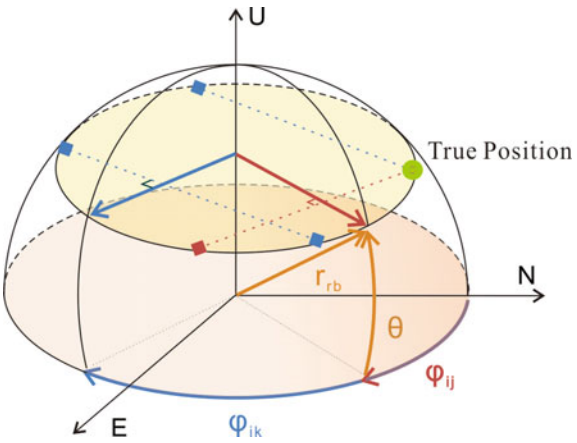


Fig. 1 Relationship between the projected baseline vector in heading plane and the fractional parts of the DD carrier phase measurements

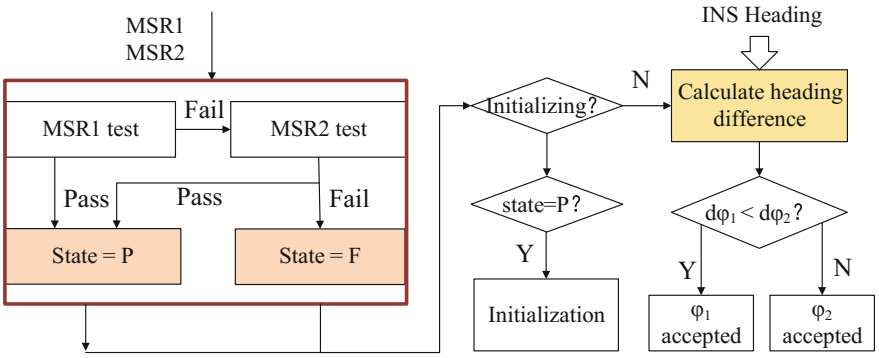


Fig. 2 Acceptance test method with INS heading assistance

$$MSR2: |\tilde{r}_{rb2} - \hat{r}_{rb}| < \delta r_{thresh}, \left|\tilde{\theta}_2 - \theta_{INS}\right| < \delta\theta_{thresh}, |\tilde{r}_{rb1} - \hat{r}_{rb}| > \delta r_{thresh} \quad (15)$$

where $\tilde{r}_{rb1}$ and $\hat{r}_{rb}$ are the GNSS estimated and pre-defined baseline length respectively, δr_{thresh} is the corresponding error threshold, $\tilde{\theta}$ and θ_{INS} are the GNSS estimated and the INS reference heading angles respectively, $\delta\theta_{thresh}$ is similarly the single epoch heading error threshold. Either MSR_{1st} or MSR_{2nd} passes the acceptance test, the GNSS attitude determination state will be set as 'P' (Pass), otherwise, set as 'F' (Fail).

Secondly, initiate the two-antenna GNSS attitude determination system. It is worth to be noted that the INS yaw angle aiding is unavailable at this stage, due to lacking of initial values from GNSS receivers. Therefore, several times of iteration are required, during which the state flag must be 'P', to avoid miss-initialization.

At last, in situations where the referential INS heading is available, select the statistically most likely (ML) baseline vector, which is the one with smaller relative error with respect to the reference.

It has to be mentioned that, even though considerably reducing the misdetection probability of the AR problem, the acceptance test method will unavoidably increase the probability of false AR detection. This problem will be solved in the following section.

3 Innovation Pre-filtered Extended Kalman Filter

Compared with traditional GNSS/INS integrated systems, there are some additional requirements demanded by SOTM terminals. (1) The pointing accuracy should be prioritized in the sensor fusion algorithm, since it is the key to a stable signal tracking of SOTM terminals. (2) The ability of detecting faulty measurements is required to deal with wrong AR solutions.

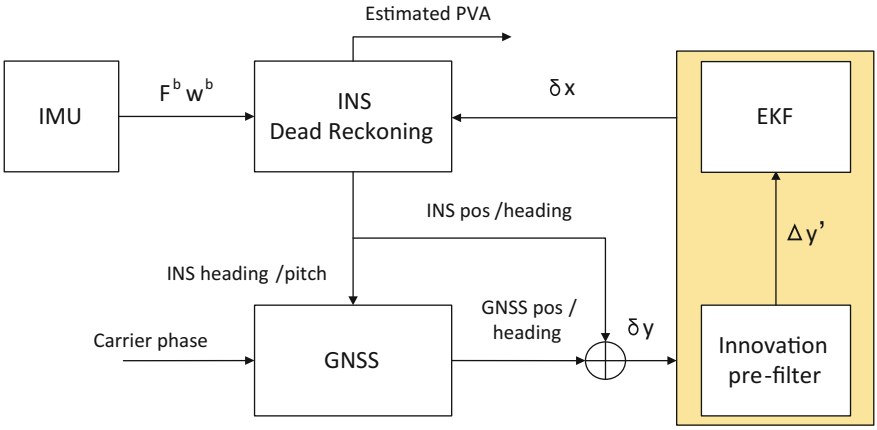


Fig. 3 Process of the Innovation Pre-filtered EKF (IPEKF)

In this section, we describes an innovation pre-filtering based Extended Kalman Filter (IPEKF) as depicted in Fig. 3. In the innovation pre-filter, the position and heading difference between INS prediction and two-antenna GNSS measurements, which is a signal that contains two noise component-the predominantly low-frequency INS component and the predominantly high-frequency differential GNSS component [7], is checked. Depending on the pre-filter outputs, measurement equation of the Kalman Filter is adjusted accordingly.

1. EKF models

The EKF algorithm is applied to nonlinear discrete-time systems of the following form:

$$\tilde{\mathbf{x}}_{k+1} = \Phi_k \hat{\mathbf{x}}_k + \boldsymbol{\omega}_k \tag{16a}$$

$$\mathbf{y}_{k+1} = \mathbf{H}_{k+1}\mathbf{x}_{k+1} + \boldsymbol{\nu}_{k+1} \tag{16b}$$

where in Eq. (16a), $\hat{\mathbf{x}}_k = \begin{bmatrix} \mathbf{p}_n^T & \mathbf{v}_n^T & \boldsymbol{\alpha}^T & \mathbf{x}_a^T & \mathbf{x}_g^T \end{bmatrix}^{\mathrm{T}}$ is the estimated state of epoch k, containing positions (BLH), velocities (NED), attitudes (Euler), accelerometer biases and gyroscope biases of the system, the Φ_k-matrix represents the linearized discrete transition matrix, and $\boldsymbol{\omega}_k$ denotes the process noise.

Equation (16b) is the measurement equation, consisting of the relationship between $\mathbf{x}_{k+1}$ and the measured GNSS-derived states $\mathbf{y}_{k+1}$, given by the design matrix $\mathbf{H}_{k+1}$ and a white Gaussian noise $\boldsymbol{\nu}_{k+1}$. Statistical properties of the noise term $\boldsymbol{\omega}$ and $\boldsymbol{\nu}$ are given by:

$$\begin{cases} E[\boldsymbol{\omega}_k] = \mathbf{0}, E[\mathbf{v}_k] = \mathbf{0} \\ E\left[\boldsymbol{\omega}_k \mathbf{v}_j^T\right] = \mathbf{0} \end{cases}, \quad \begin{cases} E\left[\boldsymbol{\omega}_k \boldsymbol{\omega}_j^T\right] = \mathbf{Q}_k \delta_{kj} \\ E\left[\mathbf{v}_k \mathbf{v}_j^T\right] = \mathbf{R}_\mathbf{k} \delta_{kj} \end{cases} \tag{17}$$

where $\mathbf{Q}_k$ and $\mathbf{R}_\mathbf{k}$ are covariance of the noises, δ_{kj} is the Kronecker delta function.

The question of components of vector $\mathbf{y}_{k+1}$ is addressed here. For fault-free GNSS measurements, $\mathbf{y}_{k+1}$ is augmented by $\mathbf{y} = [\delta L \quad \delta l \quad \delta h \quad \delta\varphi]^{\mathrm{T}}$ with the according noise vector $\mathbf{v} = \left[\xi_L \quad \xi_l \quad \xi_h \quad \xi_\varphi\right]^{\mathrm{T}}$. However, if coarse error of GNSS-derived states is detected by the pre-filter, the system will accordingly decide to either switch off the heading dimension or skip the entire EKF process of the current epoch.

2. Innovation Pre-filtering

The measurement innovation is defined as the difference between estimated states and the measurements:

$$\delta \mathbf{y}_k^- = \mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^- . \tag{18}$$

It provides an indication of whether the measurements and state estimates are consistent and can be used to detect large faults immediately.

The covariance of the innovations $\mathbf{C}_k^-$, comprises the sum of the measurement noise and the error covariance of the state estimations is transformed into measurement space [8]. Thus:

$$\mathbf{C}_k^- = \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \tag{19}$$

which is the dominator of the EKF gain.

The normalized innovations are defined as

$$\mathbf{z}_{k,j}^- = \frac{\delta \mathbf{y}_{k,j}^-}{\sqrt{\mathbf{C}_{k,j,j}^-}} \tag{20}$$

In an ideal Kalman Filter, the distribution of $\mathbf{z}_{k,j}^-$ is a zero-mean Gaussian distribution. In fact, even though differences exists between true and the modelled system noises, an empirical innovation threshold can still be used for fault detection.

Results of the innovation pre-filter of the yaw angle during a dynamic experiment with 0.267 m baseline is presented in Fig. 4. As can be seen, the GNSS measurement at point P1 is noisy, but the one at point P2 is faulty and should be eliminated.

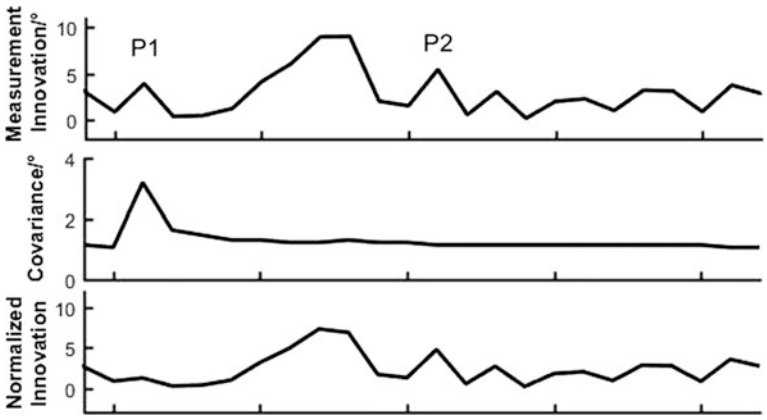


Fig. 4 Faults detection of innovation pre-filtering

4 Experiments

The SOTM GNSS/INS attitude and position determination system was test on the experimental platform in Fig. 5. The yellow SOTM antenna array is the objective of our system. It was at each side of the array, the base and rover GNSS antennas are mounted, with baseline length of 0.267 m. The left part of the figure is a prototype FPGA module, consisting of two GNSS receivers (Unicore UM220-III H), one INS (Xsens MTi-3-8A7G6) and an ARM micro-processor (STM32F427). As can be seen in the figure, to reduce the installation space, a fix pitch angle bias from GNSS baseline vector to the FPGA plane is introduced, which should be corrected in the IPEKF process.

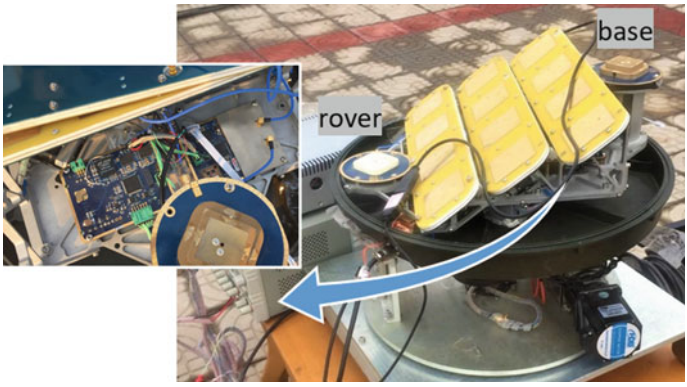


Fig. 5 Experiment platform of the GNSS/INS attitude and position determination system

1. Static Experiments

To validate performance of the GNSS/INS system, standard deviations of the platform position and Euler angles are presented in Table 1. Both two data sets were collected statically in GNSS-friendly environments.

As can be seen in the table, the standard deviation of SOTM positioning is smaller than 2 m, while the static heading accuracy is about 0.01°.

2. Kinematic Experiments

A kinematic test was conducted to evaluate stability and robustness of the GNSS/INS system along the route depicted in Fig. 6. Poor GNSS signal with leaf cover and building canyon was suffered in the red color labeled area. Reference solutions were provided from the NovAtel's high precision GNSS/INS integration system: SPAN-CPT.

Our attention here is to highlight attitude estimation performance of the proposed system.

- Attitude availability

Figure 7 compares the estimated Euler angles from the proposed system with the reference solutions of SPAN-CPT. It is shown in the photo that from second 1437 s to 1455 s, the vehicle turned around twice and caused sharp changing of the

Table 1 Standard deviations of the estimated SOTM positions and Euler angles

Data	North (m)	East (m)	Down (m)	Heading (°)	Pitch (°)	Roll (°)
1	1.65	0.77	0.49	0.009	0.137	0.030
2	1.35	1.06	0.97	0.011	0.006	0.003

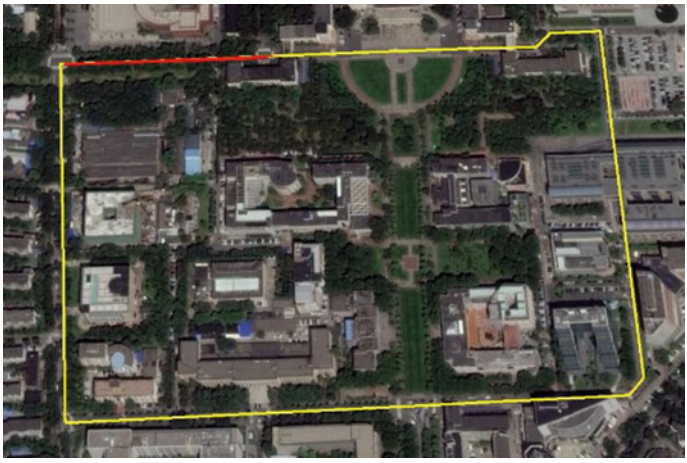


Fig. 6 Driving route of the kinematic experiment

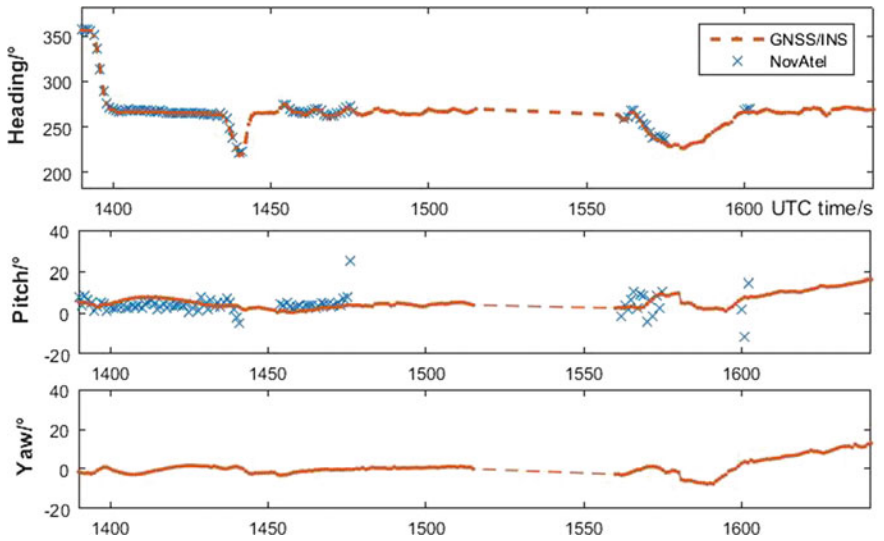


Fig. 7 Comparison of Euler angles from the GNSS/INS integrated system and SPAN-CPT (The roll angle is unavailable from SPAN-CPT)

heading angle. During this period, our integrated system succeeded to output attitude solutions continuously, while the reference system failed for more than 10 s. Later on, the SOTM system experienced a long time of GNSS outage start from around 1475 s. After SPAN-CPT stopped working, our GNSS/INS integrated system lasted approximately 40 s longer until 1515 s.

The dynamic heading and pitch angle errors of our proposed system in GNSS-friendly areas are plotted in Fig. 8 with the standard deviation of 0.277° and 0.753°, respectively.

- Attitude robustness

Figure 9 demonstrates the performance of our AR acceptance test method. The relationship between the selection of GNSS optimal/suboptimal heading results and

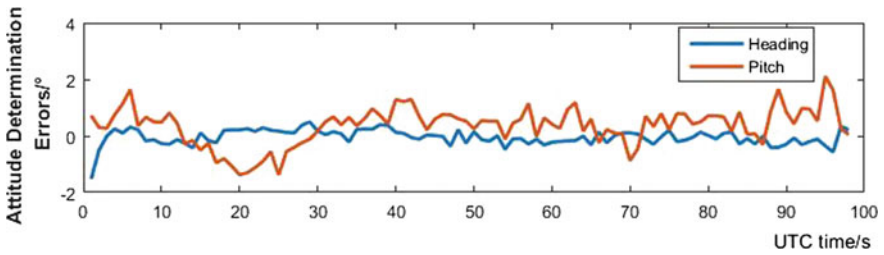


Fig. 8 Heading and pitching errors of the GNSS/INS integrated system

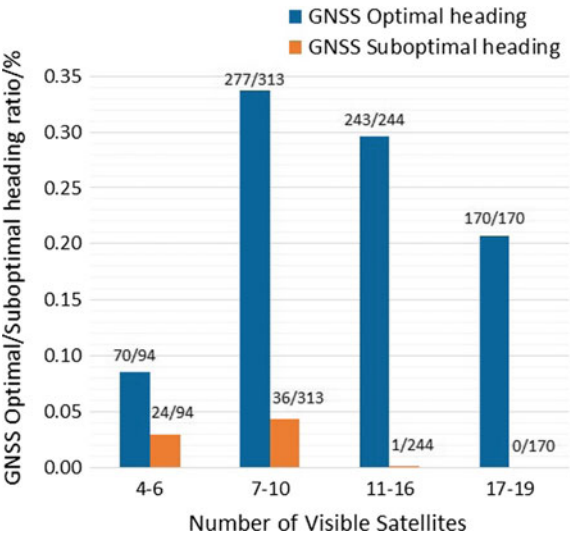


Fig. 9 Relationship between the selection of GNSS optimal/sub-optimal headings and the number of visible satellites

the number of visible satellites is plotted. Here we define the optimal solution as the candidate GNSS heading of state 'P', and the other one defined as a suboptimal solution.

As can be seen, when the number of visible GNSS satellites is less than 10, the two-antenna system is significantly affected by the observation noises and a high possibility of GNSS miss detection appears. With the assistance of INS heading information, GNSS heading result selection is optimized and the AR solution success rate is improved considerably.

5 Conclusions

In this paper, a low-cost GNSS/INS attitude and position determination system with an ultra-short baseline (0.2–0.5 m) is demonstrated, utilizing consumer-grade GNSS receivers and MEMS INS. Contrast with traditional sensor fusion algorithms, the INS and GNSS systems are no longer independent to each other. The INS dead reckoning heading and pitch angles are introduced into the process of GNSS attitude determination to increase AR success rate and decrease misdetection possibilities. Measurement outliers are detected and eliminated in the IPEKF process. Experimental results show that the system is able to output high accurate state estimates of the SOTM system with strong robustness and high stability.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 41474027).

References

1. Beljour H et al (2009) Army SATCOM on the move technology initiatives. In: MILCOM 2009, 2009 IEEE military communications conference, Boston, MA, pp 1–7
2. Fish T (2017) Communication complications. Digital Battlespace 9(3):14–20
3. Wu Z, Yao M, Ma H, Jia W, Tian F (2013) Low-cost antenna attitude estimation by fusing inertial sensing and two-antenna GPS for vehicle-mounted satcom-on-the-move. IEEE Trans Veh Technol 62(3):1084–1096
4. Girrbach F et al (2017) Optimization-based sensor fusion of GNSS and IMU using a moving horizon approach. Sensors 17(5):1159
5. Misra P (2011) Global positioning system: signals, measurements, and performance. Ganga-Jamuna Press, USA
6. Li W, Cui X, Zhao S, Lu M (2017) Single-epoch ambiguity resolution for urban ultra-short baseline attitude determination using low-cost GNSS receivers. In: Proceedings of the 30th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS + 2017), Portland, Oregon, pp 4026–4037
7. Farrell JA, Givargis TD, Barth MJ (2000) Real-time differential carrier phase GPS-aided INS. IEEE Trans Control Syst Technol 8(4):709–721
8. Groves PD (2013) Principles of GNSS, inertial, and multisensor integrated navigation systems. Artech House, USA

Implementation of Mixed Sequential Kalman Filters for Vision-Aided GNSS/INS Integrated Navigation System

Shanhui Mo, Zebo Zhou, Shuang Du, Changgan Xiang, Changhong Kuang and Jin Wu

Abstract This paper deals with the integration problem for GNSS/INS integration navigation with the aid of vision sensor. The basic principle of visual navigation is presented and discussed. Given the triple sampling rates of sensors, we design a vision-aided GNSS/INS fusion method based on the sequential Kalman filter structure. The sequential Kalman filters are mixed with both extended Kalman filter (EKF) and linear Kalman filter (LKF). The fusion computation contains three layers. (i) The first layer is INS mechanization and its 15-dimensional error state model, including three-dimensional position, velocity, attitude (PVA), gyroscope drift and accelerometer bias. (ii) The second layer is monocular camera model containing a sliding window where Fast features will be extracted and tracked within a set of corrected frames. If the parallax is sufficient, then use triangulation and Gauss-Newton optimization to estimate their coordinates in navigation frame. Considering each feature tracked by multiple frames, the state of monocular camera is naturally constrained with a nonlinear model via each landmark. By linearizing the nonlinear model of monocular camera, we apply the EKF to estimate the error state and compensate for INS PVA. (iii) The third is the lowest sampling-rate GNSS PV model. With a corrected INS state based on (ii), a traditional GNSS/INS loosely integration will be implemented through LKF. Finally, the real dataset (KITTI) of a land vehicle is employed to demonstrate the efficiency of our proposed fusion method. The results show the competitive accuracy and reliability of our method. With the aid of VO, the performance of GNSS/INS is significantly

S. Mo · Z. Zhou (✉) · S. Du · J. Wu
School of Aeronautics and Astronautics, University of Electronic Science Technology of China (UESTC), Chengdu, People's Republic of China
e-mail: klinsmann.zhou@gmail.com

Z. Zhou
Laboratory of BeiDou Integrated Applications, University of Electronic Science Technology of China (UESTC), Chengdu, People's Republic of China

C. Xiang · C. Kuang
Sichuan Jiuzhou Beidou Navigation and Location Service Co., Ltd, Chengdu, People's Republic of China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_54

improved. Especially for these GNSS challenging environments, corrected by monocular camera, INS can endure a much longer duration without GNSS signals.

Keywords Visual odometry · GNSS/INS · Extended Kalman filter
Multi-sensor fusion

1 Introduction

In modern navigation campaigns, inertial navigation system (INS) is very attractive and has many merits for integrating with global navigation satellite systems (GNSS). A typical GNSS/INS integrated system, besides bridging the GNSS navigation gaps with inertial sensors in many challenging environment e.g. city canyons, underwater, signal jamming etc. it can efficiently overcome the drawbacks of inertial drifts and biases with the aid of GNSS. However, the integrated navigation accuracy is strongly dependent on the inertial instruments. There are mainly three ways to improve the integrated navigation system performance: the first is using a high-end laser or fiber optic gyro which is very expensive and occupies a lot of volume for motion platforms. The second way is to introduce non-holonomic constraints, e.g. length, velocity and attitude constraints [1]. The third way is using auxiliary sensors, e.g. magnetometer, barometer, cameras, LOCATA etc. [2, 3]. Recently, with the increasing demands of seamless navigation, visual odometry (VO) is drawn intensive attentions from researchers in autonomous simultaneous localization and mapping (SLAM) fields [4–7]. It generates the accurate and consistent egomotion estimation by utilizing the tracked features in consecutive frames. Thus it can also continuously correct the INS error in case GNSS signals are absent. Such, it naturally becomes a good auxiliary for improving the overall performances of GNSS/INS navigation system. This stands in a new perspective to contribute to the accuracy and robustness with the aid of vision sensor.

Different sampling rates should be considered before fusing inertial measurement unit (IMU) and visual sensor. Forster proposed a pre-integration method by combining inertial measurements between two consecutive frames on SO(3) manifolds [8], which has been used as reference for other VIO algorithm such as ORB-SLAM+IMU [9], SVO+IMU [8] etc. Great deals of research efforts are made to integration of visual-inertial odometry (VIO) and IMU. These can be mainly divided into two categories: loosely and tightly coupled integrations. Very few literatures are related to loosely coupled integration for VIO. On the contrary, the tightly coupled integration methods are intensively investigated and developed, the most representative methods are MSCKF [10, 11], ROVIO [12] which are derived in the framework of Kalman filter (KF) [10–12]. Others e.g. OKVIS [13, 14], VINS [15], ORB-SLAM+IMU [9], SVO+IMU [8] are proposed based on the framework of nonlinear optimization [8, 9, 13–15]. MSCKF augments the state and covariance matrix at the time of a new image arrival [10]. It triangulates the features to obtain their positions in global reference frame and establishes the observation model

based on the landmarks. The latter work referred to as MSCKF 2.0 further analyses system observability and addresses the issue of inconsistency of EKF approaches [11]. VINS, proposed by Shen et al. in 2016, optimizes the visual, inertial measurements within a sliding window. It is run real time in an embedded device and show superior results. According to reference [16], optimization based VIO generates the more accurate estimation result compared with the KF based VIO. However, it consumes much more computation time. Overall, vision-aided GNSS/INS navigation system has a significant merit and will play an important role in the modern-time navigation, especially for those GNSS challenging environments. With the use of visual sensors, the INS will endure a much longer duration to bridge the navigation gaps. In this paper, a mixed sequential Kalman filtering structure is proposed to combine the information from the triple sampling rates of VO/GNSS/INS sensors. This paper is arranged as follows. Section 2 briefs the fundamentals of vision navigation. The fusion algorithm for vision-aided GNSS/INS integrated system is proposed in Sect. 3. The experiment and analysis are given in Sect. 4. Finally the concluding remarks are drawn.

2 Vision Navigation Fundamentals

2.1 Fast Feature Detector and Optical Flow

In practice, scale-invariant feature transform (SIFT) is unable to operate the computations on CPU in real time. Likewise, ORB consumes nearly 20 ms of time [5]. Therefore, we chose the fast feature rather than SIFT, SURF, ORB etc. to ensure a real-time running. The fast feature is a simple corner, which detects the changes in the gray level of the local pixels. Herein, we capture fast features by using L-K optical flow [16], which is a sparse-type method based on pixel gradient by assuming brightness invariant as follows:

$$\begin{bmatrix} [I_x, I_y]_1 \\ \vdots \\ [I_x, I_y]_k \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_{t_1} \\ \vdots \\ I_{t_k} \end{bmatrix} \tag{1}$$

where the pair (u, v) denote the velocity vector of pixel on the x and y components respectively; I_x, I_y and I_t represent the derivatives of the image at the coordinate (x, y) and t respectively. Let $\boldsymbol{A} = \begin{bmatrix} [I_x, I_y]_1 & \cdots & [I_x, I_y]_k \end{bmatrix}^{\mathrm{T}}$ and $\boldsymbol{I}_t = \begin{bmatrix} I_{t_1} & \cdots & I_{t_k} \end{bmatrix}^{\mathrm{T}}$, Eq. (1) can be further rewritten as the matrix-form equations,

$$\boldsymbol{A} \begin{bmatrix} u \\ v \end{bmatrix} = -\boldsymbol{I}_t \tag{2}$$

With the derivatives of multiple pixels in a given window, the least-squares principle can be employed to estimate the velocity (u, v) of tracked features by

$$\begin{bmatrix} u \\ v \end{bmatrix} = -\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{I}_t \tag{3}$$

2.2 Triangulation

After optic flow tracking, the arbitrary feature's position (containing the depth information) in the navigation frame (n-frame) is then calculated with a set of consecutive frames.

As shown in Fig. 1, the feature p_1 in the 1st frame matches p_2 in the 2nd frame. Ideally, the intersection between o_1p_1 and o_2p_2 should be point p. For convenience, let p_1 and p_2 be normalized coordinates of the features respectively, $(\boldsymbol{R}_2^1, \boldsymbol{d})$ represents rotation matrix and translation vector, then we have

$$s_1\boldsymbol{p}_1 = s_2\boldsymbol{R}_2^1\boldsymbol{p}_2 + \boldsymbol{d} \tag{4}$$

where s_1 and s_2 are the vision depth of p_1 and p_2 respectively. By multiplying orthogonal vector $p_1^{\wedge}$ on both sides of Eq. (4), it arrives

$$\begin{aligned} s_1\boldsymbol{p}_1^{\wedge}\boldsymbol{p}_1 &= s_2\boldsymbol{p}_1^{\wedge}\boldsymbol{R}\boldsymbol{p}_2 + \boldsymbol{p}_1^{\wedge}\boldsymbol{d} \\ s_1\boldsymbol{p}_1^{\wedge}\boldsymbol{p}_1 &= 0 \end{aligned} \tag{5}$$

Evidently, the s_1 and s_2 are trivially computed. By setting the p1 and p2 as the initial values, we can obtain the feature's position by utilizing the Gauss-Newton minimization with inverse-depth [16].

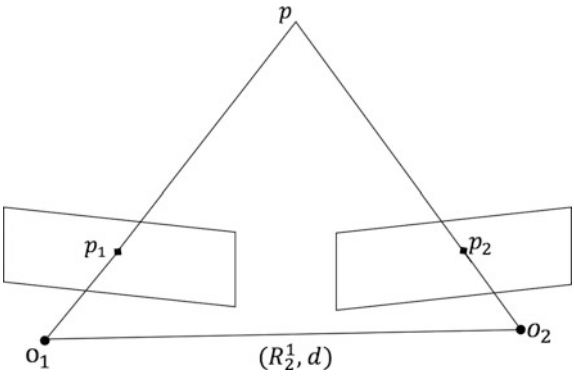


Fig. 1 The triangulation sketch for two frames

3 Vision-Aided GNSS/INS Navigation Estimation

In this section, we will investigate the Kalman filtering theory for solving the vision-aided GNSS/INS integrated navigation estimation problem.

3.1 State Representation for Integrated System

Similar to MSCKF [10], the full state $\boldsymbol{x}$ can be partitioned into two blocks. One part is the 16-dimensional IMU position, velocity and attitude (PVA) state. The rest is the N positions and poses of camera in a sliding window,

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_{\mathrm{IMU}}^{\mathrm{T}} & \boldsymbol{x}_{C_{l-1}}^{\mathrm{T}} & \boldsymbol{x}_{C_{l-2}}^{\mathrm{T}} & \cdots & \boldsymbol{x}_{C_{l-N}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{6}$$

where $\boldsymbol{x}_{\mathrm{IMU}} = \begin{bmatrix} \left(\boldsymbol{q}_{\mathrm{n}}^{\mathrm{b}}\right)^{\mathrm{T}} & \boldsymbol{b}_{\mathrm{w}}^{\mathrm{T}} & \left(\boldsymbol{v}^{\mathrm{n}}\right)^{\mathrm{T}} & \boldsymbol{b}_{\mathrm{a}}^{\mathrm{T}} & \left(\boldsymbol{p}^{\mathrm{n}}\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ $\boldsymbol{q}_{\mathrm{n}}^{\mathrm{b}}$ represents the unit quaternion from the navigation frame (n-frame) to the body frame (b-frame); $\boldsymbol{v}^{\mathrm{n}}$ and $\boldsymbol{p}^{\mathrm{n}}$ are defined as the velocity and position; $\boldsymbol{b}_{\mathrm{w}}$ and $\boldsymbol{b}_{\mathrm{a}}$ are the gyroscope drift and accelerometer bias vectors. The i-th state of camera $\boldsymbol{x}_{C_i} = \begin{bmatrix} \left(\boldsymbol{q}_{\mathrm{n}}^{C_i}\right)^{\mathrm{T}} & \left(\boldsymbol{p}_{C_i}^{\mathrm{n}}\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, where $\boldsymbol{q}_{\mathrm{n}}^{C_i}$ and $\boldsymbol{p}_{C_i}^{\mathrm{n}}$ are the camera's attitude and position respectively. Then the IMU error state is naturally described as $\delta\boldsymbol{x} = \begin{bmatrix} \delta\boldsymbol{x}_{IMU}^{\mathrm{T}} & \delta\boldsymbol{x}_{C_{l-1}}^{\mathrm{T}} & \delta\boldsymbol{x}_{C_{l-2}}^{\mathrm{T}} & \cdots & \delta\boldsymbol{x}_{C_{l-N}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ in which the 15-dimensional IMU error state reads

$$\delta\boldsymbol{x}_{\mathrm{IMU}} = \begin{bmatrix} \left(\delta\boldsymbol{\theta}_{\mathrm{IMU}}\right)^{\mathrm{T}} & \delta\boldsymbol{b}_{\mathrm{w}}^{\mathrm{T}} & \left(\delta\boldsymbol{v}^{\mathrm{n}}\right)^{\mathrm{T}} & \delta\boldsymbol{b}_{\mathrm{a}}^{\mathrm{T}} & \left(\delta\boldsymbol{p}^{\mathrm{n}}\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{7}$$

and the 6-dimensional error pose and position state of camera $\delta\boldsymbol{x}_{C_i} = \begin{bmatrix} \left(\delta\boldsymbol{\theta}_{c_i}\right)^{\mathrm{T}} & \left(\delta\boldsymbol{p}_{c_i}^{n}\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$.

3.2 IMU Error State Propagation

The IMU state and its error propagate occurs upon receiving a new IMU data. The inertial navigation differential equations are

$$\dot{\boldsymbol{q}}_{\mathrm{n}}^{\mathrm{b}} = \frac{1}{2}\boldsymbol{\Omega}(\hat{\boldsymbol{w}})\boldsymbol{q}_{\mathrm{n}}^{\mathrm{b}} \quad \boldsymbol{\Omega}(\hat{\boldsymbol{w}}) = \begin{bmatrix} -(\hat{\boldsymbol{w}}\times) & \hat{\boldsymbol{w}} \\ -(\hat{\boldsymbol{w}})^{\mathrm{T}} & 0 \end{bmatrix} \tag{8}$$

$$\dot{\boldsymbol{p}}^{\mathrm{n}} = \boldsymbol{v}^{\mathrm{n}}, \quad \dot{\boldsymbol{v}}^{\mathrm{n}} = \hat{\boldsymbol{a}}, \quad \dot{\boldsymbol{b}}_{\mathrm{w}} = \boldsymbol{n}_{\mathrm{w}_g}, \quad \dot{b}_{\mathrm{a}} = \boldsymbol{n}_{\mathrm{w}_a} \tag{9}$$

where $\hat{\boldsymbol{a}} = \boldsymbol{R}_{\mathrm{b}}^{\mathrm{n}}(\boldsymbol{a}_{\mathrm{m}} - \boldsymbol{b}_{\mathrm{a}})$ and $\hat{\boldsymbol{w}} = \boldsymbol{w}_{\mathrm{m}} - \boldsymbol{b}_{\mathrm{w}}$; $\boldsymbol{a}_{\mathrm{m}}$ and $\boldsymbol{w}_{\mathrm{m}}$ represent the measurements of accelerometer and gyroscope respectively. The 4-th Runge-Kutta method is used for gyro integration. The IMU error state model can be modelled as,

$$\delta\dot{\boldsymbol{x}}_{\mathrm{IMU}} = \boldsymbol{F}\delta\boldsymbol{x}_{\mathrm{IMU}} + \boldsymbol{G}\boldsymbol{n}_{\mathrm{IMU}} \tag{10}$$

The coefficient matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are given by

$$\boldsymbol{F} = \begin{bmatrix} -(\hat{\boldsymbol{w}}\times) & -\boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ -\boldsymbol{R}_{\mathrm{b}}^{\mathrm{n}}(\hat{\boldsymbol{a}}\times) & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & -\boldsymbol{R}_{\mathrm{b}}^{\mathrm{n}} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \end{bmatrix},$$

$$\boldsymbol{G} = \begin{bmatrix} -\boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & -\boldsymbol{R}_{\mathrm{b}}^{\mathrm{n}} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \end{bmatrix}$$

3.3 State Covariance Propagation

Since the state estimate involves states of both IMU and camera, accordingly the covariance of state $\boldsymbol{x}$ be partitioned into four parts as

$$\boldsymbol{P}_{k|k} = \begin{bmatrix} \boldsymbol{P}_{\mathrm{II}_{k|k}} & \boldsymbol{P}_{\mathrm{IC}_{k|k}} \\ \boldsymbol{P}_{C\mathrm{I}_{k|k}} & \boldsymbol{P}_{CC_{k|k}} \end{bmatrix} \tag{11}$$

where $\boldsymbol{P}_{\mathrm{II}}$ denotes the covariance of the current IMU state. $\boldsymbol{P}_{\mathrm{CC}}$ is the covariance of the camera positions and poses within the sliding window. $\boldsymbol{P}_{\mathrm{IC}}$ indicates the cross-correlations of the IMU state and the N camera state. Assuming IMU sampling time interval as Δt, referring to [14], further propagating the covariance information into Eq. (12) yields

$$\boldsymbol{P}_{k+1|k} = \begin{bmatrix} \boldsymbol{\Phi}_{\Delta t}\boldsymbol{P}_{\mathrm{II}}\boldsymbol{\Phi}_{\Delta t}^{\mathrm{T}} + \boldsymbol{G}\boldsymbol{Q}_{\mathrm{IMU}}\boldsymbol{G}^{\mathrm{T}}\Delta t & \boldsymbol{\Phi}_{\Delta t}\boldsymbol{P}_{\mathrm{IC}_{k|k}} \\ \boldsymbol{P}_{\mathrm{IC}_{k|k}}^{\mathrm{T}}\boldsymbol{\Phi}_{\Delta t}^{\mathrm{T}} & \boldsymbol{P}_{CC_{k|k}} \end{bmatrix} \tag{12}$$

where the state transition matrix $\boldsymbol{\Phi}_{\Delta t} = \boldsymbol{I}_{15\times15} + \boldsymbol{F}\Delta t$ and $\boldsymbol{Q}_{\mathrm{IMU}}$ is the covariance matrix of $\mathrm{n}_{\mathrm{IMU}}$.

3.4 State Augmentation with Camera State

Upon a new image frame obtained, the state $\boldsymbol{x}$ and its corresponding covariance will be augmented in the following way,

$$\boldsymbol{x} \leftarrow \begin{bmatrix} \boldsymbol{x}^{\mathrm{T}} & \boldsymbol{x}_{C_{l-(N+1)}}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \quad \boldsymbol{p}_{k|k} \leftarrow \begin{bmatrix} \boldsymbol{I}_{(6N+15)\times(6N+15)} \\ \boldsymbol{J} \end{bmatrix} \boldsymbol{p}_{k|k} \begin{bmatrix} \boldsymbol{I}_{(6N+15)\times(6N+15)} \\ \boldsymbol{J} \end{bmatrix}^{\mathrm{T}} \tag{13}$$

where J is computed by

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{R}_{\mathrm{b}}^{C} & \boldsymbol{0}_{3\times 9} & \boldsymbol{0}_{3\times 3} & \boldsymbol{0}_{3\times 6N} \\ \left[(\boldsymbol{R}_{\mathrm{b}}^{\mathrm{n}} \boldsymbol{p}_{C}^{\mathrm{b}})\times\right] & \boldsymbol{0}_{3\times 9} & \boldsymbol{I}_{3\times 3} & \boldsymbol{0}_{3\times 6N} \end{bmatrix} \tag{14}$$

In which, $\boldsymbol{p}_{C}^{\mathrm{b}}$ is a baseline vector from central of camera to IMU and it can be calibrated beforehand. Accordingly, the state of i-th camera frame $\boldsymbol{x}_{C_i} = \begin{bmatrix} \left(\boldsymbol{q}_{\mathrm{n}}^{C_i}\right)^{\mathrm{T}} & \left(\boldsymbol{p}_{C_i}^{\mathrm{n}}\right)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ is computed by

$$\begin{aligned} \boldsymbol{q}_{\mathrm{n}}^{C_i} &= \boldsymbol{q}_{\mathrm{b}}^{C_i} \otimes \boldsymbol{q}_{\mathrm{n}}^{\mathrm{b}} \\ \boldsymbol{P}_{C_i}^{\mathrm{n}} &= \boldsymbol{P}^{\mathrm{n}} + \boldsymbol{R}_{\mathrm{b}}^{\mathrm{n}} \boldsymbol{P}_{C_i}^{\mathrm{b}} \end{aligned} \tag{15}$$

3.5 Camera Observation Model Construction

With extracting corners from a set of images, the features points are well tracked by using L-K optical flow mentioned in Sect. 2. Then we estimate the reliable feature's position in the n-frame using the Gauss-Newton minimization with inverse-depth. For instance, $\boldsymbol{P}_{j,c_i}^{\mathrm{n}}$ represents the j-th feature's position observed by the i-th frame. Since the camera observation is nonlinear, we implement the linearization operation to the original camera observation equation as

$$\boldsymbol{z}_{\mathrm{m}j}^{i} - \boldsymbol{z}_{j}^{i} = \boldsymbol{r}_{j}^{i} \approx \boldsymbol{H}_{j}^{\delta \boldsymbol{x}_i} \delta \boldsymbol{x}_i + \boldsymbol{H}_{j}^{c_i} \delta \boldsymbol{p}_{j,c_i}^{\mathrm{n}} + \boldsymbol{n}_{j}^{i} \tag{16}$$

Here, $\boldsymbol{z}_{m_j}^{i}$ is the j-th feature's coordinate observed by the i-th image frame. $\boldsymbol{z}_{j}^{i}$ is the calculated one by transforming $\boldsymbol{P}_{j,c_i}^{\mathrm{n}}$ into the normalized camera coordinate system. The matrix $\boldsymbol{H}_{j}^{\delta \boldsymbol{x}_i}$ and $\boldsymbol{H}_{j}^{c_i}$ are the observation Jacobian matrices with respect to the error state $\delta \boldsymbol{x}_i$ and the feature position error $\delta \boldsymbol{p}_{j,c_i}^{\mathrm{n}}$. The details of the Jacobians

can be found in [10, 16]. By stacking all the observation residuals of j-th feature point, we obtain

$$\boldsymbol{r}_j \approx \boldsymbol{H}_j^{\delta x}\delta\boldsymbol{x} + \boldsymbol{H}_j^{c}\delta\boldsymbol{p}_j^{\mathrm{n}} + \boldsymbol{n}_j \tag{17}$$

$\boldsymbol{H}_j^{c} = \begin{bmatrix}\boldsymbol{H}_j^{c_1} & \boldsymbol{H}_j^{c_2} & \cdots & \boldsymbol{H}_j^{c_M}\end{bmatrix}^{\mathrm{T}}$ and $\mathrm{C_M}$ represents the j-th feature point observed by the M-th image frame in the sliding window. To marginalize the feature position error $\delta\boldsymbol{p}_j^{\mathrm{n}}$, we introduce the matrix $\boldsymbol{\Psi}$ to make $\boldsymbol{\Psi}\boldsymbol{H}_j\delta\boldsymbol{p}_j^{\mathrm{n}} = 0$. Then we have

$$\begin{aligned}\breve{\boldsymbol{r}}_j &= \boldsymbol{\Psi}\boldsymbol{r}_j \approx \boldsymbol{\Psi}\boldsymbol{H}_j^{\delta x}\delta\boldsymbol{x} + \boldsymbol{\Psi}\boldsymbol{H}_j^{c}\delta\boldsymbol{p}_j^{\mathrm{n}} + \boldsymbol{\Psi}\boldsymbol{n}_j \\ &= \breve{\boldsymbol{H}}_j\delta\boldsymbol{x} + \breve{\boldsymbol{n}}_j\end{aligned} \tag{18}$$

By stacking all the residuals and Jacobians of K feature points tracked in the current frame, we obtain the camera observation model

$$\breve{\boldsymbol{r}}_C = \breve{\boldsymbol{H}}\delta\boldsymbol{x} + \breve{\boldsymbol{n}}_C \tag{19}$$

with $\breve{\boldsymbol{r}}_C = \begin{bmatrix}\breve{\boldsymbol{r}}_1 & \breve{\boldsymbol{r}}_2 & \cdots & \breve{\boldsymbol{r}}_K\end{bmatrix}^{\mathrm{T}}$, and $\breve{\boldsymbol{H}} = \begin{bmatrix}\breve{\boldsymbol{H}}_1 & \breve{\boldsymbol{H}}_2 & \cdots & \breve{\boldsymbol{H}}_K\end{bmatrix}^{\mathrm{T}}$.

3.6 GNSS Observation Model

The GNSS Observation model provides the three-dimensional position and velocity (PV) information which can be used to generate the correction measurements by

$$\boldsymbol{r}_{\mathrm{GNSS}} = \begin{bmatrix}\boldsymbol{p}_{\mathrm{GNSS}} - \boldsymbol{p}^n \\ \boldsymbol{v}_{\mathrm{GNSS}} - \boldsymbol{v}^n\end{bmatrix} = \boldsymbol{H}_{\mathrm{GNSS}}\delta\boldsymbol{x} + \boldsymbol{n}_{\mathrm{GNSS}} \tag{20}$$

In which $\boldsymbol{H}_{\mathrm{GNSS}} = \begin{bmatrix}\boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & 0_{3\times(6\times N)} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & 0_{3\times(6\times N)}\end{bmatrix}$, and $\boldsymbol{n}_{\mathrm{GNSS}}$ is the PV noise with the variance of $\boldsymbol{\Sigma}_{\mathrm{GNSS}}$.

3.7 Sequential Kalman Filtering

Considering the triple sampling rates of sensors involved in this integration, we design a vision-aided GNSS/INS fusion architecture based on the sequential

Kalman filtering method. The first layer of Kalman filter implements the traditional GNSS/INS loosely integration computation with Kalman filter gain

$$\boldsymbol{K}_1 = \boldsymbol{P}_{k+1|k}\boldsymbol{H}_{\text{GNSS}}^{\text{T}}\left(\boldsymbol{H}_{\text{GNSS}}\boldsymbol{P}_{k+1|k}\boldsymbol{H}_{\text{GNSS}}^{\text{T}} + \boldsymbol{\Sigma}_{\text{GNSS}}\right)^{-1} \tag{21}$$

and the error state estimate

$$\delta\boldsymbol{x} = \boldsymbol{K}_1(\boldsymbol{r}_{\text{GNSS}} - \boldsymbol{r}_{\text{IMU}}) \tag{22}$$

with its covariance matrix

$$\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{K}_1\boldsymbol{H}_{\text{GNSS}})\boldsymbol{P}(\boldsymbol{I} - \boldsymbol{K}_1\boldsymbol{H}_{\text{GNSS}})^{\text{T}} + \boldsymbol{K}_1\boldsymbol{\Sigma}_{\text{GNSS}}\boldsymbol{K}_1^{\text{T}} \tag{23}$$

where $\boldsymbol{\Sigma}_{\text{GNSS}}$ is the variance of GNSS PV noise vector. It should be mentioned that the error state $\delta\boldsymbol{x}_{\text{IMU}}$ will be only used to correct the navigation results of IMU. The second layer is camera correction. Here we make use of extended Kalman filter (EKF) to process the nonlinear camera observation model with the error state equation. Denoting $\boldsymbol{\Sigma}_{\breve{\boldsymbol{n}}_C}$ as the covariance of camera observation noise vector $\breve{\boldsymbol{n}}_c$, it is easy to derive the error state estimate and its covariance with the EKF gain matrix,

$$\boldsymbol{K}_2 = \boldsymbol{P}_{k+1|k}\breve{\boldsymbol{H}}^{\text{T}}\left(\breve{\boldsymbol{H}}\boldsymbol{P}_{k+1|k}\breve{\boldsymbol{H}}^{\text{T}} + \boldsymbol{\Sigma}_{\breve{\boldsymbol{n}}_C}\right)^{-1} \tag{24}$$

and the error state is estimated by

$$\delta\boldsymbol{x} = \boldsymbol{K}_2\breve{\boldsymbol{r}}_C \tag{25}$$

with its corresponding covariance matrix

$$\boldsymbol{P}_k = \left(\boldsymbol{I} - \boldsymbol{K}_2\breve{\boldsymbol{H}}\right)\boldsymbol{P}_{k+1|k}\left(\boldsymbol{I} - \boldsymbol{K}_2\breve{\boldsymbol{H}}\right)^{\text{T}} + \boldsymbol{K}_2\boldsymbol{\Sigma}_{\breve{\boldsymbol{n}}_C}\boldsymbol{K}_2^{\text{T}} \tag{26}$$

4 Experiment

To demonstrate the accuracy and efficiency of our algorithm, the experiment is carried out on the real dataset KITTI (file name "2011_09_26_drive_0036") of a land vehicle. Figure 2 shows one image frame of testing scenario sampled with a monocular camera installed on a vehicle platform. GNSS and IMU are also mounted on the platform. To evaluate the performances of proposed methods, the GNSS high accuracy PV information is used as the reference value. Three computation schemes are employed: (1) traditional GNSS/INS loosely coupled Kalman

Fig. 2 The scene snapshot of test field

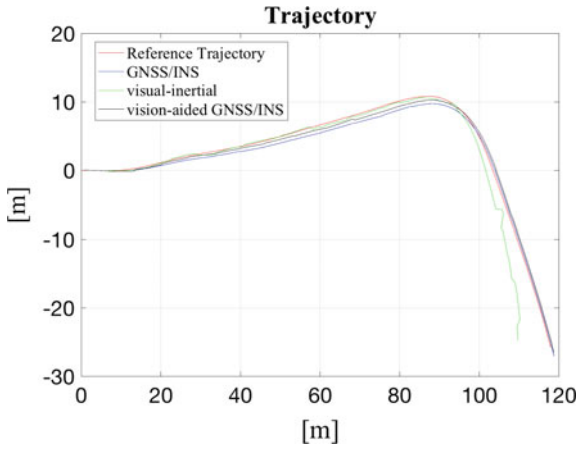


Fig. 3 The computed motion trajectories and referenced trajectory

filtering algorithm; (2) Visual-inertial odometry Kalman filtering algorithm; (3) Our proposed vision-aided GNSS/INS mixed sequential Kalman filtering algorithm.

The motion trajectories computed by three algorithms are drawn in Fig. 3. The reference trajectory is drawn with red line. Generally, all the three algorithms generate the consistent results. Particularly, the performance of GNSS/INS is very close to the vision-aided GNSS/INS algorithm. This exhibits the importance of GNSS in the overall integrated navigation system. The trajectory errors are also computed and shown in Fig. 4.

To be more clear, the X, Y and Z components are all computed 5–7. The statistics of root mean-square errors (RMSE) are also computed and presented in Table 1. It needs to be pointed that the X, Y and Z components are defined in the coordinate frame of vehicle's forward, left and up (FLU) at the first recorded epoch. Therefore, to conveniently compare all the algorithms aforementioned, the computed navigation results in the frame of LBH and ENU, also the referenced GNSS information in WGS-84 are all transformed into the FLU frame. Evidently, seen

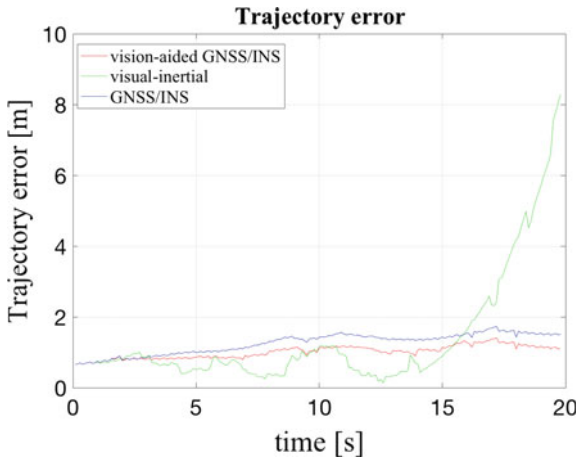


Fig. 4 The trajectory errors of VO-GNSS/INS, VIO and GNSS/INS

Table 1 RMSE of VO-GNSS/INS, VIO and GNSS/INS

RMSE	X	Y	Z
GNSS/INS	0.9178	0.9014	0.5432
VO-GNSS/INS	0.9161	0.4629	0.5403
VIO	2.1054	0.3874	0.1513

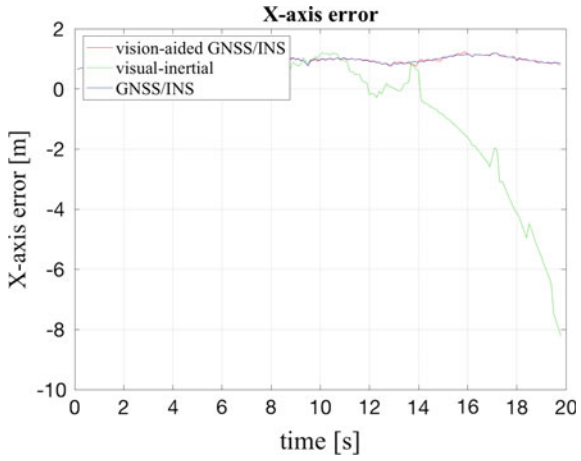


Fig. 5 The X errors of VO-GNSS/INS (red), VIO (green) and GNSS/INS (blue)

from these figures and table, in the direction of X, GNSS/INS has the comparative performance with vision-aided GNSS/INS. In Y and Z directions, with the aid of vision sensor, the system performance is significantly improved compared with other two schemes (Figs. 5, 6 and 7).

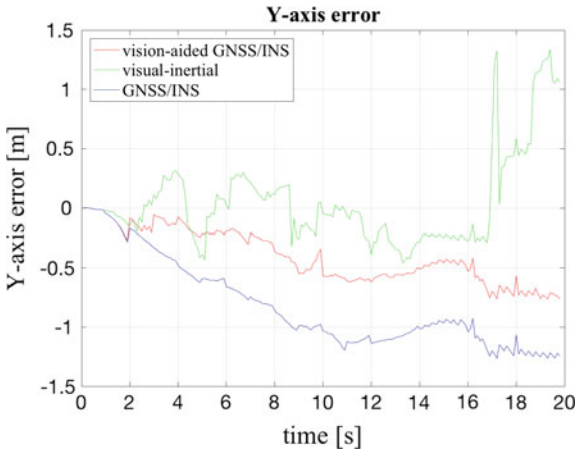


Fig. 6 The Y errors of VO-GNSS/INS (red), VIO (green) and GNSS/INS (blue)

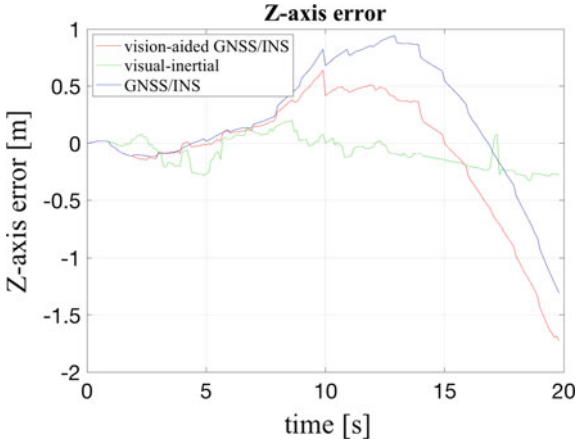


Fig. 7 The Z errors of VO-GNSS/INS (red), VIO (green) and GNSS/INS (blue)

5 Conclusion

In this paper, we propose an approach of vision-aided GNSS/INS fusion based on the sequential Kalman filter structure which is able to achieve higher precision and more reliable navigation results. Compared with traditional GNSS/INS navigation, our proposed algorithm shows its efficiency and performs more reliably and accurately, especially in the fields where GNSS signal is unavailable. Furthermore, GNSS still plays a crucial role to restrain the accumulate errors not only from inertial sensors but also the vision sensors. This makes the system endure a much longer duration in contrast to VIO.

Acknowledgements This work is supported by the National Natural Science Funds of China (No. 41604025 and 41704029), Sichuan Province Science and Technology Project (No. 2016GZ0062), the State Key Laboratory of Geodesy and Earth's Dynamics (Institute of Geodesy and Geophysics, CAS) Grant No. SKLGED2018-3-2-E and the Key Laboratory of Precise Engineering and Industry Surveying of National Administration of Surveying, Mapping and Geoinformation (PF2015-11).

References

1. Niu X, Li Y, Zhang Q, Cheng Y, Shi C (2012) Observability analysis of non-holonomic constraints for land-vehicle navigation systems. GPS Solutions 11(1):80–88
2. Bak, T. (2001). Vision-GPS fusion for guidance of an autonomous vehicle in row crops. In: Proceedings of 14th international technical meeting of the satellite division of the Institute of Navigation (ION GPS 2001), Institute of Navigation, Manassas, VA, pp 423–429
3. Zhou Z, Yang L, Li Y (2013) An adaptive dual Kalman filtering algorithm for Locata/GPS/INS integrated navigation. In: China satellite navigation conference, vol 245, pp 527–541
4. Klein G, Murray D (2007) Parallel tracking and mapping for small AR workspaces. In: IEEE and ACM international symposium on mixed and augmented reality, pp 1–10
5. Mur-Artal R, Tardos JD (2017) Orb-slam2: an open-source slam system for monocular, stereo, and rgb-d cameras. IEEE Trans Rob **33**(5), 1255–1262
6. Caruso D, Engel J, Cremers D (2015) Large-scale direct slam for omnidirectional cameras. In: IEEE/RSJ international conference on intelligent robots and systems, vol 25, no 4, pp 141–148
7. Engel J, Sturm J, Cremers D (2014) Semi-dense visual odometry for a monocular camera. In: IEEE international conference on computer vision, pp 1449–1456
8. Forster C, Carlone L, Dellaert F, Scaramuzza D (2015) On-manifold preintegration for real-time visual-inertial odometry. IEEE Trans Rob 33(1):1–21
9. Mur-Artal R, Tardós JD (2017) Visual-inertial monocular slam with map reuse. IEEE Robot Autom Lett 2(2):796–803
10. Mourikis AI, Roumeliotis SI (2007) A multi-state constraint Kalman filter for vision-aided inertial navigation. In: IEEE international conference on robotics and automation, vol 22, pp 3565–3572
11. Li M, Mourikis A (2013) High-precision, consistent EKF-based visual-inertial odometry. Int J Robot Res 32(6):690–711
12. Bloesch M, Burri M, Omari S, Hutter M, Siegwart R (2017) Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback. Int J Robot Res 36 (10):1053–1072
13. Stefan L, Simon L, Michael B (2015) Keyframe-based visual-inertial odometry using nonlinear optimization. Int J Robot Res 34(3):314–334
14. Leutenegger S, Furgale P, Rabaud V, Siegwart R (2013). Keyframe-based visual-inertial slam using nonlinear optimization. In: Robotics: science and systems, vol 34, pp 789–795
15. Yang Z, Shen S (2017) Monocular visual-inertial state estimation with online initialization and camera–IMU extrinsic calibration. IEEE Trans Autom Sci Eng 14(1):39–51
16. Clement LE, Peretroukhin V, Lambert J, Kellye J (2015) The Battle for filter supremacy: a comparative study of the multi-state constraint Kalman filter and the sliding window filter. In: Computer and robot vision, IEEE, pp 23–30

Real-Time Dead Reckoning and Mapping Approach Based on Three-Dimensional Point Cloud

Shuaixin Li, Guangyun Li, Yanglin Zhou, Li Wang and Jingyang Fu

Abstract Light detection and ranging equipment (LiDAR) has wide applications in the field of mobile surveying, autonomous driving, unmanned aerial vehicle and military, etc., since its abilities of fast three-dimensional environmental information acquisition, robustness to variable illumination conditions as well as wide measurement range. The main difficulty for localization and mapping based on LiDAR is the registration between successive point clouds, which is caused by the continuous motion and undetermined trajectory of LiDAR. A real-time dead reckoning and three-dimensional environment mapping method based on 3D-LiDAR is presented in this paper. The proposed system adopts a divide and parallel method, which performs high frequency pose estimation and low frequency mapping on parallel threads to ensure the real-time performance. In point preprocess section, the efficient elimination of point cloud distortion is processed and feature points are extracted. In dead reckoning section, point cloud registration employs generalized iterative closest point (GICP) algorithm which increases accuracy of registration by taking local covariance information of each point into consideration. In environmental mapping section, the multi-channel GICP (MCGICP) algorithm which adds point cloud intensity information into GICP framework is applied to align local point cloud with the global point cloud map. The robustness and convergence are promoted, since the fusion of additional point in-formation. Finally, the performance of the proposed approach is evaluated by the KITTI dataset. The experimental results show that our proposed simultaneous dead reckoning and 3D environment mapping method based on 3D laser scanner is an effective and feasible on-line solution which achieves high accuracy under 1% in all kinds of

S. Li (✉) · G. Li · Y. Zhou · L. Wang · J. Fu
School of Geospatial Information, Information Engineering University, Zhengzhou, Henan, China
e-mail: lsx_navigation@sina.com

S. Li
Key Laboratory of Precise Engineering and Industry Surveying of National Administration of Surveying, Mapping and Geoinformation, Wuhan, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_55

environment scenarios. Moreover, the ability of self-localization and real-time mapping is improved dramatically in those environments with poor geometric feature patterns.

Keywords LiDAR · Point cloud · Iterative closest point · Simultaneous localization and mapping

1 Introduction

The development of Global Navigation Satellite System (GNSS), inertial sensor technology and Artificial Intelligence has promoted location service related industry to become one of the fastest growing domains [1]. Location information is a significant constituent of the big data required by space research, containing more than just location. There is a wide range of applications in the mobile mapping, autonomous driving as well as unmanned aerial vehicle for which three dimensional is an essential technology. The most fundamental task of these applications is obtaining environmental information which can be used for the autonomous vehicles' decision as well as the representation of real-world scenarios. Precisely determined 6 degree of freedom (DOF) trajectory of sensor is the precondition of mapping, since sensors collecting data in the local coordinates of themselves. GNSS related approach is the most accurate and efficient positioning method in outdoor environments where satellite signals are not blocked or disturbed. However, satellite signals are always unavailable not only in indoor environments but also city canyons and outer space where the mission of self-positioning and environment mapping is urgent for autonomous vehicles to drive or land. In indoor environments, there is not a best positioning solution. A score of localization solution have been presented, such as visual odometry [2, 3] (VO), navigation based on Inertial Navigation System (INS) and positioning method based on Ultra-Wide Band [4] (UWB) radio. All these approaches have its own limitation, for example, VO and INS suffer from trajectory drift over time, and UWB need enough signal base stations. So, it is hard to achieve indoor/outdoor integration positioning as the limited application scenarios and weakness of these approaches. The advent of 3D-LiDAR changed the situation. It has attracted considerable attention and become a hot topic in this research area due to its advanced properties of high measurement resolution, anti-interference and robustness to illumination conditions.

Simultaneous localization and mapping, which is a core technology in terms of indoor/outdoor integrative mapping based on LiDAR. Two kinds of approaches can be distinguished according to different processing scheme, namely incremental and global ones [5]. Incremental approaches used a limited series of successive point clouds to estimate the relative motion of vehicle, or sometimes employed different kinds of filtering algorithms that derived by recursive Bayes Filter [6] (BF), like Extended Kalman Filter (EKF) and Particle Filter (PF) to perform accumulated trajectory estimation and simultaneous mapping. Although this kind of method

owns good real-time performance, it suffers from accumulated errors and trajectory drift. Moreover, only with a limited history of measurements, the incremental map obtained by the method lacks of consistency. Global optimization [7] employs a batch optimization utilizing a set of non-linear optimized equations which represent the constrains between pose and measurements. If we regard all LiDAR poses and landmarks as nodes and all measurements as edges linking pose nodes and landmarks nodes, a node-edge graph can be obtained. Then the location estimation problem can be transformed into a graph optimization problem, which is usually solved by bundle adjustment. Comparison with incremental approaches, batch optimization increases the accuracy of positioning as well as the consistency of global map. However, it is obvious that the method will cost a mass of computing resources restricting its application in some scenes where real-time performance is highly required.

As for 2D SLAM solution, Gmapping [8] proposed by Grisetti and Hector SLAM [9] proposed by Kohlbrecher is the state of art. Santos and Vincent deeply compared these two methods in terms of positioning accuracy, mapping accuracy as well as operation efficiency [10, 11]. Gmapping applies a novelty mutual distribution calculating approach exploiting both odometry data and 2D laser scans to compute precise proposal distribution of vehicle. Then an improved Rao-Blackwellized Particle Filter (RB-PF) is employed to estimate the orientation and planar coordinates of LiDAR. Moreover, a selective resampling method is also presented to decrease the complexity of computation. This mutual distribution calculation approach requires that Gmapping can only works well when odometry data is available. The utilization of 2D LiDAR means it can only achieve reasonable result in planar area. Hector SLAM is a fast online SLAM solution based on single LiDAR. The scheme fully exploits the high update rate and the low distance measurement noise from 2D LiDARs. The 2D pose estimation is based on optimization of the alignment of beam endpoints with the current global map. The endpoints are projected in the global map and the occupancy probabilities can be estimated using bilinear interpolation scheme. To avoid the potential risk of getting stuck in local minima, a multi-resolution map representation approach which is similar to image pyramid is presented to perform a coarse-to-refined location estimation. However, it might have problems when scan rate is relative low. The issue of cross-section drawn of environment mapping result is something that all 2D SLAM solutions have in common. On the one hand, localization based on 2D map can only used in structured indoor area, since the alignment of laser scan is rather unreliable in unstructured outdoor space. On the other hand, 2D map is quite inadequate for practical applications.

There are three main thoughts of 3D range sensing in mobile mapping implementation. Firstly, LiDAR can be movable in many ways, like periodic nodding and continuous spinning. Secondly, many LiDARs which not on a common scan plane can be attached together on a platform and rely on motion of the platform to produce 3D point clouds [12]. Thirdly, the most direct solution to 3D mapping is using an 3D-LiDAR sensor. Actuated 2D-LiDAR methods which usually incorporate some mechanical devices are common and easy to realize. Bosse presented a

2D-LiDAR/INS combined approach [13], in which a 2D range scanner is attached with a passive linkage mechanism, such as a spring. As the platform moving, the acceleration or vibration of platform will leverage an irregular motion, thereby extending the device's field of view outside of its original scanning plane and obtaining 3D point clouds. However, disadvantages of this method are obvious and serious. Unpredictable LiDAR motion causes a heavy dependence on attitude and position sensors, like Inertial Measurement Unit (IMU) and the accuracy of range measurements strongly relies on the instrumental precision. Furthermore, the mechanism of this design makes it hard to apply in flat terrain. LiDAR Odometry and Mapping (LOAM) [14], proposed by Zhang, achieves 3D mapping using a continuous spinning or periodical rotation 2D LiDAR representing the state of art. The core idea of this solution is the division of the complex problem of SLAM, which performs a global optimization after a long period of measuring, into two algorithms. The one is in charge of high-frequency odometry but also with low fidelity to location estimation of LiDAR. The other one runs at a lower frequency for location refinement and map update. Running the two algorithms on parallel threads allows LiDAR self-localization and map creation in real-time. But when registration between successive point clouds, this solution is more likely to divergency employing a point-to-line and point-to-plane integrated nonlinear optimization algorithm. Moosmann proposed Velodyne SLAM using a 3D range scanner [15]. A distorted point cloud offset algorithm causing by continuous motion of LiDAR is deeply discussed with details as well. Although this approach can obtain good trajectory and mapping result in some structural scenarios, it suffers from mismatch of point cloud in some similar and unstructured places, such as highway, since the registration of point clouds only depending on geometrical features instead of taking color or reflectance information into account.

How to align two point clouds obtained in different location to the same coordinate frame is the most significant technology in LiDAR odometry and mapping. Iterative Closest Point (ICP) is one of the most well-known registration methods for points in 3D space. Pomerleau et al. [16] presented an assessment system for all kinds of variation of ICP algorithm and compared many typical ICP variants including point-to-plane, random sampling and differential methods etc. A probabilistic frame work is proposed in [17] to describe the registration of point clouds. This so called Generalized Iterative Closest Point (GICP) algorithm explain standard ICP algorithm from the statistic perspective and revealed the essential difference between standard ICP and point-to-plane ICP algorithm. Moreover, a plane-to-plane ICP implementation is also raised based on GICP frame. The experimental results showed that the performance of this method was superior in comparison with other ICP variants, since it fully uses the geometrical characteristics in both source point cloud and target point cloud. To incorporated color and intensity information into the GICP frame, Servos proposed a Multi-Channel Generalized Iterative Closest Point (MCGICP) [18] algorithm. It added color and intensity information into the GICP frame by projecting additional information to the plane of local points pattern. MCGICP increases the accuracy and precision of

registration in unstructured scenarios by adding additional information and extending the closest distance between two points from 3D space to multi-dimensional space.

To improve current solution to 3D simultaneous localization and mapping, a 3D-LiDAR dead reckoning and 3D environmental mapping approach is presented. The approach refers to the division of SLAM problem and runs the different frequency sectors on parallel threads to guarantee an online performance of the system. Firstly, a preprocess procedure is proposed to extract surf features from input point clouds and to roughly de-skew point cloud motion distortion. Then, a plane-to-plane GICP registration algorithm is applied at the same frequency with LiDAR output to coarsely estimate the relative motion between point clouds. Finally, a registration of local point cloud and global map using MCGICP is operated at a lower frequency, after which the location of LiDAR will be refined meanwhile the global map will be updated. The rest of paper is organized as follow. In Sect. 2, we restate the problem and define the symbols and varieties. In addition, the overview of our system is listed and we briefly introduce each sector of the system. In Sect. 3, we illustrate how to extract features and ensure the distribution of features is uniform. Distortion de-skews method for spinning 3D-LiDAR is also introduced in detail. In Sect. 4, we introduce the GICP frame work and discuss the odometry of LiDAR using GICP algorithm. In Sect. 5, we focus on the location estimate refinement and map update by alignment of laser scans and global map. In Sect. 6, we analyze the accuracy, robustness and time cost using KITTI dataset. Finally, conclusions are given and improvement of the system is discussed.

2 System Description

2.1 Notations and Task Description

The indoor/outdoor simultaneous localization and 3D environmental mapping system presented in the paper utilize a 3D range scanner sensor, such as Velodyne HDL64E, and HESAI Pandar 40. We assume that the angular and linear velocities of the LiDAR are smooth and continuous over time, without any abrupt changes.

As a convention in this paper, we use right uppercase superscription to indicate the coordinate system for points, and use right subscription to indicate source coordinate system while right superscription to indicate target coordinate system for transformation matrix. A complete sweep defines as $\{P\}_k$ of whole 360° scanning from the start time t_k^S to the end time t_k^E. The LiDAR coordinate system L defines as a right-hand coordinate with x-axis pointing to the right, y-axis pointing to the left and z-axis pointing to the forward. The global coordinate system W is the initial LiDAR coordinate frame, namely $W = L_0$. We denote a point i in the k-th sweep point cloud $\{P\}_k$ as $\boldsymbol{X}_i^{L_k}$. $\boldsymbol{T}^{L_k}$ denotes a 3×4 transformation matrix from L to W at the end of each full sweep and $\{T\}_k$ consisting of a series of $\boldsymbol{T}^{L_k}$ denotes the 6DOF

trajectory. We denote $\boldsymbol{X}_i^W$ as a point i in the global map $\{M\}_k$ at the end of time t_k^E. With the notations and assumptions defined before, the problem on LiDAR self-localization and 3D environment mapping can be paraphrased as: given a sequence of scanning point clouds $\{P\}_k, k \in \mathbb{Z}^+$, estimate complete 6DOF trajectory $\{T\}_k$ and 3D map of scenario $\{M\}_k$.

2.2 System Overview

Figure 1 shows a diagram of the software system. The whole system framework can be divided into three sections: *a.* point cloud preprocess and features extraction, *b.* LiDAR dead reckoning, and *c.* 3D mapping.

a. **Point clouds preprocess and features extraction**

When the system receives a point cloud $\{\ddot{P}\}_k$ with distortion causing by LiDAR motion, the first work to do with it is de-skew. Then the surf feature patterns can be extracted utilizing principal component analysis (PCA) method. Details will be introduced in Sect. 3.

b. **LiDAR dead reckoning**

Relative motion can be estimated by aligning the last time fully de-skewed point cloud $\{P\}_{k-1}$ with current initial de-skewed point cloud $\{\dot{P}\}_k$. We apply GICP algorithm to perform a coarse-to-refined registration. Therefore, we can estimate the relative motion $\boldsymbol{T}_{L_{k-1}}^{L_k}$ during the period of time $[t_{k-1}, t_k]$, and the current LiDAR pose $\dot{\boldsymbol{T}}^{L_k}$ in world coordinate. After that, the de-distortion process can be operated

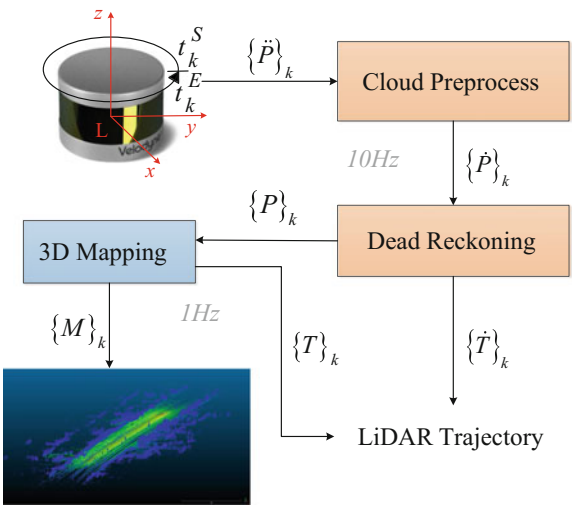


Fig. 1 Block diagram of the system

again to produce a fully de-skewed point cloud $\{P\}_k$ at the present. The output frequency of LiDAR odometry is exactly as the same as that of data receiving. Details will be introduced in Sect. 4.

c. **Environment 3D mapping**

This sector is in charge of the global map update as well as LiDAR pose optimization. To increase the robustness of the system to unstructured environment, we employ MCGICP which adds additional intensity information into GICP to achieve the registration between local point cloud $\{P\}_k$ and global map $\{M\}_{k-1}$. However, 3D mapping will cost amount of computation resources. To keep the performance of real-time, 3D mapping runs at a lower frequency referring to the system framework presented in [14]. Details will be introduced in Sect. 5.

3 Point Clouds Preprocess and Features Extraction

3.1 Elimination of Motion Distortion

3D LiDAR receives laser signal emitting from a spinning laser emitter which spins continuously around the central axis of sensor. Therefore, the exact time t_{k-1}^i of every point scanned by laser scanner is different as the sensor movement during scan acquisition. Namely, the exact corresponding coordinates over which every point acquired are slightly different. Actually, LiDAR coordinates at the end of $(k-1)$-th scan t_{k-1}^E is equal to the start of k-th scan t_k^S. If we indistinguishably regard them as point cloud obtained in the same time, the point cloud must contain distortion which will increase with the increase of platform velocity.

Since the assumption of motion with constant linear and angular velocity, we can linearly predict current LiDAR velocity based on last time velocity. This characteristic gives us an opportunity to efficiently eliminate the distortion of point cloud. We interpolate LiDAR pose $\dot{\boldsymbol{T}}_i^{L_k}$:

$$\dot{\boldsymbol{T}}_i^{L_k} = (\boldsymbol{T}^{L_{k-2}} \ominus \boldsymbol{T}^{L_{k-1}}) \cdot \frac{t_k^i}{t_k^E - t_k^S} \tag{1}$$

where $\ominus$ denotes the inverse Lie algebraic operation of transformation in the special Euclidean group. After this, the receiving distorted point cloud $\left\{\ddot{P}\right\}_k$ can be initially de-skewed by unifying all points to the estimated current LiDAR coordinate frame $\dot{\boldsymbol{T}}^{L_k}$. Later when odometry finish, the more precise LiDAR transformation will be used for furtherly de-skewing. Finally, current point cloud can be unified into the same coordinate frame $\boldsymbol{T}^{L_k}$:

(a) Distorted point cloud (b) Undistorted point cloud

Fig. 2 Comparison of de-skewed point clouds

$$\boldsymbol{T}_i^{L_k} = (\boldsymbol{T}^{L_{k-1}} \ominus \boldsymbol{T}^{L_k}) \cdot \frac{t_k^i}{t_k^E - t_k^S} \tag{2}$$

Figure 2 shows a point cloud before and after elimination of motion distortion. Theoretically, this operation should be employed once when the pose is updated each time after nonlinear optimization iteration. But the experimental results in [15] proved that it is sufficient to de-skew only twice: before and after registration.

3.2 Feature Points Extraction

Feature points extraction has two main functions. On the one hand, it can downsize the scale of points cloud and get rid of points with redundant information to keep system's efficiency and real-time performance. On the other hand, it also contributes to the precision of registration, because single noisy points can be rejected. PCA algorithm is the most common method to analyze the geometrical structure of points cluster. It illustrates that the eigen value of covariance of local point cluster will denote the geometrical configuration of point pattern. We compute the covariance $\boldsymbol{M}$ of a specific points and its k nearest neighbors, and decompose the matrix using sing value decompose (SVD) algorithm to obtain the matrix's eigen value $\boldsymbol{V}$ and eigen vectors $\boldsymbol{E}$. If there are two values that much greater than the other one, this cluster of points locates on a plane and the eigen vector corresponding to the smallest eigen value is the normal vector of this plane.

Zhang and Singh [14] indicates that the distribution of point cloud in space can influence the precision of point cloud alignment, which is that the more homogeneous points distribute, the more precise the alignment will be. To ensure feature points homogeneously distribute in the space, we present an improved feature

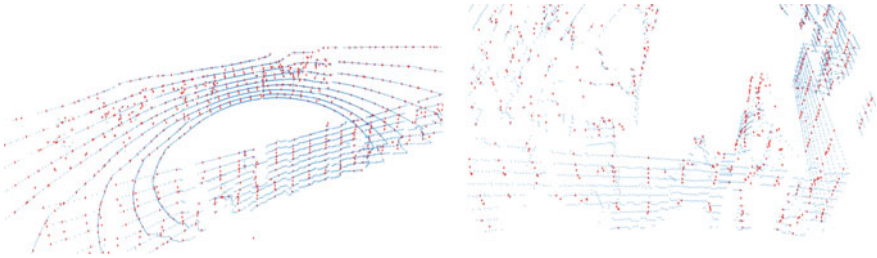


Fig. 3 Homogeneously distributed planar feature points

extraction method to optimize the extracted points. 3D mesh segmentation is applied to point cloud and we traverse each cube $\boldsymbol{V}(3)$ to find feature points. N points that the minimum eigen value of their covariance less than a specific threshold α are extracted from each cube $\boldsymbol{V}(3)$. Figure 3 shows the extracted feature points. Blue points are receiving point cloud, and red points are extracted homogeneous feature points.

4 LiDAR Dead Reckoning

4.1 GICP Framework

ICP probabilistic framework describes the standard ICP algorithm and its variants. We assume that points in $\{A\}$ and $\{B\}$ are one to one correspondence. Since the existence of measurement noises, each point $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ in $\{A\}$ and $\{B\}$ can be represented as a Gaussian distribution $\boldsymbol{a}_i \sim \mathcal{N}(\hat{\boldsymbol{a}}_i, \boldsymbol{\Sigma}_i^A)$ and $\boldsymbol{b}_i \sim \mathcal{N}(\hat{\boldsymbol{b}}_i, \boldsymbol{\Sigma}_i^B)$ respectively. Assume that the transformation between two point cloud is $\boldsymbol{T}^*$:

$$\hat{\boldsymbol{b}}_i = \boldsymbol{T}^* \hat{\boldsymbol{a}}_i. \tag{3}$$

The error equations of (3) can be rewritten as:

$$\boldsymbol{d}_i = \boldsymbol{T}\boldsymbol{a}_i - \boldsymbol{b}_i = f(\boldsymbol{T}), \tag{4}$$

where $\boldsymbol{T}$ denotes the estimation of $\boldsymbol{T}^*$. Therefore, we can obtain the probabilistic distribution of random variety $\boldsymbol{d}_i$: $\boldsymbol{d}_i \sim \mathcal{N}(0, \boldsymbol{\Sigma}_i^B + \boldsymbol{T}\boldsymbol{\Sigma}_i^A\boldsymbol{T}^{\mathrm{T}})$. And $\boldsymbol{T}$ can be computed using the maximum likelihood estimation (MLE) iteratively:

$$\boldsymbol{T} = \underset{\boldsymbol{T}}{\operatorname{argmax}} \prod_i \mathrm{p}(\boldsymbol{d}_i^T) = \underset{\boldsymbol{T}}{\operatorname{argmax}} \sum_i \ln[\mathrm{p}(\boldsymbol{d}_i^T)], \tag{5}$$

Equation (5) can be simplified as:

$$\boldsymbol{T} = \underset{\boldsymbol{T}}{\operatorname{argmin}} \sum_i \boldsymbol{d}_i^T (\boldsymbol{\Sigma}_i^B + \boldsymbol{T}\boldsymbol{\Sigma}_i^A \boldsymbol{T}^{\mathrm{T}})^{-1} \boldsymbol{d}_i. \tag{6}$$

Here, Eq. (6) can represents all kinds of ICP algorithm by setting appropriate covariances $\boldsymbol{\Sigma}_i^A$ and $\boldsymbol{\Sigma}_i^B$.

In practice, we cannot obtain two laser points reflected from the exact same point since laser points are discrete, which means that there is no point-to-point constraint between two points. Therefore, plane-to-plane ICP variant has more practical significance, because the only kind of constraints is that those along the surface normal. Based on the principle of PCA, we can know that the smallest eigen value of covariance of a points cluster represents the surface normal, while other two greater value denotes directions along the surface. Hence, we can construct a kernel covariance $\boldsymbol{\Sigma} = \mathrm{diag}(1, 1, \delta)$ to express point cloud $\{A\}$ and $\{B\}$:

$$\boldsymbol{\Sigma}_i^A = \boldsymbol{R}_i^A \boldsymbol{\Sigma} \left(\boldsymbol{R}_i^A\right)^{\mathrm{T}}, \quad \boldsymbol{\Sigma}_i^B = \boldsymbol{R}_i^B \boldsymbol{\Sigma} \left(\boldsymbol{R}_i^B\right)^{\mathrm{T}}, \tag{7}$$

where $\boldsymbol{R}_i$ denotes the rotation matrix from δ to the normal of point surface. The covariance $\boldsymbol{\Sigma}_i$ can be approximated by the covariance of local points cluster $\bar{\boldsymbol{\Sigma}}_i$. We search 20 nearest neighbors i through KD tree, which can achieve this swiftly and exactly. Then, we use SVD algorithm to estimate the rotation matrix $\boldsymbol{R}_i$:

$$\bar{\boldsymbol{\Sigma}}_i = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}} = \boldsymbol{R}_i \boldsymbol{S} (\boldsymbol{R}_i)^{\mathrm{T}}. \tag{8}$$

Hence, to compute covariance $\boldsymbol{\Sigma}_i$, the only thing we need to do is replacing $\boldsymbol{S}$ with kernel covariance $\boldsymbol{\Sigma} = \mathrm{diag}(1, 1, \delta)$.

4.2 Relative Motion Estimation

The circumstance that iterations stuck into local optimum during ICP can be avoided to the maximum, because plane-to-plane ICP variant takes normal information of points into consideration which accords with the practice better.

In Fig. 4, black dashed line ellipses represent covariances of each point, red dashed line links a point and its closest correspondence. The registration depends more on those corresponding pairs which have the same covariance information, since points with the same covariance will be assigned higher weight, while points with different covariances only be assigned lower weight. Therefore, the precision of registration can be guaranteed.

Given point cloud $\{P\}_{k-1}$ at t_{k-1} and initially de-skewed point cloud $\left\{\dot{P}\right\}_k$ at t_k, LiDAR relative motion $\boldsymbol{T}_{L_{k-1}}^{L_k}$ during $[t_{k-1}, t_k]$ can be computed using plane-to-plane GICP algorithm, and the closest correspondence can be searched using KD tree.

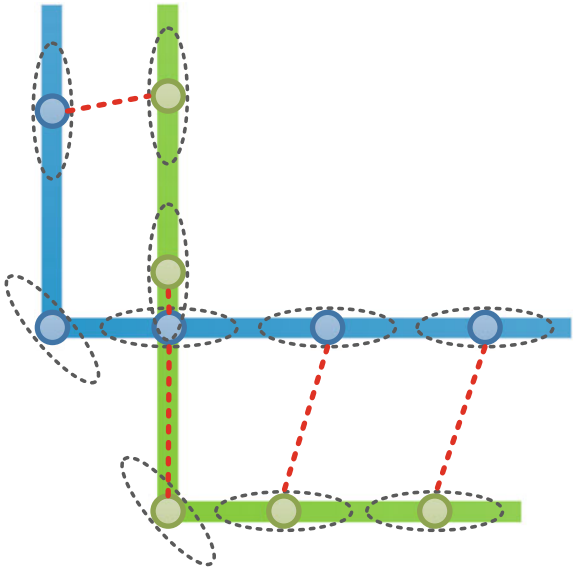


Fig. 4 Diagram of plane to plane alignment

Then, we can use Gauss-Newton non-linear optimized algorithm to solve this problem. The solution to Eq. (6) is:

$$\boldsymbol{T}_{L_{k-1}}^{L_k} + (\boldsymbol{J}^T\boldsymbol{J})^{-1}\boldsymbol{J}^T\boldsymbol{d} \rightarrow \boldsymbol{T}_{L_{k-1}}^{L_k}. \tag{9}$$

The Jaccobian matrix $\boldsymbol{J}$ denotes the derivate of coefficient matrix of Eq. (4) over transformation matrix $\boldsymbol{T}_{L_{k-1}}^{L_k}$, namely, $\boldsymbol{J} = \partial f / \partial \boldsymbol{T}_{L_{k-1}}^{L_k}$. We update $\boldsymbol{T}_{L_{k-1}}^{L_k}$ after each iteration until $\boldsymbol{d}_i \rightarrow 0$, or the iteration number come to the maximum threshold. Finally, we update current LiDAR pose as the transformation matrix representing over the global coordinate system:

$$\dot{T}^{L_k} = \boldsymbol{T}_{L_{k-1}}^{L_k} \cdot T^{L_{k-1}}. \tag{10}$$

5 Three-Dimensional Environment Mapping

5.1 MCGICP Algorithm

MCGICP algorithm incorporates additional information, like intensity and color, to extend GICP framework. Namely, it is an extended GICP in a higher dimensional space. The increased problem space not only solves the degeneracy problem but also improves accuracy, convergence, and robustness of the scan registration results without significantly increasing computational complexity.

We introduce $\boldsymbol{p}_i = [\boldsymbol{p}_i^p, p_i^d]^{\mathrm{T}}$ to denote selected points, where $\boldsymbol{p}_i^p \in \mathbb{R}^3$ represents 3D location information, and $p_i^d \in \mathbb{R}$ represents additional information.

a. **Covariance calculation**

We assume $\boldsymbol{q} = [\boldsymbol{q}^p, q^d]^{\mathrm{T}}$ is an index point, and $\bar{\boldsymbol{\Sigma}}_i$ is the local covariance of $\boldsymbol{q}$ and its k nearest neighbor points $\boldsymbol{l}_i = [\boldsymbol{l}_i^p, l_i^d]^{\mathrm{T}}$, $i \in [1, \ldots, k]$. We then project these neighbor points $\boldsymbol{l}_i^P$ onto the span surface of index point $\boldsymbol{q}$ along the normal according to Eq. (8). The projections are denoted by $\boldsymbol{z}_i = [\boldsymbol{z}_i^p, z_i^d]^{\mathrm{T}}, \boldsymbol{z}_i \in \mathbb{R}^3$:

$$\boldsymbol{z}_i^p = \begin{bmatrix} \boldsymbol{U}_1^{\mathrm{T}} \\ \boldsymbol{U}_2^{\mathrm{T}} \end{bmatrix} \boldsymbol{l}_i^p, \quad z_i^d = l_i^d, \tag{11}$$

where $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ are the first and second columns of the SVD matrix $\boldsymbol{U}$. After transform, the new population covariance $\boldsymbol{\Sigma}_w \in \mathbb{R}^{2\times 2}$ of the projection is the diagonal matrix of the largest two singular values of $\boldsymbol{S}$.

In this paper, we only consider additional intensity information of point cloud, and introduce standard Gauss distribution $\mathcal{N}(z_i^d, \sigma)$ to express intensity information. And we define weight of projection as λ_i:

$$\lambda_i = \exp\left(-\frac{1}{2}(z_i^d - q^d)^{\mathrm{T}} \sigma^{-1} (z_i^d - q^d)\right). \tag{12}$$

Then, the intensity information can be incorporated into point information by computing the means and covariance of projections according to their weights.

$$\boldsymbol{\mu}^p = \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i \boldsymbol{z}_i^p \tag{13}$$

$$\boldsymbol{\Sigma}_d = \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i (\boldsymbol{z}_i^p - \boldsymbol{\mu}^p)(\boldsymbol{z}_i^p - \boldsymbol{\mu}^p)^{\mathrm{T}} \tag{14}$$

Theoretically, $\boldsymbol{\mu}^p$ and $\boldsymbol{\Sigma}_d$ models the uncertainty of the intensity information along the surface locally and reflect the similarity between index point $\boldsymbol{q}$ and its neighbors $\boldsymbol{l}_i$, but it can be biased if the original nearest samples are biased. To compensate for this biased distribution, we normalize the population covariance as:

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma}_w^{-\frac{1}{2}} \boldsymbol{\Sigma}_d \boldsymbol{\Sigma}_w^{-\frac{1}{2}}. \tag{15}$$

Now we have the population covariance which not only contains constraints along the surface plane, but contains intensity information of this points cluster. Finally, to use this information in the GICP framework, we replace the top left corner of original covariance as $\boldsymbol{\Omega}$:

$$\boldsymbol{\Sigma}_i = \mathrm{diag}[\,\boldsymbol{\Omega} \quad \delta\,]. \tag{16}$$

Then the relative transform relationship $\boldsymbol{T}$ can be estimated replace (6) with (16).

b. **Nearest point search**

Due to the incorporation of intensity information, the nearest point search scheme based on KD tree in three-dimensional Euclidean space should be extended into a higher four-dimensional space. In other words, the nearest point search of a specific point $[q_x^p, q_y^p, q_z^p, q^d]^{\mathrm{T}}$ based on points' distance in four-dimensional space:

$$d = \sqrt{(a_x^p - b_x^p) + (a_y^p - b_y^p) + (a_z^p - b_z^p) + (a^d - b^d)}. \tag{17}$$

To increase the accuracy of nearest points research for 4D laser points, we introduce a weighting vector $\boldsymbol{w} = [\varpi_x, \varpi_y, \varpi_z, \varpi_d]^{\mathrm{T}}$ to weight the positional and intensity information.

5.2 *Registration Between Local and Global Point Clouds*

We run mapping sector at a lower frequency to allow map creation in real-time, which only updates the global map at 2 Hz while the odometry sector updates at 10 Hz. After odometry, we obtain undistorted point cloud $\{P\}_k$ in k-th sweep, and the LiDAR transformation $\dot{T}^{L_k}$ at t_k.

Figure 5 illustrates the process of mapping when local point cloud $\{P\}_k$ is aligned with global map $\{M\}_{k-1}$. Firstly, we transform $\{P\}_k$ to $\{Q\}_k$ which is in the global coordinate system using the estimate of LiDAR transformation $\dot{T}^{L_k}$. Affected by drift, $\{Q\}_k$ and $\{M\}_{k-1}$ cannot be aligned exactly. We estimated the transform error ΔT^{L_k} by aligning $\{Q\}_k$ with $\{M\}_{k-1}$ using MCGICP algorithm, so that both global map and LiDAR pose can be updated according to the transform error.

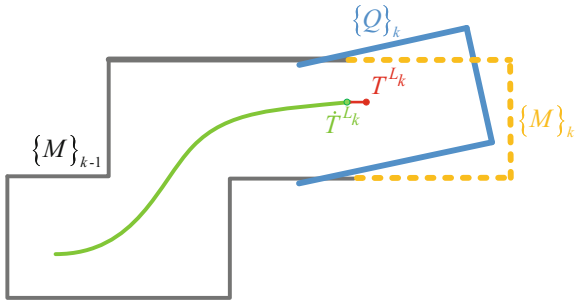


Fig. 5 Illustration of mapping process

To avoid indexing among a huge number of points in the global map, we use 3D mesh segmentation strategy again. We save the global point cloud map in a $W \times H \times D$ cube according to their different locations. The valid sub-cubes can be determined using the cube indexed by $\dot{T}^{L_k}$. Only points in the LiDAR's field of view will be used for mapping optimization.

As we discussed before, the registration only based on geometrical feature points, such as line and surf, is unreliable in unstructured environment where geometrical feature points are sparse. To fully exploit multi-channel information of point cloud and increase the divergency as well as robustness of registration, we employ the alignment using MCGICP algorithm to estimate the transform error ΔT^{L_k} and update global map as well as LiDAR trajectory:

$$\begin{aligned} T^{L_k} &= \Delta T^{L_k} \cdot \dot{T}^{L_k} \\ \{M\}_k &= \{M\}_{k-1} + T^{L_k}\{P\}_k \end{aligned} \tag{18}$$

6 Performance Evaluation and Analysis

6.1 Experiment Introduction

The performance of our proposed solution is evaluated by well-known KITTI dataset [19, 20] which provides both testing and benchmark dataset. The data contains Velodyne HDL-64E 3D laser scanning data and precise trajectory obtained by INS/GPS combined navigation system whose kinematic positional accuracy can reach 5 cm combining GPS/GLONASS/IMU with real-time kinematic (RTK) carrier phase difference technique. We use laser data to perform dead reckoning and 3D mapping and use the trajectory provided by INS/GPS combined system as the ground truth to evaluate the accuracy of positioning. Velodyne HDL-64E LiDAR outputs 64 laser scans in total and range from 0–360° in horizontal direction, 2–24.8° in vertical direction. The output frequency of LiDAR is 10 Hz and the valid measurement range is 120 m. The maximum running speed of the experimental platform is 85 km/h. There are three main experimental scenarios including highway, urban and countryside. The software is programmed by C++ language, running on a Linux PC with Intel i7 2.5 GHz CPU. In Fig. 6, the dead reckoning result is on the left viewport, while 3D mapping result is on the right.

Evaluation based on the error of the trajectory end-point is misleading, as it depends strongly on the point in time where the error has been made. We use the evaluation method proposed in [19], which computes the average of rotation or translation errors in a fixed distance or speed range. The functions below are evaluation functions of the trajectory length and platform speed respectively:

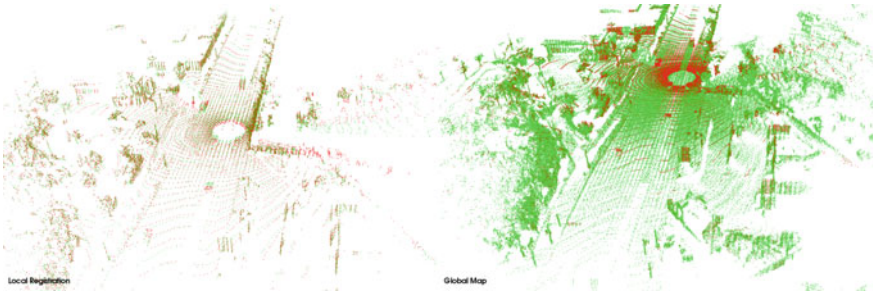


Fig. 6 Diagram of system operation

$$E_{rotation}(\mathcal{P}) = \frac{1}{|\mathcal{P}|}\sum_{(i,j)\in\mathcal{P}} \angle\left[(\hat{\boldsymbol{p}}_i \ominus \hat{\boldsymbol{p}}_j)\ominus(\boldsymbol{p}_i \ominus \boldsymbol{p}_j)\right]. \tag{19}$$

$$E_{translation}(\mathcal{P}) = \frac{1}{|\mathcal{P}|}\sum_{(i,j)\in\mathcal{P}} \left\|(\hat{\boldsymbol{p}}_i \ominus \hat{\boldsymbol{p}}_j)\ominus(\boldsymbol{p}_i \ominus \boldsymbol{p}_j)\right\|_2. \tag{20}$$

where $\mathcal{P}$ is a set of points, $\hat{\boldsymbol{p}}$ and $\boldsymbol{p}$ are estimated and true LiDAR poses respectively, i, j are indices of points, $\angle[\cdot]$ denotes the rotation vector, and $\|\cdot\|_2$ denotes the distance between two points.

6.2 Accuracy Analysis

According to Eqs. (19) and (20), we set the fixed distance and speed as 100 m and 5 km/h. Table 1 lists the average error over full length for different datasets. Figures 7, 8 and 9 illustrate the mean positional error ratios, mean angular error ratios in each trajectory length and platform speed zones for each dataset. Figures 10, 11 and 12 shows the experimental scenes and corresponding 3D point cloud maps obtained by our method. Positional error ratios for all datasets are no

Table 1 Average error over full length for different datasets

Data no.	Data configuration		Mean positional error (%)	Mean operating time (s)
	Length (m)	Environment		
#1	834	Highway	0.64	0.075
#2	296	Countryside	0.96	0.083
#3	417	Urban	0.26	0.081

Column 4 shows mean positional errors ratio among all distance zones (%). Column 5 shows the mean operating time (s) of two successive sets of points

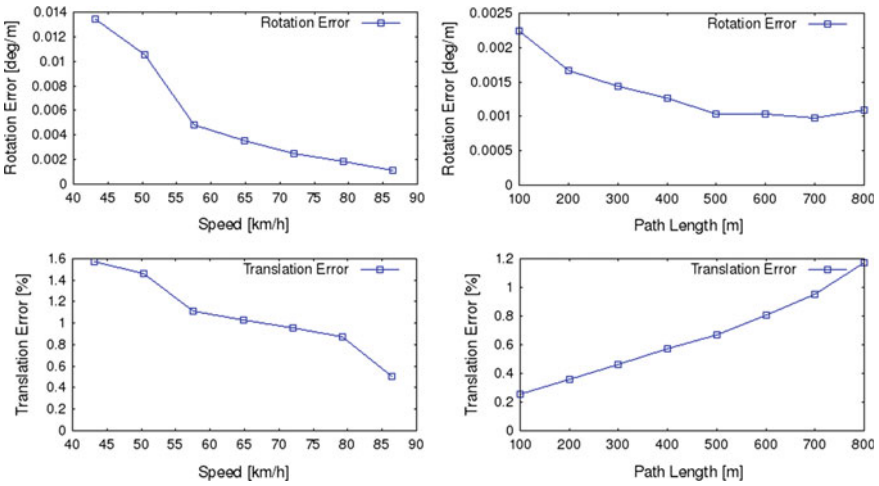


Fig. 7 Rotation and translation errors over different path length and vehicle speed in #1 highway environment

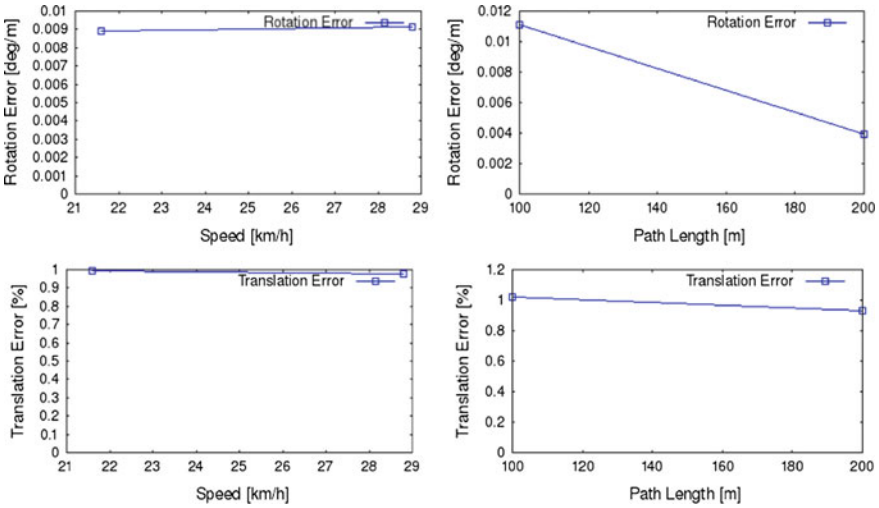


Fig. 8 Rotation and translation errors over different path length and vehicle speed in #2 countryside environment

more than 1.0%, and operating time cost less than 0.1 s, which proves that our method is a feasible on-line dead reckoning and 3D mapping method.

Table 1 shows that our proposed method can achieve high accuracy in all three experimental areas. By comparing the data in Table 1, we can find that the positioning accuracy is the highest in urban city environment, since there are sufficient

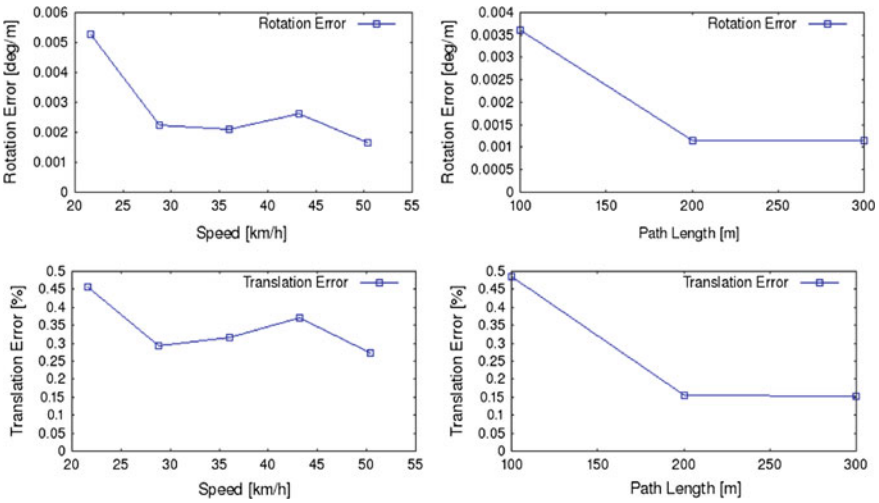
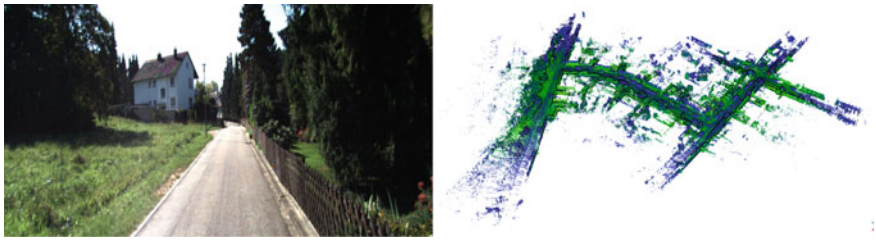


Fig. 9 Rotation and translation errors over different path length and vehicle speed in #3 urban environment

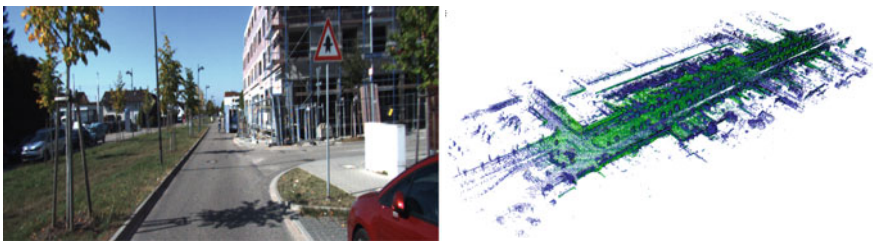
(a) highway scene (b) highway point cloud

Fig. 10 Scene and sample results of #1 highway dataset



(a) countryside scene (b) countryside point cloud

Fig. 11 Scene and sample results of #2 countryside dataset



(a) urban scene (b) urban point cloud

Fig. 12 Scene and sample results of #3 urban dataset

structured features in the experimental area which insures the accuracy of point cloud registration. In highway environment, accuracy suffers from a little loss, because the experimental area is open, where feature points are sparse and rare. As for countryside, points of weeds and trees are lack of geometrical characteristics, which also limits the accuracy of location estimation.

Analysis based on Figs. 7, 8 and 9:

#1 Highway dataset: Platform makes a turn first, and then goes straight at high speed along the highway. Angular error is sensitive to the speed, which means that high turning speed makes the angular error high. However, the angular error declines as the linear speed increases when it keeps forward. Consequently, the error of angle is high at the beginning of the trajectory, and keeps steady at around 0.001°/m later. The relationship between positional error and speed is similar to that between angular error and speed, however, positional error rises linearly with path length extending because of the earlier accumulated angular error.

#2 Countryside dataset: Platform runs at a low speed, thus both angular and positional error remain almost unchanged with speed increases. The initial angular error is relative high, because the data start from turning around. In addition, due to the short length of whole trajectory, there is not obviously drift can be checked and positional error stay at approximately 1%.

#3 Urban dataset: Platform drives to-and-fro along the avenue, which makes error of angle and position decrease as the driving distance extending. During the latter trajectory, experimental platform actually locates itself in a known environment, so that accuracy of positioning stays at a fairly high level.

7 Conclusion

This paper presents a high-accuracy real-time LiDAR dead reckoning and 3D environment mapping method to solve the problem on simultaneous localization and 3D mapping. We apply the divide-and-conquer scheme to keep the real-time performance. The elimination of point cloud distortion and the extraction of surf points using PCA are performed. GICP and MCGICP algorithm are employed in

odometry and mapping section respectively. The proposed method is evaluated by KITTI dataset. The experimental results show that our proposed method based on 3D laser scanner is an effective and feasible on-line solution which achieves high accuracy (under 1%) in all kinds of environment scenarios.

From experimental result, we can also find that the angular accuracy declines dramatically with the turning speed increasing. Consequently, if pose measurements from IMU can be incorporated into the system to provide an initial guess for point cloud registration, the accuracy and efficiency of ICP can be greatly enhanced. Although it has many advantages using 3D laser scanner to conduct self-positioning and 3D mapping, like robustness to illumination condition and audio-visual 3D map, it still suffers from drift over time. To solve this problem, the loop closure and global optimization should be employed in next step.

Acknowledgements This work is supported by "the Open Foundation of Key Laboratory of Precise Engineering and Industry Surveying of National Administration of Surveying, Mapping and Geoinformation" (Grant No. PF2017-1)

References

1. Deng Z (2016) The present situation and development of navigation and location services. Satell Appl 2016(2):41–45
2. Li Z, Wang J, Toth C (2017) A geometric correspondence feature based-mismatch removal in vision based-mapping and navigation. Photogram Eng Remote Sens 83(10):693–704
3. Li S, Li G, Zhou Y et al (2017) An improved monocular simultaneous localization and mapping solution. Chin J Sci Instrum 38(11):2849–2857
4. Li S, Li G, Wang L et al (2017) A three-dimensional robust ridge estimation positioning method for UWB in a complex environment. Adv Space Res 60(12):2763–2775. https://doi.org/10.1016/j.asr.2017.10.040
5. Thrun S, Burgard W, Fox D (2006) Probabilistic robotics. MIT Press, Cambridge, MA, p 2006
6. Durrant-Whyte H, Bailey T (2006). Simultaneous localization and mapping: part I. IEEE Robot Autom Mag 13(2):99–180
7. Kümmerle R, Grisetti G, Strasdat H et al (2011) G2o: a general framework for graph optimization. In: IEEE international conference on robotics and automation, vol 7, pp 3607–3613
8. Grisetti G, Stachniss C, Burgard W (2007) Improved techniques for grid mapping with Rao-Blackwellized particle filters. IEEE Trans Rob 23(1):34–46
9. Kohlbrecher S, Stryk OV, Meyer J, Klingauf U (2011) A flexible and scalable SLAM system with full 3D motion estimation. In: IEEE international symposium on safety, security, and rescue robotics, pp 155–160
10. Santos JM, Portugal D, Rui PR (2014) An evaluation of 2D SLAM techniques available in robot operating system. In: IEEE international symposium on safety, security, and rescue robotics, pp 1–6
11. Vincent R, Limketkai B, Eriksen M (2010) Comparison of indoor robot localization techniques in the absence of GPS. In: Proceedings of SPIE, vol 7664, pp 76641Z–76641Z-5
12. Zlot R, Bosse M (2014) Efficient large-scale three-dimensional mobile mapping for underground mines. J Field Rob 31(5):731–752

13. Bosse M, Zlot R, Flick P (2012) Zebedee: design of a spring-mounted 3-D range sensor with application to mobile mapping. IEEE Trans Rob 28(5):1104–1119
14. Zhang J, Singh S (2016) Low-drift and real-time lidar odometry and mapping. Auton Rob 41 (2):1–16
15. Moosmann F, Stiller C (2011) Velodyne SLAM. In: Intelligent vehicles symposium, vol 32, pp 393–398
16. Pomerleau F, Colas F, Siegwart R et al (2013) Comparing ICP variants on real-world data sets. Auton Rob 34(3):133–148
17. Segal A, Haehnel D, Thrun S (2009) Generalized-ICP. In: Robotics: science and systems, Seattle, USA, pp 26–27
18. Servos J, Waslander SL (2017) Multi-channel generalized-ICP: a robust framework for multi-channel scan registration. Robot Auton Syst 87:247–257
19. Geiger A (2012) Are we ready for autonomous driving? The KITTI vision benchmark suite. In: IEEE conference on computer vision and pattern recognition, vol 157, pp 3354–3361, IEEE Computer Society
20. Geiger A, Lenz P, Stiller C et al (2013) Vision meets robotics: the KITTI dataset. Int J Robot Res 32(11):1231–1237

Method of Smartphone Navigation Heading Compensation Based on Gravimeter

Shijie Zeng, Qinghua Zeng, Qian Meng, Jianye Liu and Ruizhi Chen

Abstract Smartphone based pedestrian inertial navigation system usually uses the Pedestrian Dead Reckoning (PDR) algorithm to achieve navigation and positioning. The traditional PDR algorithm requires inertial devices to remain a stationary position relative to the human body, however, in the field of smart pedestrian navigation, smartphone can't always keep a stationary posture when it is being used. When the way of using smartphone has been changed, traditional PDR algorithm will give the wrong heading estimation which is caused by the change of smartphone's attitude. Therefore, in order to satisfy the navigation demand for smartphone pedestrian inertial positioning especially when the using way changed, a Heading Compensation with Gravity Assisted (HCGA) method is proposed in this paper. This method will judge the change of smartphone's attitude and compensate heading based on data from gravity sensor. The result of experiment shows that using HCGA method, it can be distinguished whether the change of heading caused by the attitude change automatically when the change of heading is detected, then compensate the result of heading. In this way, the influence of navigation solution caused by different handheld modes will be reduced.

Keywords Gravity assisted · Smartphone navigation · Heading compensation

S. Zeng · Q. Zeng (✉) · Q. Meng · J. Liu
Navigation Research Center, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
e-mail: zengqh@nuaa.edu.cn

S. Zeng
e-mail: zengshijie@nuaa.edu.cn

R. Chen
State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_56

1 Introduction

At present, in the field of pedestrian navigation, the research on outdoor navigation mainly uses GNSS (Global Navigation Satellite System, GNSS) or GNSS/INS (Inertial Navigation System, INS) integrated navigation [1]. While in the field of indoor navigation research. Research on pedestrian navigation technology mainly focuses on the following two categories: one is based on radio frequency devices such as indoor UWB, ZigBee, wireless networks, WiFi and Bluetooth [2]; another is based on installing the inertial sensors at the human leg, waist, shoulder and other parts [3], by researching algorithm such as Zero Velocity Update (ZUPT), step recognition, and etc. to achieve pure inertial pedestrian navigation. The two types of navigation methods mentioned above both have advantages and disadvantages. The former needs to dispose radio frequency devices as beacons before using, which increases the navigation cost, the latter needs additional navigation equipment on the human body, makes portability lacking in portability defective. With the popularization and continuous improvement of smartphones, it integrate various navigation sensors inside, including three-axis accelerometer, three-axis gravimeter, three-axis gyroscope, and magnetometer and so on, make smartphone-based pedestrian navigation positioning become an important new branch in the field of navigation [4].

Pedestrian mobile navigation as a new hot spot in the field of navigation in recent years, domestic and foreign universities and research institutes have conducted in-depth research in this area, and made some research results in this field. Researchers at the Navigation Research Center of Nanjing University of Aeronautics and Astronautics designed and implemented a seamless indoor and outdoor navigation system based on a smartphone platform [5]. Radu and Marina [6] at the University of Edinburgh proposed a combined navigation algorithm, HiMLoc, which uses a smartphone as a navigation device in combination with a WiFi fingerprinting system [7]. Wonho Kang and others from Korea Advanced Institute of Science and Technology put forward and set up a smartphone-based SmartPDR navigation system, which can use smartphones to complete PDR navigation solution. However, the disadvantage of this system is that smartphones are required to be kept in a fixed way and it can't be used where is too much movement [8]. Therefore, there is an urgent need for a navigation algorithm that can automatically determine the posture of the smartphone and does not cause a misalignment of heading due to the change of posture of the smartphone so as to improve the applicability and accuracy of the navigation in a satellite-free environment.

This paper proposes a new heading compensation algorithm based only on the smartphone's own MEMS sensor. By collecting and analyzing the output information of the MEMS sensor, the algorithm determines the handset's usage pattern and the current handset gesture, and compensates for the heading errors caused by

changing the handset's holding mode when using the handset navigation. Moreover, the algorithm realizes the automatic judgment of smartphone normal mode, landscape mode and call mode, as well as the compensation of the heading when three kinds of gestures change mutually.

2 Principle of Pedestrian Dead Reckoning

When using smartphone for navigation in Non-GNSS signals areas, inertial positioning technology with stronger autonomy is always chosen. Traditional Strapdown Inertial Navigation System (SINS) continuously integrates through the output of accelerometers and gyroscope from smartphone. In this way system can get the three-dimensional information of position, speed and heading after being switched to navigation coordinate system. However, with the MEMS inertial device's low accuracy and poor stability, continuously integration will make the position result divergence quickly. And PDR navigation method can calculate the next time's walking track, location and other information, through the measurement and statistics of the step-count, step-length and direction of pedestrians at a certain time without quadratic integral to inertial navigation data, so PDR algorithm has higher positioning accuracy than SINS, and it can achieve high accuracy navigation at low cost with its the characteristics of complete autonomy and flexibility. Generally speaking, PDR algorithm is comprised of three important parts [9]:

(1) Steps Detection;
(2) Step Size Estimation;
(3) Heading Estimation.

Figure 1 shows the principle of PDR algorithm. In Fig. 1, E means East, N means North. The recursion formula of PDR algorithm is formula (1), S_k and θ_k means walking distance and heading angle from position (E_{k-1}, N_{k-1}) to position (E_k, N_k) [10]:

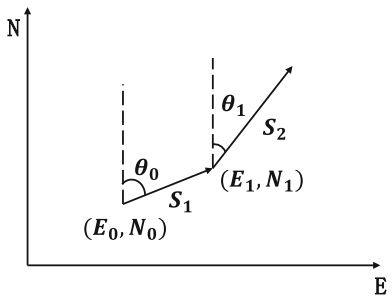


Fig. 1 Principle of Pedestrian dead reckoning

$$E_k = E_0 + \sum_{i=0}^{k-1} S_i \sin\theta_i$$
$$N_k = N_0 + \sum_{i=0}^{k-1} S_i \cos\theta_i \quad (1)$$

3 Heading Compensation with Gravity Assisted

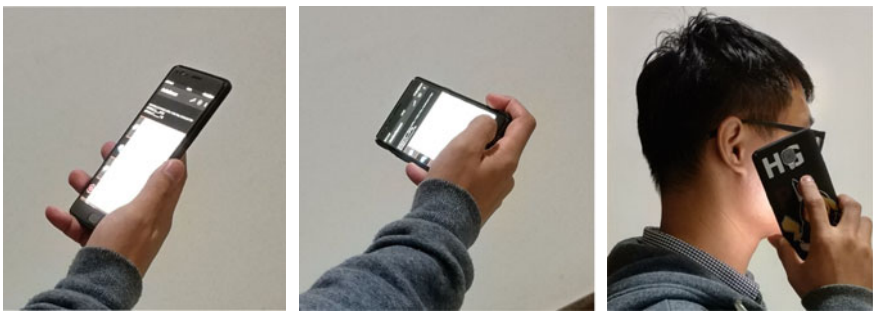
In daily life, there are many ways to use smartphone, such as shown in Fig. 2. There will be three ways to use the phone: normal mode, landscape mode and call mode.

When the postures of using smartphone change to each other from the pose mentioned in Fig. 2, the smartphone will have a larger position changes, and the internal sensors of the smartphone will record and output the acceleration and angular acceleration in the process of attitude transformation. It will make the traditional navigation algorithm to believe that the pedestrian's heading has changed mistakenly, and giving wrong navigation positioning results. Based on the analysis of pedestrian usage of smartphones and the corresponding sensor output data, a Heading Compensation with Gravity Assisted (HCGA) algorithm is proposed to compensate for the heading of the smartphone in this paper. The HCGA algorithm is divided into two parts, the attitude determination method based on gravity ball regional and heading compensation process.

Gravity data detected by gravity of smartphone has the following features:

$$\vec{g} = \vec{g}_x + \vec{g}_y + \vec{g}_z \quad (2)$$

where $\vec{g}$ means gravity acceleration resultant vector, $\vec{g}_x, \vec{g}_y, \vec{g}_z$ are the gravitational acceleration vector of the each axes output from the phone. If the collected three-axis gravimeter data is projected into a three-dimensional coordinate system



(a) normal mode (b) landscape mode (c) call mode

Fig. 2 Modes of phone holding

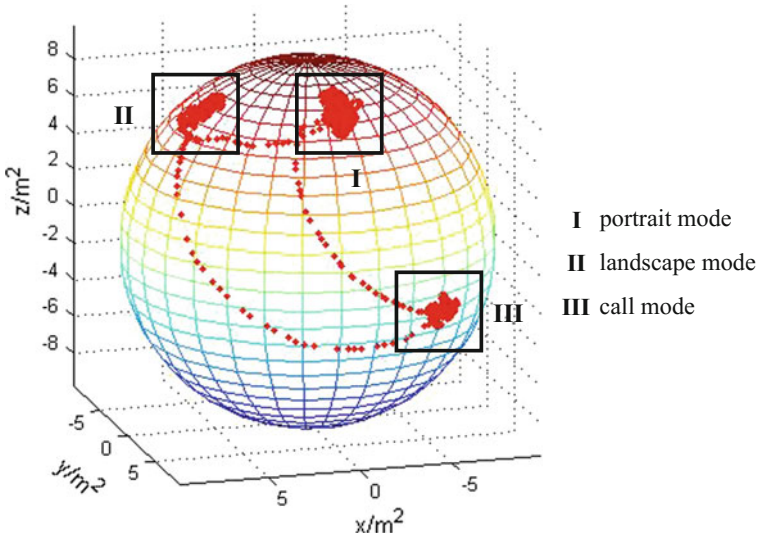


Fig. 3 Spherical point cloud

and build a sphere whose radius equals local gravity accelerometer, then we can get a spherical point cloud like the Fig. 3. In the Fig. 3, area I II III represent the projection of gravity data in 3D coordinate system of normal mode, landscape mode and call mode respectively. From the Fig. 3 it can be found that the output data of the smartphone in different attitude are projected in different areas, that is, by calculating the azimuth and elevation angle of the gravity meter data in the spherical coordinate system we can distinguish the phone's current posture.

Through formula (3) the azimuth angle and elevation angle of the three-axis gravity data in the spherical coordinate system can be calculated, where g_x,g_y,g_z means three axis's output of gravimeter, α,β means azimuth and elevation angles. The projection area getting by calculating the range of azimuth and elevation angles is using to Judge the phone's posture.

$$\begin{cases} \alpha = \arctan(\frac{g_y}{g_x}) \\ \beta = \arccos(\frac{g_z}{\sqrt{g_x^2 + g_y^2 + g_z^2}}) \end{cases} \tag{3}$$

In order to facilitate the calculation, experiments in this paper is using cell phone with right hand. According to the different ways of using smartphones, there are three different postures, and the judgment thresholds of those postures are as follows, where J means posture flags, 1, 2, 3, represent the normal mode, landscape mode and call mode respectively.

$$J = \begin{cases} 1 & -12^\circ < \alpha < 108^\circ, 14^\circ < \beta < 59^\circ \\ 2 & -12^\circ < \alpha < 108^\circ, 14^\circ < \beta < 59^\circ \\ 3 & 108^\circ < \alpha < 132^\circ, 78^\circ < \beta < 100^\circ \end{cases} \quad (4)$$

Based on the research on the relationship between pedestrian using smartphone and smartphone gesture, this paper presents a method to compensate the heading based on smartphone attitude. The specific process of this method is divided into the following four steps:

(1) Obtain the data of the smartphone's MEMS device, and judge the pace moment;
(2) Real-time monitoring the changes of the heading between two paces;
(3) According to the size of change-value to determine whether compare the pace of the phone at before and after steps.
(4) Contrast the results before and after the phone gesture, according to the comparison results to correct the heading.

4 Experiment and Analysis

4.1 Accuracy Analysis of Smartphone's Inside Sensor

In this paper, the experiment is done with ONEPLUS A3000 smartphone, whose inside sensors are consumer-grade MEMS sensors, limited by the cost and process, its performance such as zero bias and bias stability can not be guaranteed, so the accuracy of the accelerometer and gyroscope needs to be analyzed before experimental verification.

In this paper, the Allan variance method is used to analyze Angle Random Walk (ARW) and Bias Instability (BI) of mobile phone inertial sensors. The results of ARW and BI analysis are shown in Tables 1 and 2, and N and B in the table represent the ARW and the BI respectively.

Table 1 Results of Gyroscope Allan variance

Gyroscope	N (°/h$^{1/2}$)	B (°/h)
X-axis	0.4482	23.6184
Y-axis	0.5655	9.6157
Z-axis	0.42377	10.2627

Table 2 Results of accelerometer Allan variance

Accelerometer	N (m/s$^{3/2}$)	B (m/s^2)
X-axis	2.0883e−05	2.53e−4
Y-axis	2.017e−05	4.33e−4
Z-axis	2.2924e−05	5.68e−4i

4.2 Walking Experiment Analysis

The gesture-based attitude determination and heading compensation method proposed in this paper is a smartphone navigation method aimed at pedestrian walking state, so the researchers conducted indoor walking experiments in the buildings of Nanjing University of Aeronautics and Astronautics Automated institute, and the way of using the smartphone is changed during walking so as to increase the different using gestures of the smartphone for comparison experiments. The use of smartphones in Fig. 2 in three modes. In this experiment, the normal mode, the landscape mode, the call mode, and the normal mode are successively used when walking, and the judgment result of the posture of the cell phone is shown in Fig. 4, and the three posture judgments is successfully during walking, and the accuracy is 100%.

The experimental navigation data collected using the algorithm HCGA compensation algorithm and the traditional PDR algorithm for comparative analysis, the experimental results shown in Fig. 5, the green line represents the actual walking trajectory, the blue line represents the results of the compensation algorithm navigation results, the red line represents the navigation result obtained by the traditional algorithm. As shown in Fig. 5, using the compensated algorithm, it is possible to correct the wrong heading caused by changing the attitude of the handset and return the navigation heading to normal.

Figure 6 shows the heading values using both algorithms. Among them, the dotted gray line represents the turning moment of the experimenter during the walk, while the light-blue solid line represents the moment when the posture of the smartphone changes. The blue line is the heading curve obtained by the compensation navigation algorithm. The solid line marked with "X" represents the traditional algorithm heading curve, the green solid line for the real walking heading

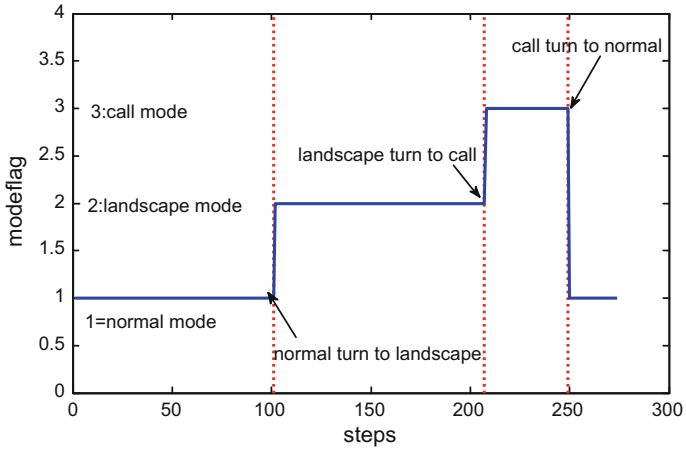


Fig. 4 Judgment of phone using mode

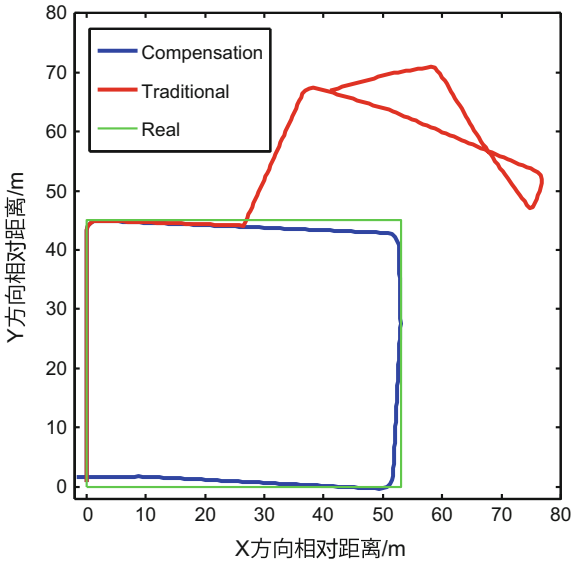


Fig. 5 Walking experiment comparison of two algorithms

curve. From Fig. 6, we can see that both the compensation algorithm and the traditional algorithm can determine the change of heading when the pedestrian is turning. However, when the posture of the smartphone changes, the traditional algorithm still considers the pedestrian to make a turn and thus outputs Wrong heading, and the compensation algorithm through the use of the attitude of the judge, the heading of the compensation, so that the heading to maintain normal.

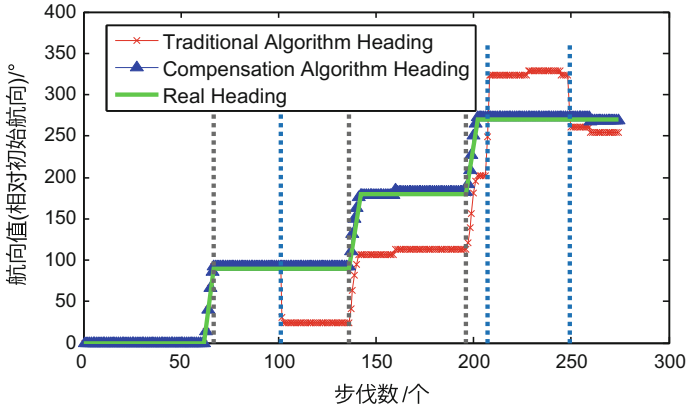


Fig. 6 Heading comparison of two algorithms

5 Summary

When using smartphone for pedestrian navigation, the changes of using way will make phone posture changes, while traditional navigation algorithm will keep the wrong heading accumulation caused by gesture change, and give a wrong navigation results. The HCGA method proposed in this paper compensates the heading by judging the posture of the smartphone and reduces the influence of different hand-held modes on navigation solution.

From the experimental results, the HCGA algorithm dynamically compensates the navigation heading according to the change of hand-held modes, eliminates navigation error caused by the smartphone posture change. The unconstrained navigation of the smartphone has been realized, and improve the accuracy of phone navigation's heading.

At present, there are three kinds of judgments about smartphone usage modes in this paper, and the threshold analysis process is a statistical data after many experiments conducted by researchers. In order to improve the analysis efficiency, machine learning will be started in subsequent research, establishing relevant judgments model, make analysis of cell phone usage modes more efficient.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (Grant No. 61533008, 61374115, 61603181), the Fundamental Research Funds for the Central Universities (No. NJ20170005 & No. NJ20170010), the Open Fund of State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (Grant No. 17P02) & the Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

1. Meng Q, Liu J, Zeng Q, Feng S, Chen R (2017) Neumann-Hoffman code evasion and stripping method for BeiDou software-defined receiver. J Navig 70(1):101–119
2. Ruiz ARJ, Granja FS, Honorato JCP, Rosas JIG (2012) Accurate pedestrian indoor navigation by tightly coupling foot-mounted IMU and RFID measurements. IEEE Trans Instrum Meas 61(1):178–189
3. Qian W, Zeng Q, Wang J et al (2015) Pedestrian navigation method based on kinematic mechanism of human lower limb. J Chin Inertial Technol 23(1):24–28
4. Chen R, Guinness R (2014) Geospacial computing in mobile devices. Artech House, Norwood, Massachusetts
5. Zeng Q, Wang J, Meng Q, Zhang X, Zeng S (2017) Seamless pedestrian navigation methodology optimized for indoor/outdoor detection. IEEE Sensors J 18(1):363–374
6. Radu V, Marina MK (2013) HiMLoc: indoor smartphone localization via activity aware Pedestrian dead reckoning with selective crowd sourced WiFi fingerprinting. In: International conference on indoor positioning and indoor navigation IEEE, pp 1–10
7. Hartmann F, Rifat D, Stork W (2016) Hybrid indoor pedestrian navigation combining an INS and a spatial non-uniform UWB-network. In: International conference on information fusion IEEE, pp 549–556

8. Kang W, Han Y (2015) SmartPDR: smartphone-based pedestrian dead reckoning for indoor localization. IEEE Sens J 15(5):2906–2916
9. Lan H, Yu C, El-Sheimy N (2015) An integrated PDR/GNSS pedestrian navigation system. Chin Satell Navig Conf 342:677–690
10. Randell C, Djiallis C, Muller H (2003) Personal position measurement using dead reckoning. In: Proceedings on IEEE international symposium on wearable computers, vol 28, pp 166–173

Research on Robust PDR Algorithm Based on Smart Phone

Jian Kuan, Xingeng Chen and XiaoJi Niu

Abstract In recent years, the demand for location based service has promoted the rapid development of pedestrian navigation and positioning technology. Among them, Pedestrian Dead Reckoning (PDR) based on inertial sensor can greatly improve the robustness of indoor and outdoor pedestrian navigation system due to its environment-friendly. This paper designs a robust PDR algorithm based on smart phone. This paper constructs multi-level constraints based on INS algorithm and establishing observation equation needed by EKF using the mutually complementary relationship among sensors, such as gyroscope accelerometer and magnetometer, and human dynamic model. The cumulative error of the system is less than 2% of the traveling distance when using different intelligent terminals, holding on hand both in indoor and outdoor test.

Keywords INS · Multi-level constraints information · Robust PDR

1 Introduction

In recent years, the indoor positioning technology of consumer-grade has become a focus and difficult research at home and abroad. Among them, the algorithm of Pedestrian Dead Reckoning (PDR) based on mobile terminals is widely used in indoor and outdoor pedestrian navigation. PDR is a relative positioning method. The basic principle of PDR is to sense the pedestrian movement by using the built-in inertial sensor of the intelligent terminal, and to obtain the pedestrian's current position by detecting the pedestrian gait and calculating the pedestrian's heading. PDR estimated position error will gradually accumulate as the time goes by, and can only provide reliable positioning results in a short time. Therefore, other positioning sources (such as Wi-Fi and iBeacon) can be integrated with PDR to improve the accuracy and stability. At present, gait detection and step length

J. Kuan (✉) · X. Chen · X. Niu
GNSS Research Center, Wuhan University, Wuhan, China
e-mail: xjniu@whu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_57

estimation models are almost perfect, and low-precision heading has become a major factor affecting the accuracy and robustness of PDR positioning systems [1]. Among them, [2] introduced Attitude and heading reference system (AHRS) to obtain better heading estimation through fusing gyroscope, accelerometer and magnetometer observation. However, the update frequency of PDR algorithm is 1–2 Hz, and it is a disadvantage when merging other positioning sources [3]. A pedestrian navigation algorithm based on INS is designed. The navigation information of high frequency is output while guaranteeing the positioning accuracy by using constraint information such as step model, non-holonomic constraints (NHC), zero angular rate update (ZARU) and zero velocity update (ZUPT). However, this method does not have any heading observation information, so the heading error quickly diverges over time. On the basis of [3], this paper constructs multi-level constraints based on INS algorithm and establishing observation equation using the mutually complementary relationship among sensors, such as gyroscope, accelerometer and magnetometer, and human dynamic model. Real time estimation of heading error and gyro bias improves the accuracy and robustness of the navigation system. The constraints are employed once quasi-static, quasi-static magnetic fields, cornering or other constraints are detected, so as to ensure the accuracy of observations and error compensation [4]. Finally, the improved PDR algorithm is analyzed and verified by experimental results.

2 The Design of Robust PDR Algorithm

2.1 The Overall Design of the Algorithm

The traditional PDR algorithm is simple, and the constraint information used is limited, resulting in low precision and short usable time. Therefore, this paper designs an improved PDR algorithm based on multi-level constraint information to improve the accuracy and robustness of MEMS sensor-based positioning system [5]. Compared with the traditional PDR algorithm, it can provide abundant real-time information (such as 3D velocity). The algorithm is shown in Fig. 1.

The multi-level constraint information considered in this paper includes ZUPT, ZARU, gravity vector, NHC, magnetic field vector and robust PDR algorithm as shown in the figure. The algorithm execution process is: first, input the accelerometer and gyroscope signals into INS mechanization [6], at the same time, conduct gait detection, step estimation and current state detection according to the MEMS sensor information. Then, the observation equation is constructed when the constraints are satisfied, and the error state is calculated according to the extended Kalman filter. Finally, the current navigation state is corrected, which improve the accuracy and robustness of the positioning system.

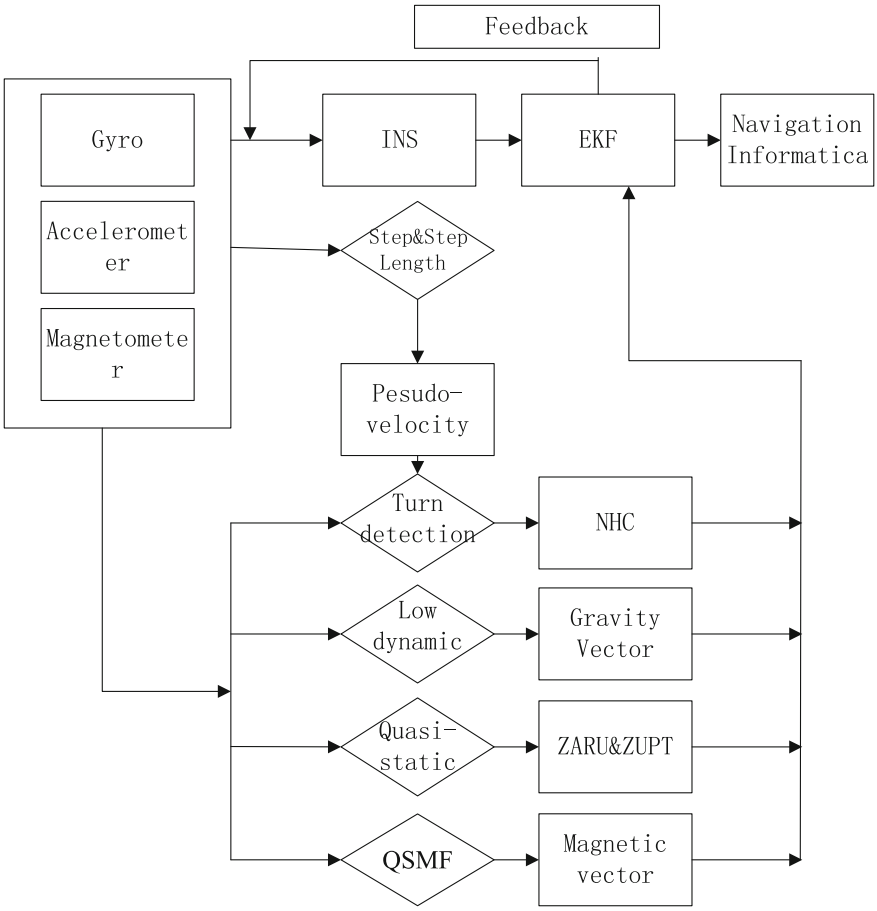


Fig. 1 Flow chart of PDR algorithm based on multi-constraint information

2.2 System State Model

In this paper, 3D position, velocity, attitude, accelerometer bias, gyro bias build 15-dimensional system state is defined as [7]:

$$\mathbf{X} = [\mathbf{p} \quad \boldsymbol{v} \quad \phi \quad \boldsymbol{b}_g \quad \boldsymbol{b}_a]^{\boldsymbol{T}} \tag{1}$$

The continuous-time state equations are given:

$$\begin{vmatrix} \delta\dot{\boldsymbol{p}} \\ \delta\dot{\boldsymbol{v}} \\ \dot{\boldsymbol{\phi}} \\ \dot{\boldsymbol{b}}_g \\ \dot{\boldsymbol{b}}_a \end{vmatrix} = \underbrace{\begin{vmatrix} \mathbf{0} & \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\boldsymbol{C}_b^n \boldsymbol{f}^b)\times & \mathbf{0} & \boldsymbol{C}_b^n \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\boldsymbol{C}_b^n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\frac{\mathbf{I}}{\mathbf{T}_{bg}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\frac{\mathbf{I}}{\mathbf{T}_{ba}} \end{vmatrix}}_{\boldsymbol{F}} \cdot \begin{vmatrix} \delta\boldsymbol{p} \\ \delta\boldsymbol{v} \\ \boldsymbol{\phi} \\ \boldsymbol{b}_g \\ \boldsymbol{b}_a \end{vmatrix} + \underbrace{\begin{vmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{C}_b^n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C}_b^n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \end{vmatrix}}_{\boldsymbol{G}} \begin{vmatrix} w_v \\ w_\psi \\ w_{b_g} \\ w_{b_a} \end{vmatrix} \quad (2)$$

where $\delta\boldsymbol{p}$, $\delta\boldsymbol{v}$, ϕ are the errors of position, velocity and attitude. $\boldsymbol{b}_g$, $\boldsymbol{b}_a$ are accelerometer bias and gyro bias. Process noise $\mathbf{w} = \begin{vmatrix} \boldsymbol{w}_v & \boldsymbol{w}_\psi & \mathbf{w}_{b_g} & \mathbf{w}_{b_a} \end{vmatrix}^T$, where $\boldsymbol{w}_v$ is the noise driven by the specific force, $\boldsymbol{w}_\psi$ is the angular velocity driving noise, $\mathbf{w}_{b_g}$ and $\mathbf{w}_{b_a}$ are the observed noise of the gyroscope and the accelerometer. $\mathbf{0}$ is a 3 × 3 zero matrix, I is a 3 × 3 identity matrix, $-\frac{\mathbf{I}}{\boldsymbol{T}_{bg}}$ and $-\frac{\mathbf{I}}{\boldsymbol{T}_{ba}}$ are 3 × 3 matrices with diagonal values $-\frac{1}{\boldsymbol{T}_{bg}}$ and $-\frac{1}{\boldsymbol{T}_{ba}}$. $\boldsymbol{C}_b^n$ is the Direction Cosine Matrix (DCM) from b-frame (i.e., the body frame) to n-frame (i.e., the navigation frame), × denotes the skew-symmetric matrix of the vector [3].

2.3 System Observation Equations Based on Multi-level Constraints Information

The multi-level constraint information used in this paper includes: ZUPT, ZARU, NHC, gravity vector, magnetic field vector.

2.3.1 Quasi-static Constraints

In pedestrian navigation, ZUPT detection has been a research focus. For example, in foot-mount pedestrian navigation, ZUPT can improve the system's available time by estimating sensor bias on-line [8]. At present, there are a lot of research on ZUPT detection. This algorithm uses the combination of, gyroscope and accelerometer signals to detect zero velocity period based on the hypothesis testing [9]. If The following conditions are satisfied: the smart phone is quasi-static.

$$\frac{1}{N}\sum_{k\in\Omega_n}\left(\frac{1}{\sigma_a^2}\left\|\boldsymbol{y}_k^a - g\frac{\bar{\boldsymbol{y}}_n^a}{\left\|\bar{\boldsymbol{y}}_n^a\right\|}\right\|^2 + \frac{1}{\sigma_w^2}\left\|\boldsymbol{y}_k^w\right\|^2\right) < -\frac{2}{N}\ln(\gamma) \quad (3)$$

N is the window size, $\boldsymbol{y}_k^a \in R^3$ represents the output of the accelerometer, $\boldsymbol{y}_k^w \in R^3$ represents the gyro output, σ_f, σ_ω represent the standard deviation of the accelerometer and the gyro noise respectively, γ is the threshold.

When the phone is in a quasi-static condition, there is a pseudo-observation velocity $\boldsymbol{v} = [\mathbf{0} \quad \mathbf{0} \quad \mathbf{0}]^T$ based on the assumption that the velocity is zero. We can get error equations:

$$\delta v = \epsilon_{\boldsymbol{v}} \tag{4}$$

$\delta v = \tilde{v} - v$, $\tilde{v}$ is the current speed estimate, $\epsilon_{\boldsymbol{v}}$ is the measuring noise.

At the same time, there exists a pseudo-angular rate $\omega_z = 0$ based on the assumption that the heading velocity is zero and the error equation

$$\delta \boldsymbol{z}_{gz} = \boldsymbol{b}_{gz} + \epsilon_{\boldsymbol{\omega}} \tag{5}$$

where $\delta \boldsymbol{z}_{gz}$ is the heading gyro output, $\boldsymbol{b}_{gz}$ is the heading gyro bias and $\epsilon_{\boldsymbol{\omega}}$ is the measuring noise.

Therefore, the design matrix under quasi-static conditions is

$$\mathrm{H} = \begin{vmatrix} 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ & 0_{1\times11} & I_{1\times1} & 0_{3\times3} & \end{vmatrix}. \tag{6}$$

2.3.2 Gravity Vector Constraint

The acceleration signals in pedestrian navigation changes periodically [4]. Therefore, the actual accelerometer observations can be modeled as gravity + noise. The noise level of the acceleration observations can be based on the size of A

$$\mathrm{A} = \left| norm\left(f^b\right) - g \right| \tag{7}$$

where g is the local gravity, $\mathrm{A} \leq |TH_{acc}|$ under the condition of small external input acceleration, then the standard deviation of accelerometer measurement noise is σ_a, which represents the accelerometer noise level; under the condition of larger external acceleration input, $\mathrm{A} > |TH_{acc}|$, accelerometer-added horizontal attitude angles are denied.

Under low dynamic conditions, accelerometer measurement information can be used to aid the attitude angle. This paper uses a tight couple to build acceleration observation equation [4], which can guarantee that it doesn't lead to Euler angles singularity problem. The acceleration error equation is as follows:

$$\delta f^n = [g^n \times]\psi + C_b^n n \tag{8}$$

where $f^n = -g^n = [\,0 \quad 0 \quad g\,]^T$, ψ is attitude error, n is the measuring noise.

Therefore, the design matrix under low dynamic conditions is

$$\mathrm{H} = |\,0_{3\times3} \quad 0_{3\times3} \quad g^n \times \quad 0_{3\times3} \quad 0_{3\times3}\,|. \tag{9}$$

2.3.3 Motion Constraints

In vehicle navigation, the NHC is widely used to improve the navigation accuracy. The basic assumption is that the vehicle only exists velocity in the forward or backward direction [10]. Based on this assumption and in combination with the practical application of indoor pedestrian navigation, such as the environment of an office promenade, if the pedestrian does not walk abnormally, the situation will proceed along the corridor. At this time, there is only the forward speed and therefore the NHC is also satisfied. However, under cornering conditions, there is a similar phenomenon as the sliding in the car navigation system, and the NHC needs to be closed in time to ensure the stability of the system. Therefore, turning detection can be used as a basis for whether or not to use the NHC, the paper uses the heading gyro observations combined with threshold to determine whether the pedestrian is turning.

$$C_b^n \tilde{g}_{\min}^b > TH_{gyro} \tag{10}$$

where, $\tilde{g}_{\min}^b$ is the current gyro observations, min represents the minimum value of a time window, TH_{gyro} is threshold.

When pedestrian is walking in a straight line, the pseudo-velocity observation $v^b = [\,v_{vx} \quad 0 \quad 0\,]^T$ can be obtained at this time. According to the NHC assumption, the following error equation

$$\delta z_v = C_n^b v^n - v^b \tag{11}$$

where v^n is the current speed estimated in the navigation frame, and C_n^b is the transformation matrix rotated from the body frame to the navigation frame as a function of attitude components. v_{vx} represents the forward speed, $stepL$, Δt are the estimated step length and footstep time interval respectively

$$v_{vx} = stepL/\Delta t \tag{12}$$

$$stepL = 4\sqrt{a_{z\max} - a_{z\min}}K \tag{13}$$

where, $a_{z\max}$, $a_{z\min}$ represents the maximum and minimum of acceleration amplitude value within the time window, K represents the coefficient.

Therefore, the design matrix under non-integrity constraints is

$$\mathrm{H} = \left| \begin{array}{cccc} 0_{3\times3} & \mathrm{C}_n^b & \left(\mathrm{C}_n^b v^n\right)\times & 0_{3\times6} \end{array} \right|. \tag{14}$$

2.3.4 Quasi-static Magnetic Field Constraints

The difficulty in using the magnetometer in the room is that the indoor steel reinforced concrete causes the disturbance of the environment magnetic field and the serious magnetic field disturbance. However, in some areas, especially in hall, the environmental magnetic field may remain stable. This phenomenon of local magnetic field stability is called quasi-static magnetic field (QSMF) [4].

In order to ensure the reliability of magnetometer measurement information, this paper probes the magnetic field of the surrounding environment and uses the magnetometer constraint only when the magnetic field is QSMF for a period of time. Quasi-static conditions are as follows:

In this paper, the observed value of the current epoch magnetometer is projected onto the n-frame to calibrate the magnetic field of the environment. By comparing the difference between the current epoch and the magnetic field in the buffer, it is determined whether the current epoch is a quasi-static magnetic field. As follows

$$\left| C_b^n \tilde{m}_k^b - \frac{1}{N}\left(\sum_{i=1}^{N} C_b^n \tilde{m}_i^b\right) \right| < TH_{mag} \tag{14}$$

where N is the number of epochs, and TH_{mag} is the threshold.

When the current epoch is in quasi-static magnetic field, the magnetic field vector can be used to assist the current attitude. This article also uses a tight couple to build magnetometer observation equation. The error equation of the magnetic field vector is:

$$\delta m^n = [m^n\times]\psi + C_b^n n \tag{15}$$

where m^n is the calibration of the magnetic field environment, n for the magnetometer measurement noise.

Therefore, the design matrix is

$$\mathrm{Z} = \left| \begin{array}{cccc} 0_{3\times3} & 0_{3\times3} & m^n\times & 0_{3\times3} \quad 0_{3\times3} \end{array} \right|. \tag{16}$$

3 Experimental Analysis and Verification

In order to verify if this algorithm designed is suitable for both indoor and outdoor environment.

We use Huawei p10, Samsung S6, Mi 4 three different smart phones, using phone's built-in sensors (such as 3D accelerometer, gyroscope, magnetometer) to collect raw data, sampling rate is 50 Hz; choose the typical indoor and outdoor scenes: Wuhan University Library, Wuhan University playground as an experimental field. Pedestrian handheld smart phones from a fixed point along a fixed direction, walking around the planning trajectories, and finally return to the starting point to form a closed loop. A total of 6 outdoor track, each phone 2 times, a single track is 390 m; indoor track 6 times, each phone 2 times, a single track is 205 m. The robust PDR algorithm designed in this paper is used to solve the pedestrian's walking trajectory (as shown in Figs. 2 and 3). Tables 1 and 2 show the difference of closing of each track.

It can be seen from Figs. 2 and 3 that the repeated tests of different terminals on the same track have good repeatability, indicating that the algorithm designed in this paper is insensitive to the sensor differences of different terminals and has good performance. It can be seen from Tables 1 and 2 that the closing errors of the two navigation positioning results of the three smartphones are all smaller and the maximum is 1.6%, which is better than the normal PDR algorithm described in the literature of the current stage with a navigation accuracy of 3–5%.

The PDR algorithm designed in this paper can effectively suppress the divergence of heading error with the time goes by. This paper compares the results of PDR of three treatment schemes: Option 1, without any constraint information; Option 2, using accelerometer and magnetometer as constraint information; Option 3, using the PDR algorithm designed in this paper. The experimental results as shown in Fig. 4.

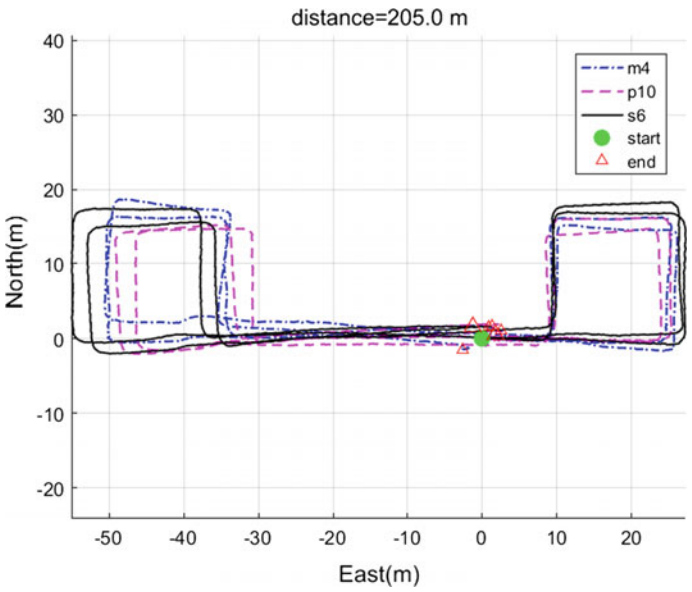


Fig. 2 The track of the experiment in the library

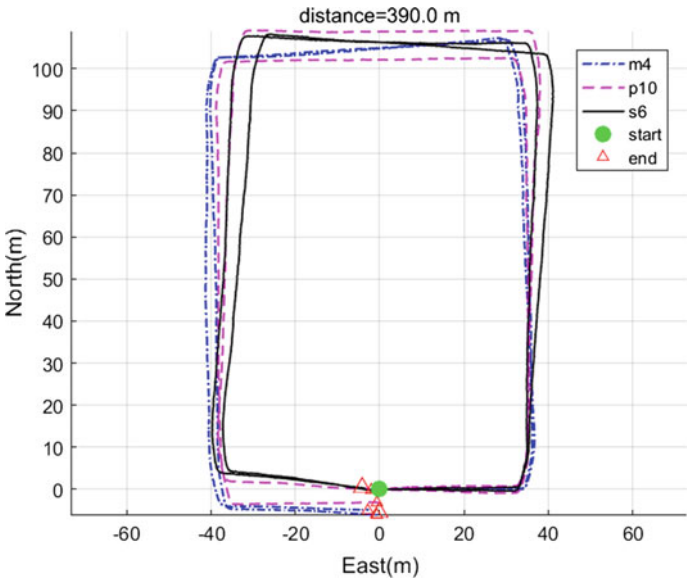


Fig. 3 The track of the experiment on the playground

Table 1 The statistical table of experimental results in the library

Phone	Distance (m)	Error 1 (m)	Error 2 (m)	Percentage (%)
P10	205	2.8	2.3	1.3
S6	205	1.9	1.9	0.9
M4	205	1.8	2.3	1.1

Table 2 The statistical table of experimental results in the playground

Phone	Distance (m)	Error 1 (m)	Error 2 (m)	Percentage (%)
P10	390	3.1	2.3	0.7
S6	390	2.1	4.3	1.1
M4	390	5.8	6.1	1.5

As can be seen from the Fig. 4, PDR algorithm designed in this paper, due to the online estimation of gyro bias and heading error using multi-layer constraint information, compared with schemes 1 and 2, the divergence of the heading over time is slow.

In addition, the proposed algorithm has the ability to estimate the accelerometer bias on-line and has more robust performance than traditional PDR algorithms. At the same time, the algorithm can provide more abundant navigation information (such as 3D velocity, Fig. 5) and can utilize more constraint information (such as motion model constraints).

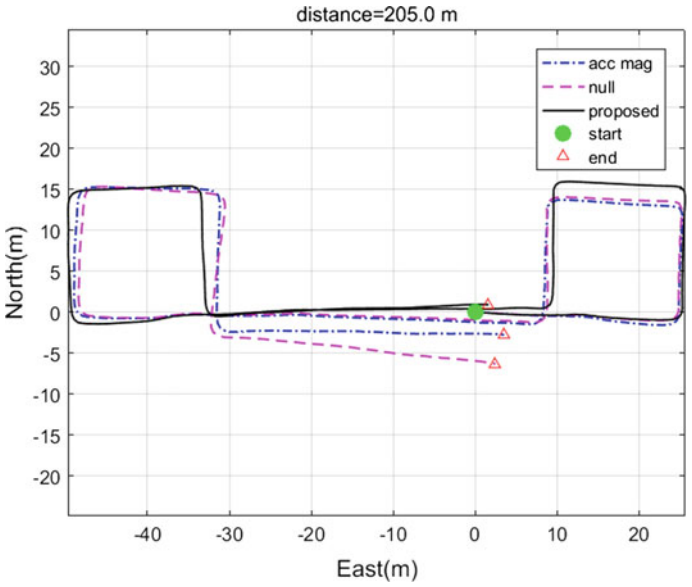


Fig. 4 The comparison of PDR result with three schemes

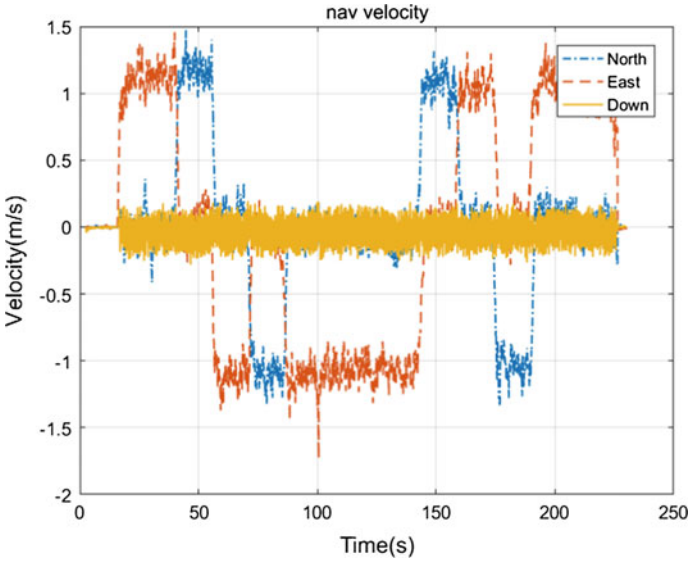


Fig. 5 3D velocity

The traditional PDR algorithm updates the position based on the gait of pedestrians, and has a low update frequency (1–2 Hz). When merging with other positioning sources (such as Wi-fi and Bluetooth), the time alignment interpolation process needs to be performed, and the error propagation lacks the strict and reliable theoretical basis. The algorithm provides 50 Hz high-frequency navigation information output, and can use the nearest neighbor information fusion of other positioning sources to solve the multi-source information fusion time not aligned; At the same time, this algorithm is based on INS, the relevant theories and methods of traditional INS navigation can be directly applied in the multi-source fusion process, which greatly reduces the algorithm complexity and workload of data fusion.

4 Conclusion

This paper presents a robust PDR algorithm based on the poor robustness and low performance of intelligent terminal pedestrian navigation algorithm. The method based on INS and fuses multi-level constraint information to ensure the navigation performance of the system. Experimental results show that the algorithm is suitable for different terminals. When hold phone on hand, the cumulative error of results for repeated are all better than 1.6% of the travel distance. Compared with the traditional PDR algorithm, the proposed PDR algorithm can improve the positioning system accuracy and robustness based on MEMS sensor. At the same time, the proposed algorithm that provides high-frequency navigation information can utilize more observation information (such as velocity observation information), and it is easy to integrate multi-source information with other positioning information sources. The next step will be to improve the robustness and navigation performance of the PDR algorithm under the existing algorithm framework for more complicated motion modes (such as calling, waving, pocket, etc.).

References

1. Chen W (2010) Research on seamless positioning algorithm for pedestrians based on GPS and self—contained sensor. University of Science and Technology of China, Hefei
2. Li Y, Zhang P, Niu X, et al (2015) Real-time indoor navigation using smartphone sensors. In: International conference on indoor positioning and indoor navigation, IEEE, pp 1–10
3. Zhuang Y, El-Sheimy N (2016) Tightly-coupled integration of WiFi and MEMS sensors on handheld devices for Indoor Pedestrian. Navig IEEE Sens J 16(1):224–234
4. Li Y (2015) Indoor intelligent Pedestrian navigation algorithm based on MEMS inertial sensor, WiFi and magnetic field. Wuhan University, Wuhan
5. Lan H et al (2015) A novel Kalman filter with state constraint approach for the integration of multiple Pedestrian navigation systems. Micromachines 6(7):926–952
6. Sun W et al (2013) MEMS-based rotary strapdown inertial navigation system. Measurement 46(8):2585–2596

7. Tao X (2017) The robust fusing strategy of multi-source information integrated navigation in the complex ground environment. In: Beidou Satellite Navigation Conference
8. Colomar DS, Nilsson J, Handel P (2012) Smoothing for ZUPT-aided INSs. In: International conference on indoor positioning and indoor navigation, IEEE, pp 1–5
9. Skog I et al (2010) Zero-Velocity Detection—An Algorithm Evaluation. IEEE Trans Biomed Eng 57(11):2657–2666
10. Chen Q (2016) Study on precision measurement technology of railway track geometry based on A—INS integrated navigation, Wuhan University, Wuhan

Research on Aircraft Landing Based on GNSS/INS/Photogrammetry Integration Using Multi-rate Extended Kalman Filter

Pengjin Zhou, Zhiwei Lv, Dianwei Cong, Zhengyang Jia and Yangjun Gao

Abstract The carrier-based aircraft system is the core component of modern aircraft carrier combat system, a maritime platform for carrier-borne fighters, whether carrier aircraft can demonstrate combat power depends on the ability of the carrier-based aircraft to take off and landing. In order to improve the safety and reliability of shipboard on warship, this paper aims at GNSS/INS/Photogrammetry integrated navigation theory, discussed from the basic level of integrated navigation system on implementation and the performance of the whole positioning. The mainly research is when satellite is not available, the inertial navigation error is corrected by the positioning of the Photogrammetry, to achieve high accuracy, high reliability and data refresh rate. The multi-rate kalman filter method is used to deal with the problem of the non-equal intervals between the INS and Photogrammetry. Finally, by simulating the accuracy of the integrated navigation positioning of INS/Photogrammetry by multi-rate EKF filter, to verify the feasibility of the aircraft landing/carrier landing.

Keywords EKF · Photogrammetry · Integrated navigation · Information fusion

1 Introduction

The take-off and landing of the aircraft is the most risky and difficulty of the pilots, how to ensure that the completion efficiency and reliability of these two processes becomes a priority. Therefore, all countries in the world have increased their investment in the technology of landing [1]. First of all, there is no flight path to track when it is landing, and can only be guided by the position of the warship. Secondly, there is no alternate airport in the sea, the deck is short and the carrier's heading must be accurate, otherwise the deviation of navigation may lead

P. Zhou (✉) · Z. Lv · D. Cong · Z. Jia · Y. Gao
Institute of Geospatial Information, Information Engineering University, Zhengzhou, Henan, China
e-mail: zhoujinera@gmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_58

to the failure of the landing. To ensure that the carrier can safely, reliably and accurately land on the deck of a moving ship or aircraft carrier, precise calculation of the relative position, speed and attitude to the warship is required. In the 1990s, the United States developed based on the relative GPS Carrier technology integrated.

Precision Approach Landing systems (Joint Precision Approach and Landing System, JPALS), to replace the existing Automatic Carrier Landing System (Automatic Carrier Landing System, ACLS) [2]. He Zhu proposes a local area difference GPS and the external highly damped SINS/RA as subsystems, using kalman filter for information fusion. The simulation results show that the scheme satisfy the accuracy requirements of aircraft landing [3]. Liu Chang designed a Visual/INS algorithm, using the attitude information of the visual system with high accuracy to restrain the divergence of the inertial navigation system's error [4]. A fuzzy adaptive kalman filter is proposed based on the tightly integration of SINS/CNS by Xie Mei Lin, correcting the SINS/CNS positioning accuracy through BD which can guarantee the navigation accuracy of the SINS/CNS for a long time [5]. However, the CNS has the weakness of starlight too weak or too few stars in the field of view, so it is possible to replace the CNS with the Photogrammetry positioning or visual positioning for the three-system integration of the aircraft landing. Although the present system can achieve the basic requirements of carrier landing, but once the GPS signal is unavailable, it's hard to meet the ship's demand.

Therefore, the SINS as the main navigation system, the GNSS and the camera positioning system, which have not accumulated over time, as an auxiliary system is proposed. And the multi-rate kalman filter is used for information fusion, the measurement model and error state model of the integration system are established. Simulation results show that when GNSS is unavailable, the integration of the INS/Photogrammetry can effectively improve the accuracy and reliability of the ship system.

2 Multi-sensor Integrated Navigation Mathematical Model

2.1 Inertial Navigation System Error Model

SINS error state parameters refer to the error variables used in the integration system, there are many error variables involved in the SINS system, such as position error, velocity error, attitude error, zero deviation of the gyroscope and zero deviation of the accelerometer, which are the main error variables. In the literature [6], the error model of INS is given by RM Rogers.

$$\begin{aligned}
\dot{\varphi} &= \varphi \times \left(\omega_{ie}^{n} + \omega_{en}^{n}\right) + \delta\omega_{en}^{n} + \delta\omega_{in}^{n} - C_{b}^{n}\varepsilon^{b} \\
\delta\dot{v}^{n} &= -\varphi \times f^{n} + \delta v^{n} \times \left(2\omega_{ie}^{n} + \omega_{en}^{n}\right) \\
&\quad + v^{n} \times \left(2\delta\omega_{ie}^{n} + \delta\omega_{en}^{n}\right) + C_{b}^{n}\nabla^{b} \\
\delta\dot{\lambda} &= \frac{v_E \tan L \sec L}{R_N + h}\delta L - \frac{V_E \sec L}{(R_N + h)^2}\delta h + \frac{\sec L}{R_N + h}\delta V_E \\
\delta\dot{L} &= -\frac{v^{n}}{(R_M + h)^2}\delta h + \frac{1}{R_M + h}\delta v^{N} \\
\delta\dot{h} &= \delta v_U
\end{aligned} \tag{1}$$

The above formula, φ represent the attitude error, δv^{n} represent the velocity error, δL, $\delta\lambda$, δh represent latitude、longitude and height error respectively, C_{b}^{n} is the attitude transfer matrix, ω_{ie}^{n} is the angular velocity of the earth's rotation, ω_{en}^{n} is the rotation angular rate of the local geographic coordinate system to the ground relatively, ∇^{b} and ε^{b} represent the accelerometer and the output error of the gyroscope respectively, R_N, R_M represent the radius of the unitary curvature and the radius of curvature of the meridian circle respectively.

$$\begin{aligned}
X(t) = [\,&\delta L \quad \delta\lambda \quad \delta h \quad \delta v_E \quad \delta v_N \quad \delta v_U \quad \varphi_E \quad \varphi_N \quad \varphi_U \\
&\nabla_x^b \quad \nabla_y^b \quad \nabla_z^b \quad \varepsilon_x^b \quad \varepsilon_y^b \quad \varepsilon_z^b\,]_{1\times 15}^{T}
\end{aligned} \tag{2}$$

2.2 Positioning Subsystem of INS/GNSS

In terms of the tactical level inertial component in this paper, we only need to consider the constant values of the gyroscope and accelerometer and the random noise items. Then:

$$\begin{cases} \delta f_{ib}^{b} = b_a + \varepsilon_a \\ \delta\omega_{ib}^{b} = b_b + \varepsilon_g \end{cases} \Rightarrow \begin{cases} \dot{b}_a = 0 \\ \dot{b}_g = 0 \end{cases} \Rightarrow \begin{cases} \delta\dot{f}_{ib}^{b} = \varepsilon_a \\ \delta\dot{\omega}_{ib}^{b} = \varepsilon_g \end{cases} \tag{3}$$

Because the state equation of SINS already contains the position, speed and other parameters, therefore, the clock correction and the clock speed of the GNSS receiver should be extended as the kalman filter state parameters. The clock correction and the clock speed can be modeled as follows [7]:

$$\begin{cases} \dot{T}_B = T_D + \varepsilon_{T_B} \\ \dot{T}_D = \varepsilon_{T_D} \end{cases} \tag{4}$$

Formula (4), T_B is the equivalent range of the receiver clock, T_D is the equivalent speed of receiver, ε_{T_B} and ε_{T_D} is the corresponding random noise items.

In this paper, the loose integration state error equation of the system is obtained by expanding the INS and GNSS system state in the Local Geographic Coordinate [7].

$$\begin{bmatrix} \delta\dot{r}^n \\ \delta\dot{v}^n \\ \dot{\varphi}^n \\ \delta\dot{b}_a \\ \delta\dot{b}_g \end{bmatrix} = \begin{bmatrix} F_{rr} & F_{rv} & 0 & 0 & 0 \\ F_{vr} & F_{vv} & F_{v\varphi} & F_{vf} & 0 \\ F_{\varphi r} & F_{\varphi v} & F_{\varphi\varphi} & 0 & F_{\varphi\omega} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta r^n \\ \delta v^n \\ \varphi^n \\ \delta b_a \\ \delta b_g \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ F_{vf} & 0 \\ 0 & F_{\varphi\omega} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_a \\ \varepsilon_g \end{bmatrix} \quad (5)$$

In Eq. (5), because the derivation process of F_{rr}, F_{rv} is more complicated and takes up more space, see the literature [7] in detail. The INS/GNSS loose integration adopts the position and velocity combination method, the difference between the position and velocity of the two systems are the observation. The measurement equation of the location and velocity in the Local Geographic Coordinate system can be obtained.

$$Z(t) = [\, r^n_{INS} - r^n_{GNSS} \quad v^n_{INS} - v^n_{GNSS} \,]^T = H(t)X(t) + V(t) \quad (6)$$

r^n_{INS}, r^n_{GNSS} represent the position figured out by the INS and GNSS system respectively, v^n_{INS}, v^n_{GNSS} represent the velocity figured out by the INS and GNSS respectively. By literature [8], the measurement equation of the location and velocity of the local geographic coordinate system can be directly obtained as follows:

$$\begin{bmatrix} r^n_I - r^n_G \\ r^n_I - r^n_G \end{bmatrix} = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta r^n \\ \delta v^n \\ \varphi^n \\ \delta b_a \\ \delta b_g \end{bmatrix} + \begin{bmatrix} \varepsilon^n_r \\ \varepsilon^n_v \end{bmatrix} \quad (7)$$

I represents the identity matrix of 3 rows and 3 columns.

2.3 Positioning Subsystem of INS/Photogrammetry

According to the principle of photogrammetry, the high resolution CCD camera can provide a high-precision location and attitude information of aircraft by shooting a mark field with a ground mark. Using photogrammetry information to estimate and correct the error of SINS, therefore, the corrected INS can maintain high accuracy position, high accuracy attitude, highly data update rate information. The integration of photogrammetry and INS can be integrated with the relative position and attitude angle of the two systems, the INS continuously provides the attitude to

camera, the difference between the SINS and the camera is calculated as the filter observation value, the error of the SINS system is corrected by the kalman filter method [9].

The state is:

$$X(t) = [\,\delta L \quad \delta\lambda \quad \delta h \quad \phi_e \quad \phi_n \quad \phi_u \quad \delta v_e \quad \delta v_n \quad \delta v_u \quad \varepsilon_{bx} \quad \varepsilon_{by} \quad \varepsilon_{bz}\,]^T \tag{8}$$

The state equation of the system is:

$$\dot{X}(t) = \Phi(t)_{12\times12}X(t)_{12\times1} + \Gamma(t)_{12\times9}W(t)_{9\times1} \tag{9}$$

The difference of the position and attitude of the INS and Photogrammetry is input to the filter as the measurement [10].

The measurement equation of the system is:

$$Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \widetilde{P}^n_{b(SINS)} - \widetilde{P}^n_{b(cam)} \\ \tilde{\phi}^n_{b(SINS)} - \tilde{\phi}^n_{b(cam)} \end{pmatrix} = H(t)X(t) + V(t) \tag{10}$$

where V(t) is measurement noise in the formula:

$$H(t) = \begin{bmatrix} \frac{-1}{\cos\theta} & \begin{bmatrix} \sin\psi & \cos\psi & 0 \\ \cos\psi\ \cos\theta & -\sin\psi\ \cos\theta & 0 \\ \sin\psi\ \sin\theta & \cos\psi\ \sin\theta & -\cos\theta \end{bmatrix} & 0_{3\times9} \\ 0_{3\times6} & diag[\,R_M \quad R_N \cdot \cos L \quad 1\,] & 0_{3\times3} \end{bmatrix} \tag{11}$$

γ, θ, ψ are the rolling, pitching and heading angle of the carrier respectively.

3 Multi-rate Kalman Filter Algorithm

Because of the differences between the sensors of the camera and INS, it can cause a different data sampling cycle; the problem of the non-equal interval sampling period and the interruption of the system sampling point are presenting in the Photogrammetry system. The traditional single rate kalman filter algorithm is not effective. Therefore, a multi-rate kalman information fusion algorithm is needed to deal with multi-rate filter [8].

(1) One step state prediction and prediction error variance matrix

$$\begin{cases} \widehat{X}_{k,k-1} = \Phi_{k,k-1}\widehat{X}_{k-1} \\ \widehat{P}_{k,k-1} = \Phi_{k,k-1}P_{k-1}\Phi^T_{k,k-1} + \Gamma_{k,k-1}Q_{k-1}\Gamma^T_{k,k-1} \end{cases} \tag{12}$$

(2) Update of filter gain matrix, estimation state and error variance matrix

$$\begin{cases} \widehat{K}_k = P_{k,k-1}H_k^T\left[H_kP_{k,k-1}H_k^T + R_k\right]^{-1} \\ \widehat{X}_k = \widehat{X}_{k,k-1} + K_k\left[Z_k - H_k\widehat{X}_{k,k-1}\right] \\ P_k = [I - K_kH_k]P_{k,k-1}[I - K_kH_k]^T + K_kR_kK_k^T \end{cases} \tag{13}$$

It is hypothesized that at the moment of k, when the quantity measurement information input is just SINS and there's no Photogrammetry navigation information, the filter thinks that the system has arbitrarily large errors, then the observed noise variance of the system R_k tends to infinity, and the filter gain matrix tends to zero, therefore the state estimation and error variance matrix are:

$$\begin{cases} \widehat{X}_k = \widehat{X}_{k,k-1} \\ P_k = P_{k,k-1} \end{cases} \tag{14}$$

Equation (13) indicates that the state estimate and the estimated error variance matrix are unchanged from the previous one, without the positioning of Photogrammetry and the observation of GNSS, the filter equation should only be time updated (status update), rather than measurement update. In the moment of t_1, t_2, t_4 the filter takes time update only. When we had GNSS observations at time t_3 only, carried out status updates and measurement update at the same times. At this point, the observation model is adjusted, select the observation matrix and observation vectors of the INS/GNSS integration to filter. At the time of t_5 there is just the Photogrammetry observation, the observation model of the INS and Photogrammetry integration subsystem is selected for status update and measurement update [8].

4 Design of Fault Tolerant Scheme and Analysis of Algorithm Complexity

4.1 Design of Fault Tolerant Scheme

The fault-tolerant system structure of GNSS/SINS/Photogrammetry integrated navigation system is shown in Fig. 1, the SINS is used as the benchmark system and assume that it's not going to fail. Two other subsystems were detected by fault identification system. First, the GNSS information and airborne Photogrammetry information are sent to the pretreatment unit through the data interface, the pretreatment unit has preprocess the spatiotemporal unification and data anomaly detection. By integrating landing guide information through fusion algorithm, to get the relative parameters of the aircraft; secondly, decision unit determines whether

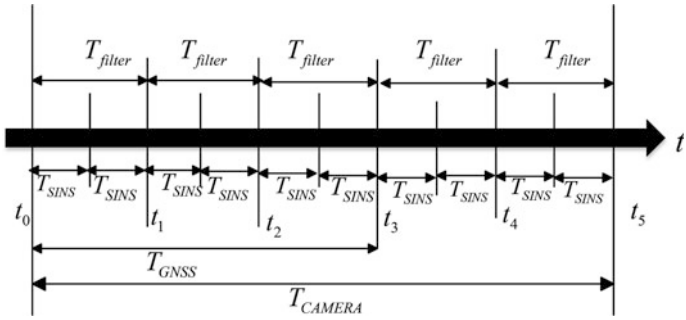


Fig. 1 The sample period of multi-rate integration system

the current bootstrap mode is available from the boot mode library. Furthermore, the model management unit is used to evaluate the performance of the currently available guidance mode, and select the optimal boot mode. Once the decision unit has identified the subsystem failure, the boot mode is isolated until normal. At this point, the main filter uses the state estimation results of another subsystem only, and the settlement is carried out (Fig. 2).

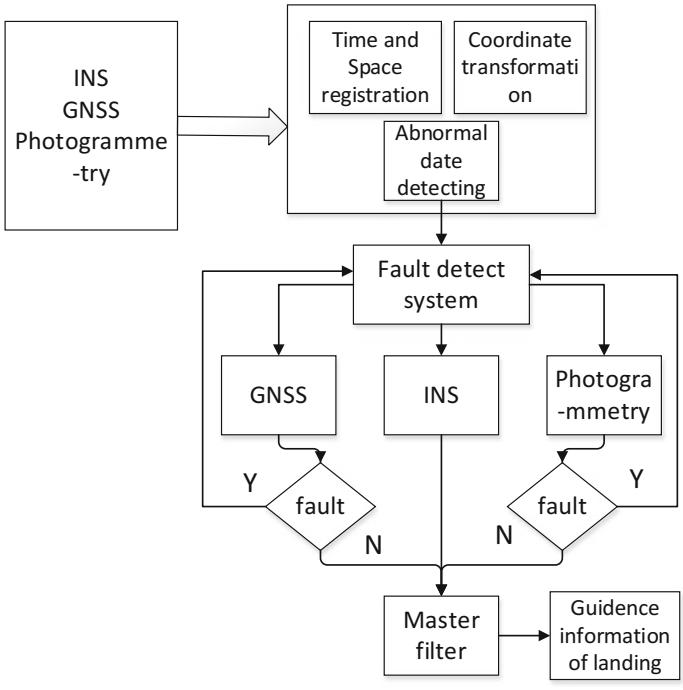


Fig. 2 Fault tolerant systems of GNSS/INS/Photogrammetry

4.2 The Analysis of Algorithm Complexity

For the feasibility of the landing, considering not only the convergence precision of the algorithm, but also the computation complexity of the algorithm must be considered. The landing of warship has strict requirement for real-time, so it is important to choose the algorithm with fast speed and low processor resource. Before the performance analysis of the INS/Photogrammetry, setting the dimension of the system equation is n, and the dimension of the measurement equation is m.

In the filtering process: if $A \in R^{n\times m}, B \in R^{n\times m}$, then the floating-point operations of $A \pm B$ is $m \times n$; if $A \in R^{n\times m}, B \in R^{m\times k}$, then the floating-point operations of AB is $2mnk - nk$; if $A \in R^{n\times n}$, then the floating-point operations of A^{-1} is n^3. The state equation of INS/Photogrammetry system has linear characteristics, therefore to calculate the jacobian matrix only, the system model is discretized before the filtering.

Equivalent discrete form of the system matrix:

$$\Phi_{k+1,k} = I + TF_k + \frac{T^2}{2!}F_k^2 + \frac{T^2}{3!}F_k^3 + \cdots$$

Generally, we take the three-order approximation algorithm, its complexity is approximately $n^3 + 4n^2$ times of floating point arithmetic.

The algorithm time complexity corresponds to the time spent by this algorithm, and the total computational complexity is:

$$f(n) = 6n^3 + 2m^3 + 9n^2 + 2m^2n + mn^2 + 4mn + 2m + n$$

The state dimension of the INS/photogrammetry system is 12, and the measurement dimension is 6, the number of floating point operations that can be calculated through the upper equation is 14500, which is a 0.23M FLOPS floating point computing power. The computational complexity of this algorithm is applicable to the real-time application of aircraft landing.

5 The Simulation Verification

To verify the feasibility of the integrated navigation system to the aircraft landing, numerical simulation is carried out. Firstly, the trajectory of the aircraft is simulated and the position parameters of the trajectory generator are: initial latitude 34°, initial longitude 113°, the initial level 116 m. The trajectory has flat flying, climbing, accelerating, turn and so on. The output of SINS devices is calculated from the generated trajectory, and the error is added as the actual output. Other simulation parameters are as follows, the simulation time is 480 s. Inertial device error of measurement indicators: gyro constant drift 0.26°/h, random white noise error of

0.26°/h (1σ), data update rate of 100 Hz. Constant value deviation of accelerometer is $1 \times 10^{-4} g(1\sigma)$, random white noise error is $1 \times 10^{-4} g(1\sigma)$. GNSS precision index: positioning error of the GNSS receiver is 10 m, and the data update rate is 10 Hz. The simulated aircraft trajectory as the true value, the camera positioning system error index: the random error of adding the standard deviation 0.5 m in the location of the photo node was used as the location result of the photogrammetry, Photogrammetry system attitude accuracy 0.2°, and the data update rate was 5 Hz.

Simulation 1: For the INS/GNSS state equation given above, the error of INSS/GNSS integrated system was estimated by EKF filter algorithm. The simulation results are shown as follows: The real trajectory of the simulation aircraft and the position and attitude error of the algorithm of INS/GNSS based on EKF filter are shown in Fig. 3.
Simulation 2: the multi-rate EKF filter algorithm is used for the INS/Photogrammetry state equation given above. The simulation results are as follows:

Figures 4, 5, 6 and 7 shows that, when GNSS is not available, Photogrammetry can replace GNSS for the divergence correction of the INS that can achieve GNSS correction effect. The correction of the INS attitude error by Photogrammetry is better than the result of GNSS correction, however, due to the low accuracy of the attitude that the camera calculated at the turning point and the accumulation of attitude error reduces the accuracy of the heading angle.

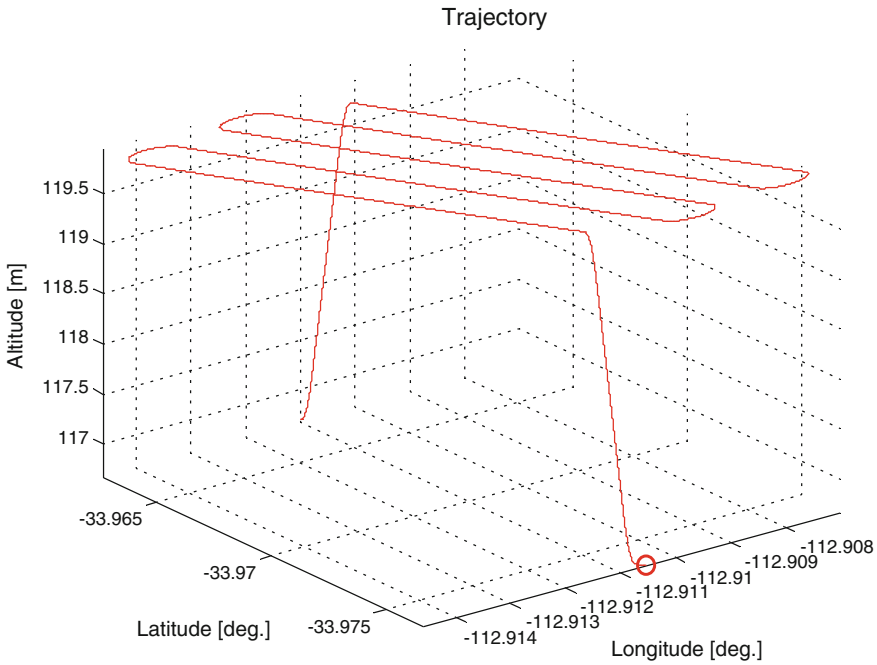


Fig. 3 Simulation vehicle trajectory

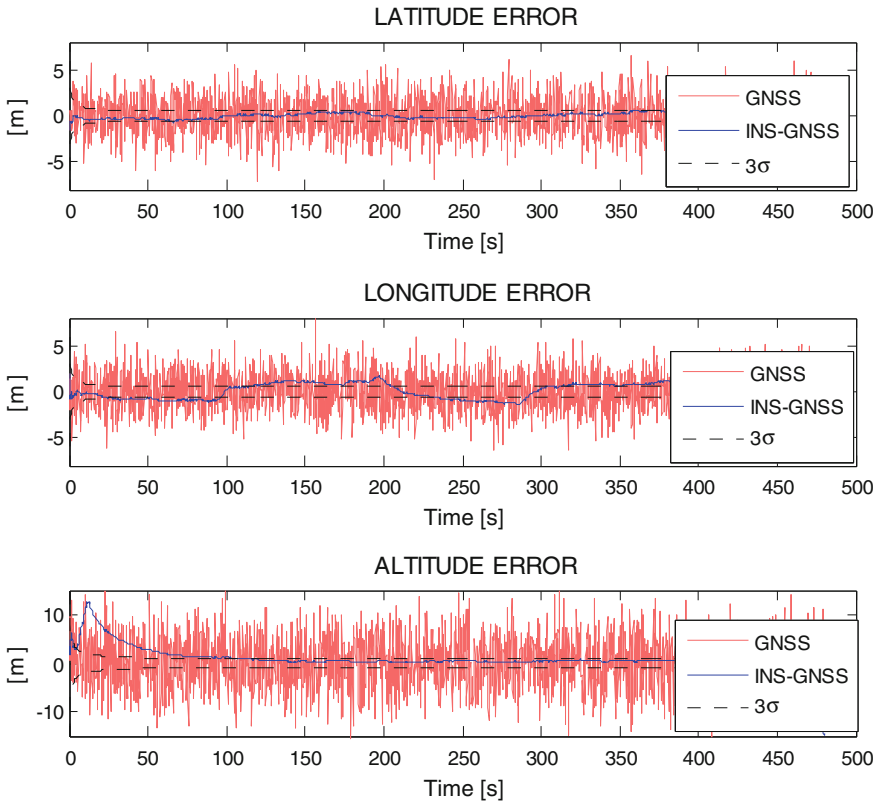


Fig. 4 Curve: the GNSS and INS/GNSS position error

It can be seen from Table 1 that when GNSS is unavailable, using integration mode of INS/Photogrammetry to replace GNSS/INS system is feasible, comparing the performance of INS/GNSS and INS/Photogrammetry. It can be concluded that the INS/Photogrammetry integration model can achieve even better positioning performance than INS/GNSS. The positioning performance can meet class II precision approach that achieve the positioning accuracy, availability and other indicators demanded (Table 2).

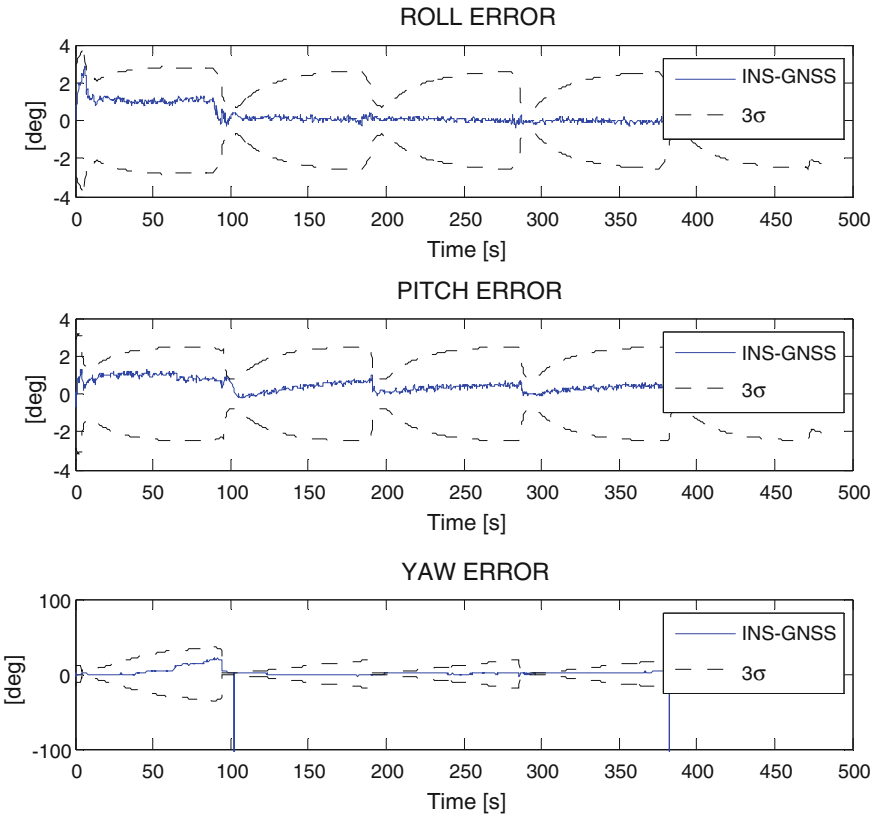


Fig. 5 Curve: the GNSS and INS/GNSS attitude error

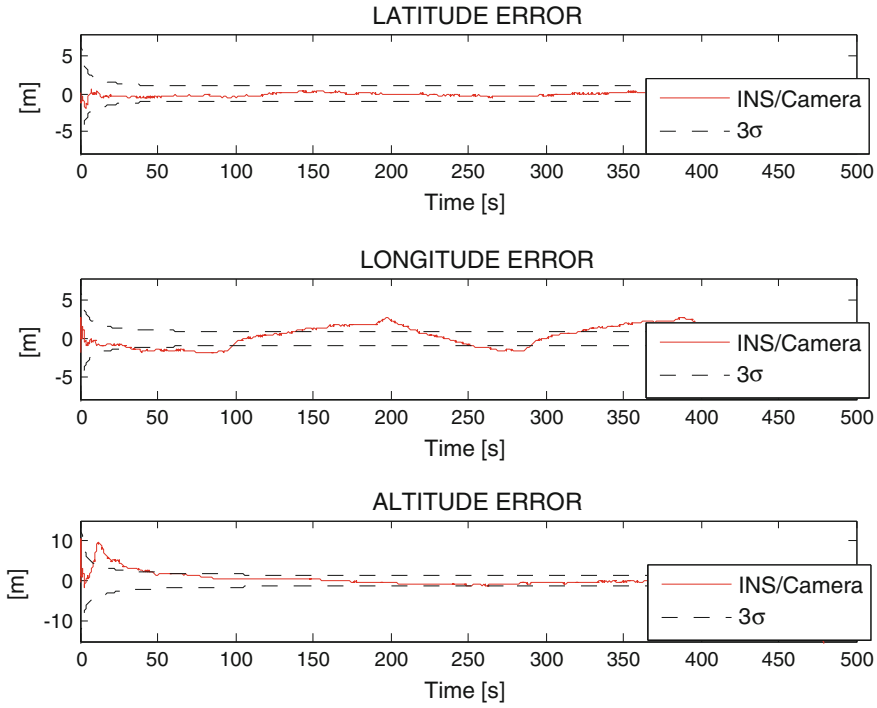


Fig. 6 Curve: the INS/Photogrammetry position error

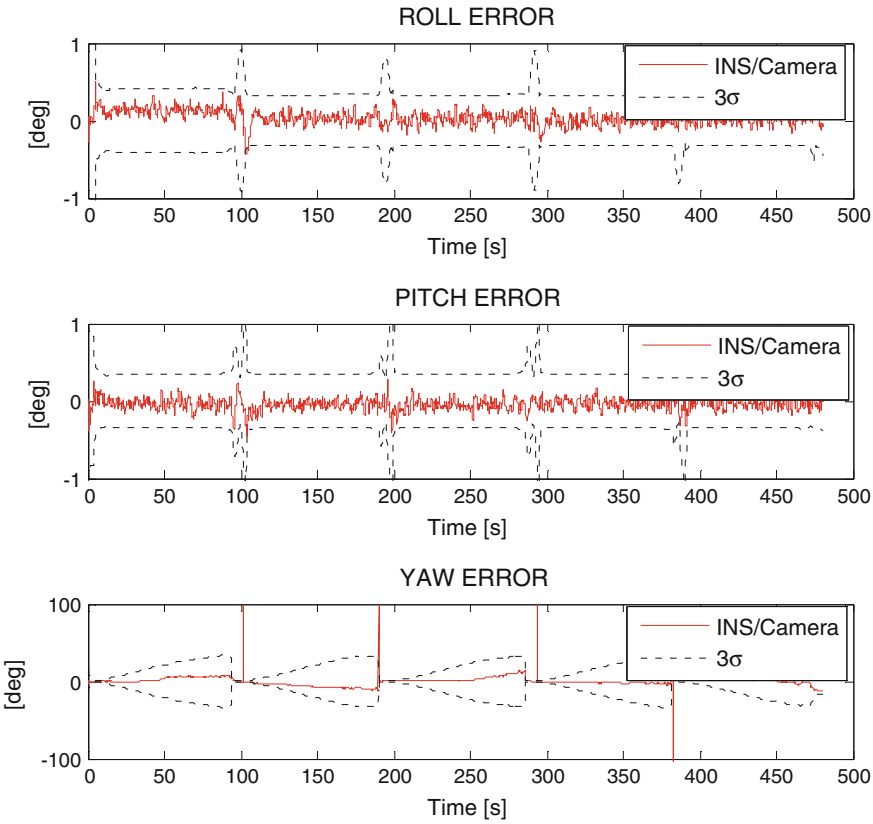


Fig. 7 Curve: the INS/Photogrammetry attitude error

Table 1 Analysis of EKF algorithm complexity based on INS/Photogrammetry

Steps of algorithm		Flops
One step prediction of state	$\widehat{X}_{k+1}^{-} = \Phi_{k,k+1}\widehat{X}_{k}^{+}$	$2n^2$
Prediction of mean square error	$P_{k+1}^{-} = \Phi_{k,k+1}P_{K}^{+}\Phi_{k,k+1}^{T} + Q_k$	$4n^3 + n^2$
Calculate the gain of filter	$K_{k+1} = P_{k+1}^{-}H_{k+1}^{T}\left(H_{k+1}P_{k+1}^{-}H_{k+1}^{T} + R_{k+1}\right)^{-1}$	$2m^3 + 2$ $m^2n + 2mn + m$
State estimation	$\widehat{X}_{k+1}^{+} = \widehat{X}_{k}^{-} + K_{k+1}\left(Z_{k+1} - H_{k+1}\widehat{X}_{k+1}^{-}\right)$	$2mn + m + n$
State covariance estimate	$P_{k+1}^{+} = (I - K_{k+1}H_{k+1})P_{k+1}^{-}$	$mn^2 + 2n^2$
Total flops	$2m^3 + 4n^3 + 5n^2 + 2m^2n + mn^2 + 4mn + 2m + n$	

Table 2 Positioning performance comparison of INS/GNSS and INS/Photogrammetry

Positioning performance of INS/GNSS and INS/Photogrammetry			
RMSE	Latitude (m)	Longitude (m)	Altitude (m)
INS/GNSS	3.5158e−01	1.1617e−01	3.0021e+00
INS/Photogrammetry	3.1489e−01	1.3810e−01	2.0818e+00
RMSE	Roll (deg)	Pitch (deg)	Yaw (deg)
INS/GNSS	7.9798e−02	2.9472e−01	3.5423e−01
INS/Photogrammetry	1.7741e−02	1.5396e−01	1.8218e−01

6 Conclusion and Outlook

Based on INS/GNSS/Photogrammetry system, the integrated navigation scheme of aircraft landing/ship is presented in this paper, and through simulation to be verified. The simulation results show that when the GNSS receiver signal is interrupted or invalid, according to the correction of the SINS error by the positioning of Photogrammetry that can guarantee the positioning accuracy of the integrated navigation system too. The feasibility of the INS/GNSS/Photogrammetry integrated system for aircraft landing/ship is demonstrated, it provides a theoretical basis for the following multi-system integration and multi-data fusion.

Acknowledgements This research work was supported by the State Key Laboratory of Geo-information Engineering (SKLGIE2015-M-2-5), the Natural Science Foundation of China (41604032).

References

1. Cong D, Xu Q, Dong M (2015) Photogrammetry/SINS integrated positioning technology used in the evaluation of GNSS dynamic positioning test. J Zhengzhou Inst Surveying Mapp 3:244–247
2. Dong M (2014) Research on key technology of GNSS/INS/Vision integrated navigation information fusion. PLA Information Engineering University, Zhengzhou
3. HE Z, Wang X (2013) A high-precision integrated navigation program for aircraft landing phase. Aero Weaponry 6:12–16
4. Liu C, Yang L, Liu F et al (2016) Navigation algorithm based on inertial/vision information fusion of UAV autonomous landing. Navig Pos Timing 3(6):21
5. Xie M, Yang X et al (2015) Research on the airborne SINS/CNS integrated navigation system assisted by BD navigation system. In: New navigation technology and application workshop
6. Feng G (2012) Research on observability analysis and dynamics filter algorithms of monocular camera/INS integrated navigation. National University of Defense Technology, Changsha, China
7. Liu S (2012) Research and implementation of GPS/INS integrated navigation algorithms. PLA Information Engineering University, Zhengzhou
8. Tang D, Bi B, Wang X (2010) Summary on technology of automatic landing/carrier landing. J Chin Inertial Technol 18(5):550–555
9. Yu Y, Xu J, Zhang L (2014) Research on SINS/binocular vision integrated position and attitude estimation algorithm. Chin J Sci Instr 35(10):2170–2176

Wi-Fi Assisted GNSS Positioning and Continuous Tracking

Changsheng Liu, Xiukui Li and Xiaoyan Liu

Abstract Currently location-based services have been widely used. It is critical to ensure the effectiveness of universe positioning. GNSS and Wi-Fi fingerprint positioning are used in outdoor and indoor environments, respectively. Both can achieve rather accurate positioning results. Nevertheless, in some situations, any of the two methods cannot be employed by user receivers to obtain positions because the satellites or Wi-Fi signals may be insufficient at the junction of two positioning systems. Hence, in this paper, we propose a GNSS and Wi-Fi positioning fusion method to implement the availability and continuity of receiver's navigation and positioning. First, a Wi-Fi-assisted GNSS positioning method is proposed. When the number of satellites is insufficient, assistant information provided by Wi-Fi systems can improve the availability of positioning. When both GNSS and Wi-Fi positioning are available, group Kalman filter technique is hired to improve the positioning accuracy. Simulation results show that the proposed positioning methods can improve the availability and accuracy of positioning.

Keywords GNSS · Wi-Fi fingerprint · Assisted positioning · Continuous tracking

1 Introduction

Location-Based Service (LBS) has attracted extensive attention due to the great potential in improving the quality of personal services [1]. People require high-performance positioning techniques because of wide utilization of navigation terminals and upgraded hardware. High-accuracy and robust positioning is of the utmost importance.

Global Navigation Satellite System (GNSS) is used for outdoor positioning, which can achieve high positioning accuracy. Nevertheless, the terminals may not be able to obtain sufficient available satellite signals for precise positioning near

C. Liu · X. Li (✉) · X. Liu
Dalian University of Technology, Dalian, China
e-mail: xli@dlut.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_59

high buildings, in the forest, in urban canyons or in indoor environments. It results in discontinuity of positioning, and reduces the availability of positioning. Wi-Fi fingerprinting based positioning is the dominating method for indoor positioning because of the wide coverage and obvious differentiation of Wi-Fi signals. However, at the entrance of a building, the terminal may not be able to collect sufficient access point (AP) signals or satellite signals for positioning, and this will result in positioning failure.

Therefore, it is challenging to achieve precise positioning and continuous tracking in the junction region between indoor and outdoor. Yan et al. in [2] propose a positioning network of combining GNSS and Wi-Fi signals. The environment is classified into multiple cases for positioning; however, the classification is simple. In [3], Jia et al. propose a method for indoor and outdoor positioning based on multiple sensors. Data measured by light sensors, GNSS modules and geomagnetic sensors are combined to calculate the probability of the receiver being located indoor or outdoor. Since a variety of sensors is required to work together, it has higher complexity. The authors in [4] propose to combine GNSS and Bluetooth signals to determine whether the receiver is indoor or outdoor, and use Bluetooth technique to obtain indoor position. A large amount of beacons are required in this method. In [5] heterogeneous positioning systems are jointly used to achieve a seamless positioning system, and this improves the positioning performance by maximizing the utilization of existing positioning systems.

Due to the wide coverage of wireless AP signals, the receiver is able to receive GNSS and Wi-Fi signals simultaneously when it is located in the junction area of indoor and outdoor (an area around the door and it includes part of indoor and part of outdoor). It will improve the availability and positioning accuracy of LBS by using those signals in the junction area. In this paper, based on the number of satellites and APs from which the receiver can receive available signals, we divide the junction area into multiple zones and employ different positioning techniques in those zones to reduce positioning failure. First, for the case of the number of satellites being less than 4, we propose to use the terminal altitude information obtained by Wi-Fi signals to assist GNSS for positioning. When both GNSS and Wi-Fi signals are sufficient for the positioning, group Kalman filter technique is hired to improve the positioning accuracy. Simulation results show that the proposed positioning methods can improve the availability and accuracy of positioning.

2 System Model

When the terminal enters the junction area, the situations will be different regarding the availability of the received GNSS satellite signals and Wi-Fi AP signals. In order to attain continuous positioning and tracking, the terminal should be able to achieve reliable positioning for different situations. Table 1 shows the positioning solutions for different situations, and note that the terminal may also use other

Table 1 Positioning solutions for different situations

N_A	N_S				
	0	1	2	3	≥4
0	DR	DR	DR	Assisted GNSS	GNSS
1	DR&Wi-Fi	DR&Wi-Fi	DR&Wi-Fi	Wi-Fi assisted GNSS	Fusion of GNSS and Wi-Fi
2	DR&Wi-Fi	DR&Wi-Fi	DR&Wi-Fi	Wi-Fi assisted GNSS	Fusion of GNSS and Wi-Fi
3	DR&Wi-Fi	DR&Wi-Fi	DR&Wi-Fi	Wi-Fi assisted GNSS	Fusion of GNSS and Wi-Fi
≥4	Wi-Fi	Wi-Fi	Wi-Fi	Wi-Fi assisted GNSS	Fusion of GNSS and Wi-Fi

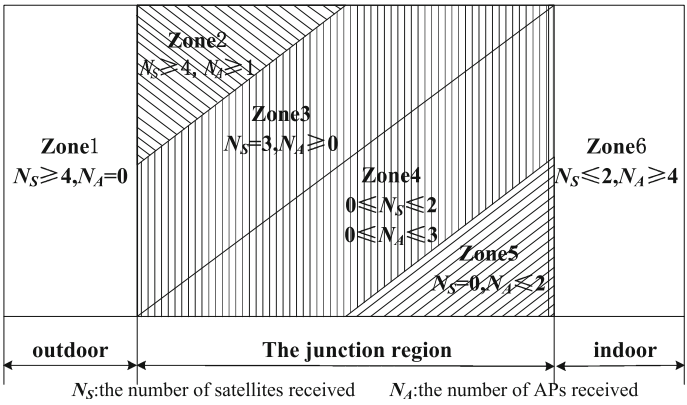


Fig. 1 Junction region of indoor and outdoor

sensors such as gyroscopes, barometers and geomagnetic sensors, etc. to further improve positioning accuracy. Here *N*s is the number of visible satellites and N_A is the number of APs which the terminal can access. According to signal availability, the junction region can be divided into multiple zones, as is shown in Fig. 1:

(1) Zone 1, denoted by Z_1, where the terminal can receive signals from more than four navigation satellites; however it cannot receive Wi-Fi signals, i.e., *N*s $\geq$ 4 and $N_A = 0$.
(2) Zone 2, denoted by Z_2, where the terminal can receive signals from three navigation satellites, it may or may not receive Wi-Fi signals, i.e., *N*s = 3, $N_A \geq 0$.
(3) Zone 3, denoted by Z_3, where the terminal can receive signals from at least four navigation satellites and at least one AP, i.e., *N*s $\geq$ 4 and $N_A \geq 1$.
(4) Zone 4, denoted by Z_4, where the terminal can receive signals from two or less satellites and cannot receive Wi-Fi signals, i.e., *N*s $\leq$ 2 and $N_A = 0$.

(5) Zone 5, denoted by Z_5, where the terminal can receive signals from two or less satellites and three or less APs, i.e., $0 \leq Ns \leq 2$ and $0 < N_A \leq 3$.
(6) Zone 6, denoted by Z_6, where the terminal can receive signals from two or less satellites and more than four APs, i.e., $N_A \geq 4$, $Ns \leq 2$.

Note that the division of those regions is based on the number of satellites and APs, and practically the geographical zone shape may not be regular.

In Z_1, GNSS signals can be used for positioning [6]. Dead reckoning (DR) [7] algorithm can be used to locate the receiver in Z_4. The results of Wi-Fi fingerprint positioning and dead reckoning can be combined to obtain the receiver's position in Z_5 [8]. In Z_6, Wi-Fi fingerprinting based method can be used for positioning [9]. According to the signal availability of Z_2 and Z_3, Wi-Fi assisted GNSS positioning algorithm for Z_2 and GNSS and Wi-Fi fusion positioning algorithm for Z_3 are proposed in this work.

2.1 Zone 2: Wi-Fi Assisted GNSS Positioning

Generally, the terminal needs to receive signals of four or more satellites to achieve positioning. When only three satellites are visible to the receiver, other information is required to assist the positioning. One of the solutions is to use receiver altitude information [10]. When Wi-Fi signals are available, the signal fingerprints can be used to determine which floor in the building the receiver is located at. The relationship between the floors and altitude can be pushed to the receiver by Wi-Fi message when the receiver enters the building. Thus, measured satellite pseudoranges and altitude can be combined to find the receiver position. Herein, we propose a fingerprint based algorithm to determine the floor on which the receiver is located. This algorithm requires building the received signal strength (RSS) database in advance, and the database will be used during positioning. The data processing is divided into the following two phases.

2.1.1 Offline Data Processing Phase

First we create an offline database of floor information. Select N location points in each layer and record accessed APs' MAC and signal strengths that can be received at those points. Then we can obtain the available APs and the range of their signal strength in each floor. The range of signal strength for all APs in the building is denoted by $\Gamma = [\xi_{\min}, \xi_{\max}]$, where ξ represents the signal strength. Dividing this range into M segments and the range of each segment is

$$\Gamma_m = [\xi^m_{\min}, \xi^m_{\max}] \tag{1}$$

where $m = 1, 2, \ldots, M$, and

$$\xi_{\min}^{m} = \xi_{\min} + \frac{\xi_{\max} - \xi_{\min}}{M} \cdot (m - 1) \tag{2}$$

$$\xi_{\max}^{m} = \xi_{\min} + \frac{\xi_{\max} - \xi_{\min}}{M} \cdot m \tag{3}$$

M is determined by the actual environment, and usually the segment is spaced by 10 dBm. Let

$$G = \{ \Gamma_1 \quad \Gamma_2 \quad \ldots \quad \Gamma_M \} \tag{4}$$

To have reliable signals, the APs accessible in each floor are sorted in a descending way according to the maximum signal strength, and the first K_1 APs are selected to form a set

$$\Omega_{\mathrm{i}} = \{ A_{i,1} \quad A_{i,2} \quad \ldots \quad A_{i,K_1} \} \tag{5}$$

where i denotes the floor identity; j denotes the AP identity and $j = 1, 2, ..., K_1$. Define $R_{i,j}^{m}$ as the occurrence number of signal strength of $A_{i,j}$ falling within Γ_m. The occurrence number matrix of AP can be given by

$$\boldsymbol{\Psi}_{i,j} = [R_{i,j}^{1} \quad R_{i,j}^{2} \quad \ldots \quad R_{i,j}^{M}] \tag{6}$$

Then the occurrence frequency of signal strength of $A_{i,j}$ falling within Γ_m is

$$\rho_{i,j}^{\mathrm{m}} = \frac{R_{i,j}^{m}}{N} \tag{7}$$

where $m = 1, 2, \ldots, M$. Thus, a frequency matrix can be formed

$$\mathbf{P}_{i,j}(\Omega_i) = [\rho_{i,j}^{1} \quad \rho_{i,j}^{2} \quad \ldots \quad \rho_{i,j}^{M}] \tag{8}$$

K_1 depends on the actual measured Wi-Fi signal strengths in the building. Denote the number of APs whose maximum signal strength is greater than a certain threshold in floor i as B_{i}. Then we select the minimum of B_{i} as the value of K_1, i.e., $K_1 = \min(B_{\mathrm{i}})$, $i = 1, 2, \ldots, H$, where H is the total number of floors in the building. Thus an offline database can be created as shown in Table 2, where F_{i} represents the ith floor and $\mathbf{P}_{i,j}$ represents the frequency matrix of the jth AP in the ith floor.

Table 2 Offline database of floor information

Floor	AP	Signal occurrences and frequency matrix
F_1	$A_{1,1}$	$\mathbf{\Psi}_{1,1}$, $\mathbf{P}_{1,1}$
…	…	…
F_i	$A_{i,1}$	$\mathbf{\Psi}_{i,1}$, $\mathbf{P}_{i,1}$
…	…	…
F_H	$A_{H,1}$	$\mathbf{\Psi}_{H,1}$, $\mathbf{P}_{H,1}$
…	…	…

2.2 Online Positioning Phase

The receiver measures the signal strength of accessible APs when it is located at a point in a floor. Sort the received signal strength descendingly, and select the first K_2 APs to form

$$\Omega_0 = \{ A_1 \quad A_2 \quad \ldots \quad A_{K_2} \} \tag{9}$$

K_2 depends on K_1. K_1 is determined by the signal strength range of APs in the entire building and K_2 only relies on the signal strength range of APs in a floor. Hence, $K_2 < K_1$. To reduce the computation complexity, generally let $K_2 = K_1/2$.

The AP matching process is divided into two phases. In the first stage, determine whether an AP whose signals are available to the receiver is in the ith floor, $1 \leq i \leq H$, Define

$$d_{i,j} = \begin{cases} 1, A_j \in \Omega_0 & \text{and} \quad A_j \in \Omega_i \\ 0, A_j \in \Omega_0 & \text{and} \quad A_j \notin \Omega_i \end{cases} \tag{10}$$

where i denotes the floor identity, and $j = 1, 2, \ldots, K_2$. After $d_{i,j}$ of all APs in Ω_0 are calculated, the coarse floor differentiation value (CFDV) D_i of floor i will be

$$D_i = \sum_{j=1}^{K_2} d_{i,j} \tag{11}$$

In the second stage, the D values of all floors are compared to determine the floor which the receiver is located at. Let

$$D_{\max} = \max(D_i), \quad i = 1, 2, \ldots, H \tag{12}$$

Define

$$f_i = \begin{cases} 1, D_i = D_{\max} \\ 0, D_i \neq D_{\max} \end{cases} \quad i = 1, 2, \ldots, H \tag{13}$$

$$F = \sum_{i=1}^{H} f_i \tag{14}$$

where F is the total number of floors that can achieve the maximum D value. Let

$$Q = \{ q_1 \quad q_2 \quad \ldots \quad q_{\mathrm{F}} \} \tag{15}$$

where Q is a set of all floors whose f_i is equal to 1, and q_{i} is the floor identity.

If $F = 1$, the only element in Q is the floor on which the receiver is located; If $F > 1$, it means that the D value of multiple floors is greater than 1; thus, the receiver cannot determine the floor on which the receiver is located. Then we need to further process data of the floors in Q to determine the receiver's location, which is called fine floor differentiation.

Fine floor differentiation is divided into two steps. In the first step, assumes that the receiver is located at floor i. Compare the signal strength of A_j received by the receiver with Γ_m. Define

$$t_{i,j}^{m} = \begin{cases} 1, \xi_j \in \Gamma_m \\ 0, \xi_j \notin \Gamma_m \end{cases} \tag{16}$$

where i floor identity and $i \in Q$, $j = 1, 2, \ldots, K_2$. The matching matrix of the *ith* floor can be defined as

$$\mathbf{T}_{i,j}(\Omega_0) = [\, t_{i,j}^{1} \quad t_{i,j}^{2} \quad \ldots \quad t_{i,j}^{M} \,] \tag{17}$$

Use the AP frequency $\mathbf{P}_{i,j}(\Omega_0)$ to weight $\mathbf{T}_{i,j}(\Omega_0)$, and the fine floor differentiation value (FFDV) E_i will be

$$E_i = \sum_{j=1}^{K_2} \mathbf{P}_{i,j}(\Omega_0) \cdot \mathbf{T}_{i,j}^{T}(\Omega_0) \tag{18}$$

FFDV of all floors in Q can be calculated. Then compare these values, the receiver should be located at the floor which has the maximum E value. By this method the floor on which the receiver is located can be determined with a high accuracy. Due to space constraint, the performance analysis is omitted here. After determining the floor and having the altitude information, the receiver can obtain the height difference h by comparing with the altitude of the previous floor where it is located. If the receiver is on the same floor at two consecutive measurement times, h is equal to 0. Thus, combining the pseudorange measurements, the receiver position can be found with the following equations

$$\begin{cases} \sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}+\delta t=\rho_c^1 \\ \sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2}+\delta t=\rho_c^2 \\ \sqrt{(x-x_3)^2+(y-y_3)^2+(z-z_3)^2}+\delta t=\rho_c^3 \\ \Delta u=h \end{cases} \tag{19}$$

where Δu is the difference in ENU coordinate system and $(x_i, y_i, z_i)(i = 1, 2, 3)$ is the ECEF coordinates of satellites; δt is the clock error and satellite and $\rho_c^i(i = 1, 2, 3)$ is the pseudorange for the receiver to satellite i. The conversion of ENU coordinates into ECEF coordinates is

$$\Delta u = \begin{bmatrix} \cos\phi\cos\lambda & \cos\phi\sin\lambda & \sin\phi & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta\delta t \end{bmatrix} \tag{20}$$

where ϕ is the longitude and λ is the latitude. The ECEF coordinates of the receiver can be calculated by combining (19) and (20).

2.3 Zone 3: GNSS and Wi-Fi Fusion Positioning

In general, receivers may see fewer satellites and receive weak Wi-Fi signals in areas close to the building. Hence it may not obtain highly accurate positions if using GNSS or Wi-Fi signals alone. Taking into account that these two systems locate the receiver independently, we can use group Kalman filter to fuse the positioning results of the two systems to improve positioning accuracy.

Assume the state transition matrix is an identity matrix. The state model will be

$$\mathbf{x}_k^- = \mathbf{x}_{k-1} + \mathbf{w}_k \tag{21}$$

where $\mathbf{x}_{k-1}$ is the optimal estimated position of the receiver at the previous moment, and $\mathbf{x}_k^-$ is a priori estimate of the current position, $\mathbf{w}_k$ is process noise.

The measured positions by Wi-Fi and GNSS signals are used to correct state prior estimates. Thus, the composite measurement vector is

$$\mathbf{y}_k = \begin{bmatrix} \mathbf{y}_{1,k}^T & \mathbf{y}_{2,k}^T \end{bmatrix} \tag{22}$$

where $\mathbf{y}_{1,k}$ and $\mathbf{y}_{2,k}$ are the position coordinates obtained by the Wi-Fi positioning and GNSS positioning method, respectively. Define Wi-Fi and GNSS positioning measurement matrix are $\mathbf{H}_{1,k}$ and $\mathbf{H}_{2,k}$. Then the group Kalman filter measurement matrix is

$$\mathbf{H}_k = [\mathbf{H}_{1,k}^T \quad \mathbf{H}_{2,k}^T] \tag{23}$$

Then the measurement model can be given by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_K \tag{24}$$

where $\mathbf{v}_k$ is measurement noise

$$\mathbf{v}_k = [\mathbf{v}_{1,k}^T \quad \mathbf{v}_{2,k}^T] \tag{25}$$

Let $\mathbf{R}_{1,k}$ and $\mathbf{R}_{2,k}$ be the covariance matrix of measurement noises of Wi-Fi and GNSS positioning, respectively, then the measurement noise covariance matrix is

$$\mathbf{R}_k = [\mathbf{R}_{1,k} \quad \mathbf{R}_{2,k}] \tag{26}$$

The gain matrix $\mathbf{K}_k$, the state estimation value $\mathbf{x}_k$ and the error covariance matrix $\mathbf{P}_k$ can refer to the standard Kalman filter equations. By group Kalman filter, positioning accuracy will be improved.

3 Simulation Results and Analysis

The floor differentiation algorithm is verified through the actual data measurements, and the GNSS and Wi-Fi positioning fusion algorithm is simulated and analyzed. We collect the Wi-Fi signals with the laptop of DELL INSPIRON which is configured with the operating system Windows10, and the software is WirelessMon V4.0. The Wi-Fi signals are measured in the fifth floor of Innovation Park Building in Dalian University of Technology, and the schematic diagram is shown in Fig. 2. The triangle symbols represent the AP locations. The receiver to be located will be

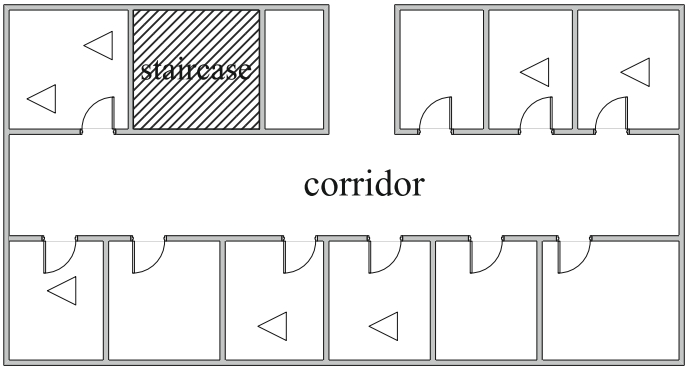


Fig. 2 The fifth floor of innovation park building

Table 3 Floor differentiation CFDV and FFDV values

Floor	CFDV/FFDV				
4th	1/–	3/–	2/–	3/0.227	2/–
5th	5/–	4/–	4/0.909	3/0.591	4/0.909
6th	4/–	3/–	4/0.5	2/–	4/0.545

'–' means that FFD is not required

in the corridor area which can be categorized as Z_2 according to the zone classification.

In offline phase, we collect the signals of APs in the 4th, 5th and 6th floor of the building, and build offline database of floor information according to Table 2, where K_1 is 10. The range of AP signal strength is [−90 dBm, −20 dBm], and the signal strength is divided into 7 sections, i.e., $M = 7$. Randomly select five unknown location points in each of the 4th, 5th and 6th floor, and sort the APs according to the signal strength. Select five APs with the largest signal strength, i.e., $K_2 = 5$. The signal strength of APs for the 15 points are matched with the data of the offline database to obtain the floor information.

We also simulate the positioning in Z_3. The measured coordinates obtained by using the Wi-Fi and GNSS signals are fused by the group Kalman filter.

Table 3 shows the floor differentiation CFDV and FFDV values of those points on the fifth floor. Table 4 shows the floor differentiation results of the 15 points. It

Table 4 Floor differentiation results

Floor	Differentiation result (Y/N)				
4th	Y	Y	Y	Y	Y
5th	Y	Y	Y	Y	Y
6th	Y	Y	Y	Y	Y

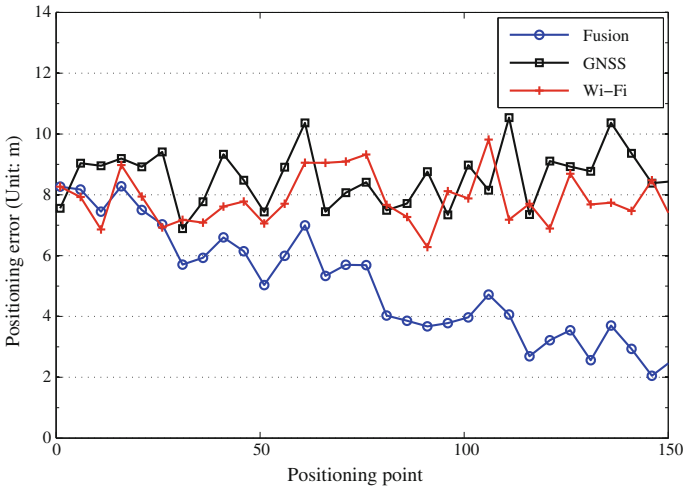


Fig. 3 Positioning errors

Table 5 Positioning errors (Unit: m)

	Wi-Fi	GNSS	Fusion
1	4.5	5	2.5936
2	4.5	10	3.4751
3	3.5	11	3.4893

is observed from Table 4 that all points are located correctly in terms of altitude, and it verifies the reliability of the proposed floor differentiation algorithm.

The GNSS, Wi-Fi and fusion positioning error are showed in Fig. 3. It is observed that the positioning accuracy is obviously improved after fusing. When Wi-Fi and GNSS positioning take different measurement errors, the group Kalman filter estimation errors are shown in Table 5. Compared with Wi-Fi positioning and GNSS positioning alone, the positioning error with group Kalman filter can be reduced by 37.5%.

References

1. Li Y, Zhang P, Niu X et al (2015) Real-time indoor navigation using smartphone sensors. In: International conference on indoor positioning and indoor navigation. IEEE, pp 1–10
2. Yan M, Chen X, Xu Y (2012) Wireless local area network assisted GPS in seamless positioning. In: International conference on computer science and electronics engineering. IEEE, pp 612–615
3. Jia M, Yang Y, Kuang L et al (2017) An indoor and outdoor seamless positioning system based on android platform. Trustcom/bigdatase/ispa. IEEE, pp 1114–1120
4. Anagnostopoulos GG, Deriaz M (2015) Automatic switching between indoor and outdoor position providers. In: International conference on indoor positioning and indoor navigation. IEEE, pp 1–6
5. Zou D (2016) Heterogeneous wireless systems based seamless positioning technologies. Harbin Institute of Technology
6. Jimenez AR, Seco F, Prieto C et al (2009) A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU. In: IEEE international symposium on intelligent signal processing. IEEE, pp 37–42
7. Hou B, Zhang X (2016) A dual-satellite GNSS positioning algorithm of high accuracy in incomplete condition. China Communications, pp 58–68
8. Ma L, Fan Y, Xu Y et al (2017) Pedestrian dead reckoning trajectory matching method for radio map crowdsourcing building in WiFi indoor positioning system. In: IEEE international conference on communications. IEEE, pp 1–6
9. Bisio I, Cerruti M, Lavagetto F et al (2014) A trainingless WiFi fingerprint positioning approach over mobile devices. IEEE Antennas Wirel Propag Lett 832–835
10. Xie G (2009) Principles of GPS and receiver design. Publishing House of Electronics Industry

An EMD—Based Fast Algorithm for INS Aided BDS Ambiguity Resolution

Junbing Cheng, Dengao Li, Wenjing Li and Jumin Zhao

Abstract Beidou satellite navigation system (BDS) signal is blocked during dynamic positioning which can lead to poor geometry structure of visible satellites and serious multipath effect. In this contribution, the advantages of inertial navigation system (INS), autonomy and short-term accuracy, are used to speed up the BDS ambiguity resolution (AR). However, the accuracy of general INS cannot meet the fixed BDS original carrier ambiguity requirement. BDS can broadcast three frequency observations, so the virtual observations of long wavelength and short wavelength are formed by linear combination, which make the INS precision meet the fixed ambiguity requirement and achieve high precision positioning, respectively. The short wavelength observations ambiguity due to the relatively large noise is difficult to reliably fixed. To solve this problem, Empirical mode decomposition (EMD) can be used to decompose the error of the short wavelength observations into two parts: systematic error and random error. Systematic error is used to correct the short wavelength ambiguity float solution and reduce the ambiguity search space, to achieve a reliable fixed purpose, finally solving the baseline solution. The experimental results show that the algorithm is more fast and reliable than the traditional algorithms, and more suitable for dynamic positioning in the complex environment.

Keywords BDS · Triple-frequency · INS · EMD

1 Introduction

Beidou satellite navigation system (BDS) is dynamically positioned in the harsh environment such as the city canyon. Because the signal is blocked by the building, the multipath effect is serious, and the geometric structure of the satellites is poor.

J. Cheng · D. Li (✉) · W. Li · J. Zhao
College of Information Engineering, Taiyuan University of Technology, Jinzhong, China
e-mail: lidengao@tyut.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_60

In order to realize high precision dynamic positioning in complex environment, the fast and reliable fixation of ambiguity is the key [1]. Parkins [2] used epochs information for partial ambiguity resolution and improves ambiguity fixed success rate. Tian et al. [3] used the ionosphere constraint to fix the ambiguity of narrow lane and improve the ambiguity fixed quality. These articles resolve ambiguity by digging the internal relations of the observations, and improve the reliability of the fixed ambiguities. However, the convergence time is long and can not meet the real-time requirement of dynamic positioning. INS is a kind of independent positioning system which has strong complementarities with BDS. It is used to assist BDS positioning as a research hotspot. Liu et al. [4] improved the combination system of INS and Global Positioning System (GPS) based on float solution to integer ambiguity, which reduces the convergence time and improves the stability of the positioning accuracy. Gao et al. [5] proposed to combine multi-satellite navigation system with micro inertial navigation system, which effectively increases the information redundancy of INS. These algorithms accelerate the resolution of ambiguity by means of INS, but the reliability of the ambiguity fixed is not effective.

The proposed algorithm uses INS to accelerate the fixing of BDS ambiguity, and introduces the EMD to decompose observation errors of narrow-lane, then obtains the systematic error to correct the ambiguity float solution, so as to improve the precision of float solution and enhance the reliability of ambiguity fixing.

2 INS Assisted BDS Fast Ambiguity Resolution

In the complex environment, multipath effect is more serious for the pseudorange than carrier, and sometimes the error of pseudorange is up to 20 m. The traditional method of calculating carrier phase ambiguity by pseudorange has a long convergence time and poor reliability. INS have strong autonomy, high output frequency and good short-term accuracy [6]. Using INS high quality positioning information to assist BDS AR can greatly reduce the influence of multipath, increase geometric information, improve the geometry structure of satellites, and provide reliable positioning information when BDS signal is blocked.

2.1 The Relationship Between INS Accuracy and BDS Ambiguity

Assuming the receiver is j, the satellite is q, then the carrier phase observation equation is expressed as:

$$\lambda\varphi_j^q = \rho_j^q + g_j^q + \delta t_j - \delta t^q - I_j^q + T_j^q + \lambda N_j^q + e_{\varphi,j}^q \tag{1}$$

where λ denotes the wavelength, φ_j^q denotes the phase value, ρ_j^q denotes the geometric distance, N_j^q denotes the integer ambiguity, δt_j denotes the receiver clock error, δt^q denotes the satellite clock error, I_j^q denotes the ionospheric delay, T_j^q denotes the tropospheric delay, $e_{\varphi,j}^q$ denotes the observed noise.

Supposing that the base station is i and the reference satellite is p, the double difference in the short baseline condition can eliminate common errors between stations and satellites, such as satellite clock error and receiver clock error, and also greatly reduce orbit error, ionospheric error and tropospheric error. After the double difference, ambiguity should still maintain the characteristic of the integer cycle, then, available carrier phase double differential observation equation is:

$$\lambda\nabla\Delta\varphi_{i,j}^{p,q} = \nabla\Delta\rho_{i,j}^{p,q} - \lambda\nabla\Delta N_{i,j}^{p,q} + e_{\nabla\Delta\varphi} \tag{2}$$

It can be shown from [2] that if the INS error accuracy meets the requirement, the positioning information of INS can be converted into the corresponding geometric distance ρ_I to substitute the equation. For convenience, the superscript and subscript of double difference can be omitted:

$$\lambda\nabla\Delta\varphi = \nabla\Delta\rho_I - \lambda\nabla\Delta N + e_{\nabla\Delta\varphi} \tag{3}$$

Base station i does not get the INS positioning parameter:

$$\nabla\Delta\rho_I = \nabla\rho_I - \nabla\rho_i \tag{4}$$

The equation for estimating the integer ambiguity $\nabla\Delta N$ becomes:

$$\nabla\Delta\hat{N} = \nabla\rho_I/\lambda - \nabla\rho_i/\lambda - \nabla\Delta\phi \tag{5}$$

Assuming that the geometric distance error of base station i is 0, the estimation accuracy and speed of AR will be affected by the error of INS positioning parameters, multipath error, atmospheric residual error and incomplete modeling error. If rounding method is used to fix ambiguity, the error of Integer ambiguity $\nabla\Delta\hat{N}$ must be less than 0.5 cycles:

$$\left|\delta_{\nabla\rho_I} - \delta_{\nabla\Delta\varphi}\lambda\right| < 0.5\lambda \tag{6}$$

Assuming that the error of the carrier phase observations is 0.01 cycles, according to the error theory, the error of the double-difference observation is 0.02 cycles, the confidence interval is set to 99.97%.

Table 1 Frequency and wavelength parameters of BDS

Signal	Frequency/MHz	Wavelength/m
B1	1561.098	0.192
B2	1207.140	0.248
B3	1268.520	0.236

$$\left|\delta_{\nabla\Delta\varphi}\right| = 0.06\lambda \tag{7}$$

$$\left|\delta_{\nabla\rho_I}\right| < 0.44\lambda \tag{8}$$

Obviously, the equivalent distance of accumulating error for INS within one epoch must be less than 0.44 times the carrier wavelength. According to Table 1, the shortest wavelength of the three frequency points broadcast by the BDS is 0.192 m, the precision of INS must be less than 0.08448 m to meet the requirement of rounding method to directly fix the ambiguity. However, the precision of the general INS is not satisfactory.

2.2 BDS Triple-Frequency Ambiguity Fixed Model

The essence of ambiguity fast fixing is to reduce the range of alternative combinations, that is to narrow the search space. The size of the search space is determined by the variance-covariance matrix of the ambiguity float solution:

$$W \propto \frac{\delta_{\nabla\rho_I}}{n\lambda} \tag{9}$$

where W denotes ambiguity alternative range, $\delta_{\nabla\rho_I}$ denotes the INS precision, λ denotes carrier wavelength, n denotes the number of epochs, $\propto$ denotes the proportional relationship. In dynamic environment, the number of epochs will not be too large, INS accuracy is determined, and the search space can only be reduced by expanding the wavelength. BDS broadcast the observations information of three frequency points, and it not only improves the information redundancy, but also obtains the virtual observations of various wavelength by linear combination of triple-frequency [7]. The long wavelength can make the INS precision meet the rounding fixed ambiguity, the short wavelength can achieve high-precision positioning.

BDS triple-frequency linear combination observations is expressed as:

$$\varphi_{a,b,c} = a\varphi_1 + b\varphi_2 + c\varphi_3 \tag{10}$$

Table 2 Wavelength and noise of combinations for BDS

$\varphi_{a,b,c}$	Wavelength/m	Error
$\varphi_{1,0,0}$	0.19	δ_0
$\varphi_{0,1,0}$	0.24	δ_0
$\varphi_{0,0,1}$	0.23	δ_0
$\varphi_{0,1,-1}$	4.88	1.41 δ_0
$\varphi_{1,-1,0}$	0.84	1.41 δ_0
$\varphi_{1,0,-1}$	1.02	1.41 δ_0
$\varphi_{1,1,0}$	0.10	1.41 δ_0
$\varphi_{1,0,1}$	0.10	1.41 δ_0

where φ_1, φ_2 and φ_3 are triple-frequency carrier phase of BDS respectively, a, b and c are the corresponding combination coefficients [11]. The frequency, the wavelength and the ambiguity of combined observations are respectively:

$$f_{a,b,c} = af_1 + bf_2 + cf_3 \tag{11}$$

$$\lambda_{a,b,c} = v_c / f_{a,b,c} \tag{12}$$

$$N_{a.b.c} = aN_1 + bN_2 + cN_3 \tag{13}$$

where f denotes the frequency, λ denotes the wavelength, N denotes the ambiguity and v_c denotes the speed of light.

Assuming that the observations noise at the three frequencies are uncorrelated and equal to each other, the noise of combined observations is:

$$\delta_{a,b,c} = \sqrt{a^2 + b^2 + c^2} \cdot \delta_0 \tag{14}$$

Table 2 is the wavelength and noise of the common BDS triple frequency combination. Triple-frequency AR is generally fixed in turn from long to short by the pseudorange according to the wavelength of the combined virtual observations. The longest wavelength is called an ultra-wide lane (EWL), which is the first to be fixed; the middle wavelength is called a wide lane (WL) and then fixed; the shortest wavelength is called the narrow lane (NL) and finally the fixed [8].

2.3 Combination Observations Ambiguity Fixed

The triple-frequency combination $(0, 1, -1)$ is taken as EWL, the combination $(1, 0, -1)$ is taken as WL13, and the wavelength of combination (1, 1, 0) is smaller than the original carriers, taken as NL12. It should be noted that the combination of three frequency points can take full advantage of redundant information of BDS tri-band observations.

For EWL and WL13 with the long wavelength, the error range of INS is sufficient to meet their precision requirements. The use of high-quality positioning information provided by INS can get a high fixed success rate by the rounding method. However, the wavelength of NL12 is only 10 cm, the observations noise is as high as 1.41 δ_0, and the multipath affects seriously [9]. The success rate of direct fixation is small.

3 Using EMD to Improve NL12 Ambiguity Fixed Reliability

The NL12 wavelength is relatively short and the noise is large, so it is difficult to be fixed reliably. In addition to high-frequency noise such as thermal noise, there are also low-frequency errors caused by multipath and other factors. EMD can decompose high-frequency errors and low-frequency errors and then correct narrow-lane ambiguity float solution with low-frequency errors.

3.1 EMD Introduction

As a tool for local time-frequency analysis, EMD has better performance of self-adaptability. Instead of decomposing the signal into several sin and cos functions with different frequency, it decomposes the signal into a finite number of intrinsic modulus functions (IMF) according to the characteristic of the signal itself, which is suitable for analyzing and processing non-stationary signal [10]. The essence of the EMD is to decompose the signal into high frequency and low frequency components, and then decompose the low frequency signal to the high and low frequency, until it meets the requirement.

3.2 EMD for NL12 Ambiguity Fixing

The errors caused by various factors in the dynamic positioning of BDS are collectively referred to as the integrated error, including multipath error with time-varying characteristics and slow changes, high frequency white noise errors caused by thermal noise, it also includes unmodelled and other factors caused by the error. These errors make the integrated error have typical non-stationary characteristic, suitable for processing with EMD. The high-frequency error caused by thermal noise and other factors is called random error. The low-frequency error such as multipath and incomplete modeling is called systematic error. The main high-frequency errors decomposed by EMD mainly exist in low-order intrinsic

mode functions. The low-frequency errors caused by multipath mainly exist in higher-order intrinsic mode functions and trend residuals. Then the NL12 observation equation is:

$$\varphi_{(1,1,0)} = \rho_{1,0,-1} - \lambda_{1.1,0} N_{1,1,0} + e_{1,1,0} \tag{15}$$

The integrated error is decomposed into systematic error and random error:

$$\varphi_{(1,1,0)} = \rho_{1,0,-1} - \lambda_{1.1,0} N_{1,1,0} + e^{\text{sys}}_{1,1,0} + e^{\text{ran}}_{1,1,0} \tag{16}$$

If the systematic error obtained by EMD decomposition as a known quantity is substituted into Eq. (16), the noise of the combined observations will be significantly reduced, and the accuracy of the ambiguity float solution will be improved, which is equivalent to narrowing the search space for NL12 ambiguity. Then use the determined WL13 ambiguity reliably fixed NL12 ambiguity and finally solve the baseline vector.

3.3 The Overall Block Diagram of The Algorithm

See Fig. 1.

4 Experimental Design and Analysis

In order to verify the feasibility and effectiveness of the proposed algorithm, experiments were conducted on a downtown street. The data sampling interval BDS triple-frequency receiver is 1 s, the data sampling length is 3000 epochs, the satellite cutoff height angle is 10°; The sampling rate of INS is 100 Hz, the gyro zero offset is 0.5°/h, and the accelerometer zero offset is 0.1 mg. In this experiment, the decomposition scale of EMD is 7 orders, the first 5 orders are regarded as random error, and the last 2 orders and trend residuals are used as systematic error. The F-Ratio criterion is used to determine whether the ambiguity is fixed successfully. If the F-Ratio is greater than 1.5, the ambiguity is successfully fixed.

4.1 Experimental Scheme Design

According to the innovation of this article, we design three kinds of comparison experiment schemes:

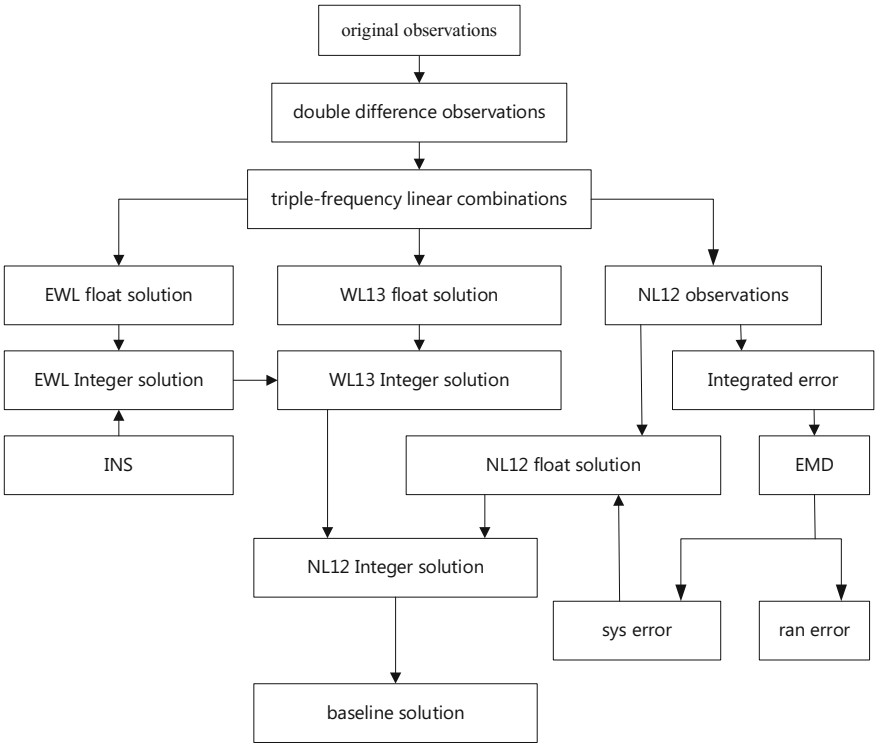


Fig. 1 EMD—based fast algorithm of INS aided BDS ambiguity resolution

Scheme 1: The proposed algorithm, that is, with the aid of INS, the EWL and WL13 ambiguities of BDS are fixed, and then the systematic error of NL12 is decomposed by EMD to correct the ambiguity float solution, and finally the NL12 ambiguity is fixed.
Scheme 2: Instead of INS assisting, the ambiguity is solved by pseudorange, the other steps are the same as the scheme 1.
Scheme 3: N12 ambiguity float solution without systematic error correction but directly fixed, the other steps are the same as the scheme 1.

4.2 Experimental Results and Analysis

4.2.1 Scheme 1 and 2 Positioning Error and the First Fixed Time

The left and right figures of Fig. 2 show the position error of scheme 2 and scheme 1, respectively. Comparison of the two figures, scheme 1 for the first time

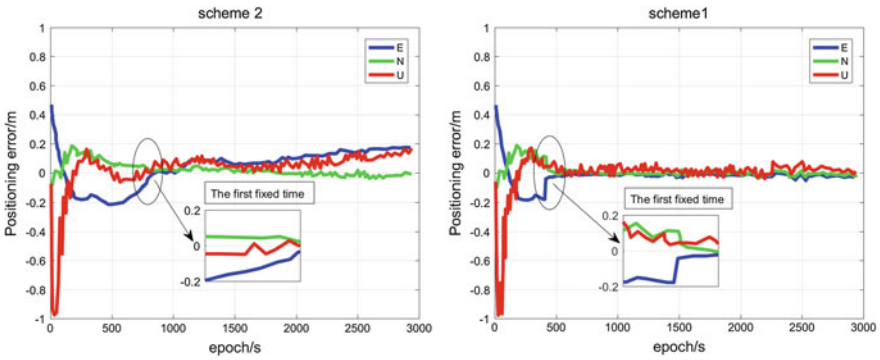


Fig. 2 Positioning error of scheme 2 and scheme 1

fixed was significantly less than the scheme 2. After the first fixer of ambiguity, the positioning error of scheme 1 is smaller than that of scheme 2. This is due to INS providing additional positioning information, demonstrating that INS can speed up the fixing of ambiguity.

4.2.2 Ambiguity Fixed Success Rate of EWL, WL13 and NL12

We choose BDS No. 1 satellite, No. 6 satellite and No. 10 satellite experiment in accordance with the scheme 3, their EWL, WL13 and NL12 ambiguity fixed success rate are shown in Fig. 3. The fixed success rate of EWL and WL13 for the three satellites are above 99%, while NL12 only about 70%. This is because the wavelength of NL12 is relatively short and the noise is relatively large, and the multipath effect is serious.

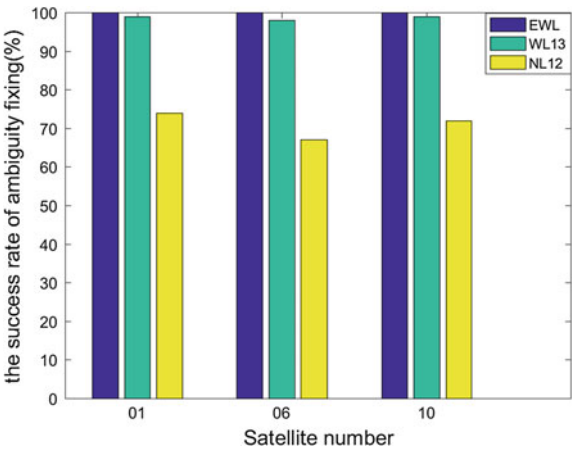


Fig. 3 Ambiguity fixed success rate of EWL,WL13 and NL12

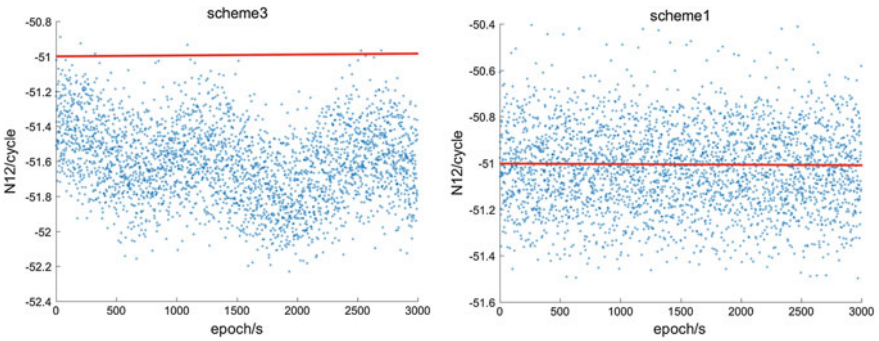


Fig. 4 Ambiguity float solution of NL12 before and after error correction

4.2.3 Accuracy of Float Solution Before and After NL12 Error Correction

The left and right figures of Fig. 4 are the ambiguity float solution of BDS No. 6 satellite conducted according to Scheme 3 and Scheme 2, after the systematic error correction, float solution accuracy improved significantly.

5 Conclusion

We investigated noise characteristics of BDS observations during dynamic positioning in the complex environment, and take into account the need of fast and reliably ambiguity fixing. The high-quality positioning information of INS is used to assist in fixing the ambiguity, the experimental results show that INS can indeed reduce the first fixing time. Then the systematic error obtained by the EMD decomposition are used to correct the float solution of narrow lane to improve the accuracy and help to fix the narrow lane ambiguity reliably. The experimental results show that the float solution accuracy is significantly improved.

Acknowledgements The study is supported by The National High Technology Research and Development Program ("863" Program) of China (2015AA016901): High linearity laser diode array and high saturation power photodiode array; The General Object of National Natural Science Foundation (61772358) Research on the key technology of BDS precision positioning in complex landform; International Cooperation Project of Shanxi Province (Grant No. 201603D421012): Research on the key technology of GNSS area strengthen information extraction based on crowd sensing; The General Object of National Natural Science Foundation under Grants (61572347): Resource Optimization in Large-scale Mobile Crowdsensing: Theory and Technology.

References

1. Teunissen PJG (1995) The least squares ambiguity decorrelation adjustment: a method for fast GPS integer estimation. J Geod 70:65–82
2. Parkins A (2011) Increasing GNSS RTK availability with a new single-epoch batch partial ambiguity resolution algorithm. GPS Solut 15:391–402
3. Tian Y, Zhao D, Chai H, Wang S (2017) BDS ambiguity resolution with the modified TCAR method for medium-long baseline. Adv Space Res 59:670–681
4. Liu S, Sun F, Zhang L; Li W, Zhu X (2016) Tight integration of ambiguity-fixed PPP and INS: model description and initial results. GPS Solut 20:39–49
5. Gao Z, Zhang H, Ge M, Niu X, Shen W, Wickert J, Schuh H (2017) Tightly coupled integration of multi-GNSS PPP and MEMS inertial measurement unit data. GPS Solut 21:377–391
6. Sasani S, Asgari J, Amiri-Simkooei AR (2016) Improving MEMS-IMU/GPS integrated systems for land vehicle navigation applications. GPS Solut 20:88–91
7. Li J, Yang Y, He H, Guo H (2017) An analytical study on the carrier-phase linear combinations for triple-frequency GNSS. J Geod 91:151–166
8. Han H, Wang J, Wang J et al (2017) Reliable partial ambiguity resolution for single-frequency GPS/BDS and INS integration. GPS Solut 21:251–264
9. Li B, Shen Y, Feng Y, Gao W et al (2014) GNSS ambiguity resolution with controllable failure rate for long baseline network RTK. J Geod 88:99–112
10. Dai W; Huang D, Cai C (2014) Multipath mitigation via component analysis methods for GPS dynamic deformation monitoring. GPS Solut 18:417–428
11. Tang W, Deng C, Shi C, Liu J (2014) Triple-frequency carrier ambiguity resolution for Beidou navigation satellite system. GPS Solut 18:335–344

Floor Recognition Based on SVM for WiFi Indoor Positioning

Shuai Zhang, Jiming Guo, Wei Wang and Jiyuan Hu

Abstract With the rapid development and popularization of WiFi technology, indoor positioning technology based on WiFi has become a research hot spot. At present, a variety of complex structures come into being along with the gradual improvement of people's living standards. The two-dimensional positioning of the room has been unable to meet people's needs. In this paper, a method of WiFi indoor positioning for floor recognition based on SVM (support vector machine) classification is proposed. Due to the obvious change during the WiFi signal through the wall or floor, floor identification is realized quickly and accurately by using SVM. We collect the RSS data of each floor with smart phone, and utilize the 10-fold cross-validation method to train classifier model and evaluate classification accuracy. The experimental results show that the high accuracy of floor discrimination, and the 99.09% floor recognition accuracy can be obtained.

Keywords WiFi · RSS · Floor recognition · Support vector machine
Indoor positioning

1 Introduction

With the scientific development and the improvement of people's living standards, the demand for LBS (location based services) is increased. Accuracy of positioning using GPS (global positioning system) has been able to meet the need of people in outdoor environment currently, but weak GPS signal can not penetrate buildings well for localization in the indoor environment. GPS is not available in indoor environment. With the rapid development of wireless network technology and the popularity of WLAN (wireless local area networks), the indoor positioning technology based on WiFi has become a hot topic of research [1–3].

S. Zhang (✉) · J. Guo · W. Wang · J. Hu
School of Geodesy and Geomatics, Wuhan University, Wuhan, China
e-mail: 2016102140015@whu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_61

At present, the structure of the city is complex and the structure of the floor is varied. People can not easily find their location in a complex indoor environment. Simple two-dimensional location can no longer meet people's needs. People are more interested in three-dimensional location. Floor recognition plays an important role in indoor positioning. On the basis of the known plane position, if we recognize the floor, the indoor location can be extended from two dimensions to three dimensions.

The built-in barometer of a smart phone can identify the floor, but the indoor environment is volatile. The indoor environment is closed relative to outdoor environment. Especially in winter, the air conditioners increase temperature of indoor environment. Which causes the change of air pressure in the environment of closed doors and windows [4]. In addition, flow of air has a certain influence on the atmospheric pressure. So the barometer is not very good at recognizing the floor.

For this demand, Li et al. [5] presented a method of identifying the floor of the building, which based on WiFi RSSI (received signal strength indication). The method first records all APs (access points) belonging to floor. By statistics of the number of AP greater than the threshold of RSSI in each floor, if the number of AP is the most in same floor, the floor is defined as the current floor. The location method reduces computation, but when AP number of mark up reduce or AP move to other floors, which will have an impact on the results.

Additionally, WiFi fingerprints can also be used to identify the floors. This way of floor discrimination based on WiFi fingerprints is to scan all the fingerprint points and calculate the Euclidean distance from test points to all fingerprint points. The floor of minimum distance of fingerprint as the floor of test point, but the method of using WiFi fingerprint to identify the floors is a lot of time-consuming. In order to resolve this problem, Deng et al. [6] proposed a method based on K-means algorithm for identification of the floor. This method uses spatial clustering algorithm, greatly reduces the computation time. But this method needs to determine the number of clustering K in advance. The value of k has a great influence for the positioning accuracy. The method of fingerprint point is still laborious.

SVM is a better classification method in pattern recognition. The support vector machine has the characteristics of high classification precision and robust to noise. Tran et al. [7] used support vector machine to recognize the pattern of pedestrians; Zhang et al. [8] identified the movement patterns of the old people in a particular group.

In this paper, SVM is applied to floor discrimination. According to the difference of WiFi signal strength between floors, a floor recognition method based on SVM classification is proposed. In this method, the SVM classification algorithm costs less time, gets higher accuracy, which reaches the 99.09% about the floor recognition. The method only need to collect a certain amount of AP RSSI value each floor.

2 The Principles of SVM

This chapter mainly discusses the principles of SVM and the process of establishing classifier model. To establish a classifier, we need to consider the linear selection of kernel function of classifier, and cost factor selection of SVM. The cost factor directly affects the classification result. In the process of establishing the classifier, we use the 10-fold cross-validation method. The method is a general method which can train the classifier well. Finally, the classification effect of SVM is evaluated by F1 indication.

2.1 Linear SVM Principle

SVM is a popular model of two classes classifier. The model is defined as the linear classifier with the largest interval in the feature space [9]. The support vector machine contains the kernel function. This paper utilizes the linear kernel function to support vector machine classification. The specific principle is as follows:

There is a series of training data $(\vec{x}_1, y_1), (\vec{x}_2, y_2), \ldots, (\vec{x}_n, y_n)$, $y_i (i = 1, 2, \ldots n)$ equal to 1 or -1 to represent different categories. $\vec{x}_i$ is a vector making up of different features. There are a lot of hyper planes to separate the two types of sample data. In Fig. 1, the pentagram and the circular scatter plot represent two kinds of sample data respectively. Line B and line C are hyper planes, but the maximum gap of the hyper plane is only one, namely optimal hyper plane. SVM uses the maximum distance to find an optimal hyper plane $f(x_i) = \vec{w} \cdot \vec{x}_i + b$. Line C is the best hyper plane in Fig. 1, line A and line D are the extreme hyper plane, line E is the largest hyper plane interval.

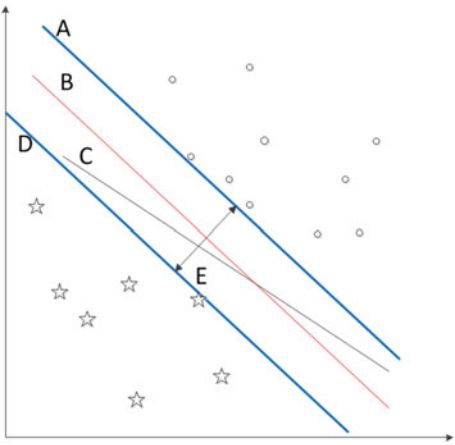


Fig. 1 SVM classification

In the SVM classification method, the maximum geometric interval of the two classes can be weighed by two times of the reciprocal of $||\vec{w}||$. The greater interval the smaller $||\vec{w}||$. In order to avoid the data falling into the interval region, an additional constraint is added as (1).

$$y_i = 1 : \vec{w} \cdot \vec{x}_i + b \geq 1$$
$$y_i = -1 : \vec{w} \cdot \vec{x}_i + b \leq 1 \tag{1}$$

Combining Eqs. (1) and (2) can be obtained:

$$1 - \mathrm{y}_i \cdot (\vec{w} \cdot \vec{x}_i + b) \leq 0 \tag{2}$$

To avoid the data points falling into the interval region, we ignore a number of data with relatively large deviations. So we consider the loss function as (3):

$$\sum_{i=1}^{n} \max(0, 1 - \mathrm{y}_i \cdot (\vec{w} \cdot \vec{x}_i + b)) \tag{3}$$

Taking into account the hyper plane interval and loss function, we minimize the Eq. (4), and C is a cost parameter:

$$\frac{1}{2}||\vec{w}||^2 + C \sum_{i=1}^{n} \max(0, 1 - \mathrm{y}_i \cdot (\vec{w} \cdot \vec{x}_i + b)) \tag{4}$$

After the minimization formula (5), the sum of the $\vec{w}$ and b values can be obtained and the (5) can be used for classification.

$$\mathrm{sgn}(\vec{w} \cdot \vec{x}_i + b), \mathrm{sgn}(x) = \begin{cases} -1 : x < 0 \\ 0 : x = 0 \\ 1 : x > 0 \end{cases} \tag{5}$$

2.2 Cost Selection

In the process of finding line E above, there are some discrete points inevitably which have large deviations. If all the sample points are separated these discrete points largely determine the objective function which is Eq. (2). So that if a discrete sample point is wrong, losses of the objective function will be very heavy. We introduce a cost factor to decide the importance of discrete points. We set a cost factor value before classification. Then we use the test data to examine the result which is good or bad. If the effect of classification is not good, we can get another classifier by changing the value of cost factor C until the best classification results

are obtained. The selected process of C value is a process of parameter optimization. After repeated tests, the C value in this paper is 0.01, which makes the floor Recognition with the highest accuracy.

2.3 K-Fold Cross-Validation

When training linear SVM model, the training data needs to be divided into two parts: a training set and an evaluation set. The training set is used to train the linear SVM model, and the evaluation set is used to evaluate the quality of the model. It is impossible to make full use of the existing data by simply dividing all the training data into two subsets. Therefore, the k-fold Cross-validation is used in the data division. It is a method which used to prevent the over fitting of the model when the model is too complex. In the k-fold cross-validation, the training data is randomly divided into k parts, $k - 1$ is used as the training set, and the remaining 1 part is used as evaluation sets. So after the K cycles, all the data has been used as an evaluation set. We can gain the average of K results. This paper uses the 10-fold cross-validation method.

2.4 Evaluation of SVM Classification Results

We usually use test data to evaluation the classifier model. There are many criteria for evaluation of classification, such as precision, the recall and the accuracy. But sometimes the precision and recall will appear contradictory, we adopt comprehensive evaluation. Index F1 is often used. When the formula (6) parameter a = 1, we can get F1, that is the precision and recall of the harmonic mean. Table 1 is the confusion matrix of samples A and B.

In the table:

TP: the class A is predicted to be class A.
FN: the class A is predicted to be class B.
FP: the class B is predicted to be class A.
TN: the class B is predicted to be class B.

$$F = \frac{(a^2 + 1)P * R}{a^2(P + R)} \tag{6}$$

Table 1 The confusion matrix of A and B

Observation	Prediction	
	A	B
A	TP	FN
B	FP	TN

Using the confusion matrix of the sample, the indications are as follows:

$$\begin{aligned} \text{precision} &= TP/(TP+FP) \\ \text{recall} &= TP/(TP+FN) \\ \text{F1} &= 2*(\text{P}*\text{R})/(\text{P}\ +\ \text{R}) \\ \text{accuracy} &= (TP+TN)/(TP+TN+FP+FN) \end{aligned}$$

3 Analyses of WiFi Signal Characteristics

In order to use the WiFi signal for the right floor recognition, we first need to understand the characteristics of WiFi signal. In the experiment we find the WiFi signal fluctuates with multi-path. With the change of distance, WiFi signal gradually decays. The body's block influence WiFi signal strength, and the strength of WiFi signal will be weakened when it goes through the wall. We can better use the WiFi signal for floor identification through the experiment.

3.1 Description of WiFi Signal Fluctuation

In order to verify the WiFi signal fluctuation, We set up an AP in the laboratory. The phone is three meters away from the AP. We collect eight minutes of data, and get an RSS data in a second. In Fig. 2, It can be seen that the range of the signal fluctuation is about 10 dBm. So we need to take a period of time on a collection point to ensure that the signal is closer to the truth.

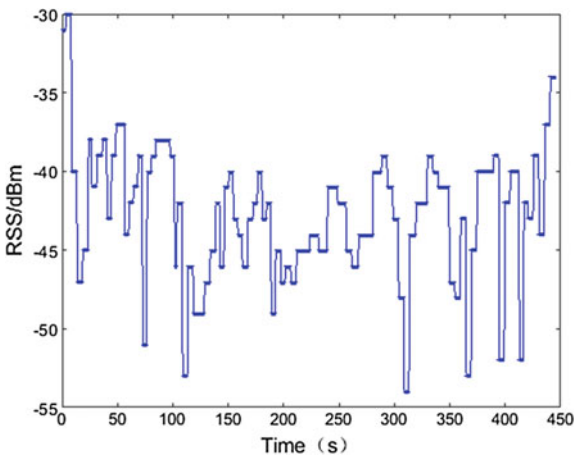


Fig. 2 The change of WiFi signal with the time in constant position

3.2 The Influence of the Body Block for WiFi Signal

The experiment was carried out in a 50 m^2 room. We kept the same position at a collection point. The RSS was received in different directions from same AP. The direction was divided into back to AP direction, direction to AP and side to AP direction. The RSS was collected in each direction for 5 min, and the histogram of the RSS distribution is shown in Fig. 3. Under the body block, the signal attenuation is more serious, the signal fluctuation range is large in the side direction. The most concentrated RSS in the face of AP direction. Therefore, data is collected in all directions to ensure that the training samples are more abundant.

3.3 Characteristics of WiFi Signal Through Wall

In order to verify RSS change of the AP through wall, we design two experiments. One experiment is that we receive signal from outdoor environment to indoor environment when AP is set outside, and the other is from outdoor environment to indoor environment when AP is in the laboratory. In Fig. 4, the change of the RSS is abrupt when people walk into the laboratory or go out of the laboratory. The biggest change of signal strength is about 30 dBm, and the average change is about 20 dBm. The degree of change can satisfy the identification between floors.

4 Building Process of SVM Classifier Model

The classification process is mainly divided into three steps: data acquisition, data samples preprocessing, classifier training and testing. The samples need preprocessing for achieve better classification results. In the training phase of the classifier, we put the samples into two parts, the training data (80%) and test data (20%).

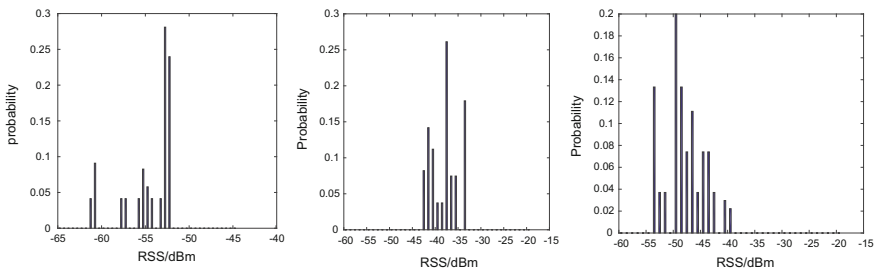


Fig. 3 **a** Back to AP. **b** Direction to AP. **c** Side direction to AP

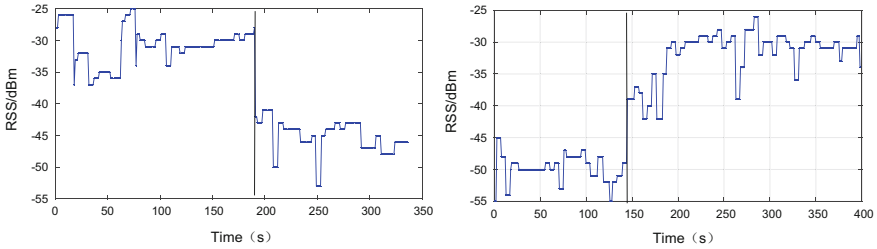


Fig. 4 First picture shows outdoor environment to indoor environment when AP in outdoors environment. Second picture shows outdoor environment to indoor environment when AP in indoor environment

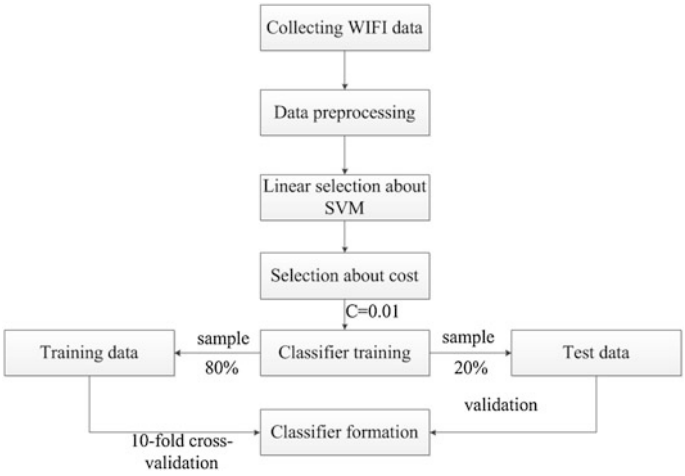


Fig. 5 The process of establishing classifier model

In this paper, we select 8 signal features as a sample to train the classifier. Classifier is evaluated by test data. The detailed process of the establishment of the classifier is shown in Fig. 5.

5 Experimental Results and Analysis

We collected RSS data from the first floor to the third floor in School Of Geodesy and Geomatics, Wuhan University. Through data preprocessing and classifier training, we made an evaluation of the classification results. The recall and precision of the method were over 99% about floor recognition.

5.1 Experimental Deployment

The experiment was carried out in the office building of School Of Geodesy and Geomatics, Wuhan University. The experimental area covered three floors of office buildings. There were three APs on each of the first floor and third floor. Two APs were deployed on the second floor. The APs of each floor were basically the same position. The position of each AP on the third floor is shown in Fig. 6. The position of red dot represents position of each AP in the figure.

Using MI4 smart phone (Android 4.4 based MIUI6 operation system) collected data in the corridor, collection position of each floor was basically the same. For example, collection position on the third floor is shown in Fig. 7, the position of the red dot represents the position of each AP, and the blue symbol represents the collection point. The interval between collection points is about 1.2 m along the corridor direction, and the interval between collection points is about 0.6 m in a direction perpendicular to corridor. 45 collection points are distributed in the corridor. Taking into account the body shelter and signal influence, we collect 30 s data at each point and change one direction in each 10 s. There is a regular collection along the S route from left to right. The mobile phone is placed horizontally when the data is collected. This method of collection makes the training samples as rich as possible.

In this experiment, we collect RSS among three floors for half an hour on each floor. One sample contains signal features of 8 APs, 1645 samples are collected

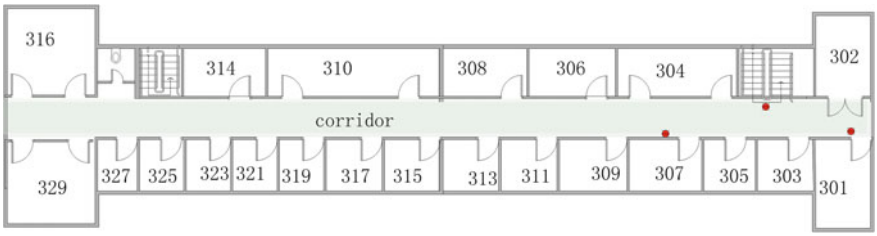
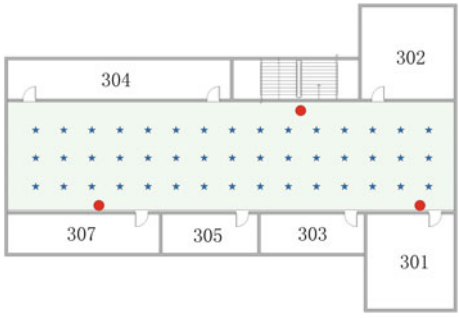


Fig. 6 AP distribution on the third floor

Fig. 7 Fingerprint distribution on the third floor



from the first floor, 1675 samples from second floor and 1679 samples from three floor. 4999 samples are collected in total. The number of samples on each floor is as consistent as possible to ensure the reliability of the classification.

5.2 Data Preprocessing

RSS of the AP may not be collected at a certain moment due to the environmental instability. In order to ensure the number of features of each sample, RSS at a time when the signal can not be received is replaced by a default value of the −95 dBm.

5.3 Data Processing and Accuracy Evaluation

Figure 8 shows the curve of the correct rate of training and cross-validation. The two curves converge to about 98.6% with the increase of the training samples. The stable classification results can be obtained at the point of convergence, which can be used to identify the floors well. As shown in the figure, the curve begins the convergence when the number of samples reaches 1500. The speed of convergence is fast. We take less time to identify the floor than other methods of identification floor.

By testing the classifier model with the test data, we get the confusion matrix of Table 2. In Table 2, there is a very little difference between the observation class and the predicted class. That shows that the predicted effect is good. The recall and precision are over 99% about floor recognition, and the lowest value of F1 reaches 98.67%. Results of evaluation indication show that the method of SVM classification using WiFi signal can achieve the desired effect for the floor recognition.

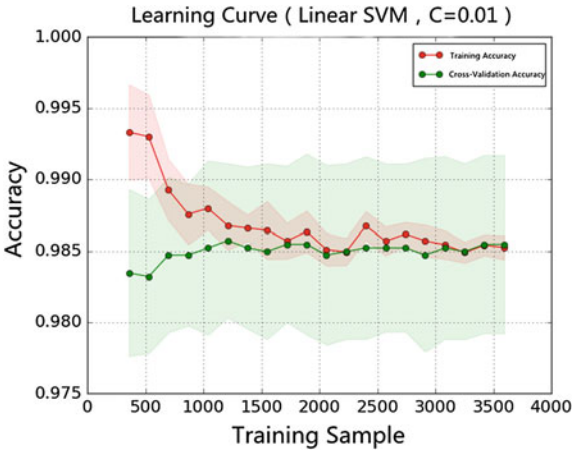


Fig. 8 The curves about the process of learning

Table 2 The confusion matrix and evaluation indication of floor recognition

		Predicted class				
		First floor	Second floor	Third floor	Recall	F1
Actual Class	First floor	330	0	0	1.0000	0.9940
	Second floor	0	326	2	0.9939	0.9924
	Third floor	4	3	334	0.9795	0.9867
	Precision	0.9880	0.9909	0.9940	/	/

6 Conclusions

In this paper, a method of floor identification with WiFi signal strength is implemented based on SVM classification. We use mobile phone to collect WiFi signal. According to the difference of WiFi signal strength between floors, we utilize the linear kernel function for support vector machine classification. The paper uses the 10-fold cross-validation method to train classifier model and evaluate classification accuracy. In this method, the SVM classification algorithm costs less time and gets higher accuracy. The method has a very good result for floor identification. The recall and precision are over 99% about floor recognition, the lowest value of F1 reaches about 98.7%. Floor recognition accuracy is 99.09% as a whole. But there is some space to improve, floor discrimination can be combined with the barometer to further improve the accuracy of floor identification.

References

1. Xi R, Li Y, Hou M (2017) Indoor positioning method on WiFi. Comput Sci 42(2):1–4
2. Shi G, Wang B, Wu B (2015) An overview of the indoor location methods based on WiFi and mobile intelligent terminals. Comput Eng 41(9):39–44
3. Karegar P A(2017). Wireless fingerprinting indoor positioning using affinity propagation clustering methods. Wirel Netw
4. Wang X, Zhang J (2017) Design and implementation of automation instrument. Air Press Altitude Measuring Syst 38 (8):59–63
5. Li H, Qi D, Wang Q (2014) A floor discrimination method based on the RSSI difference between floors
6. Deng Z, Wang W, Xu L (2012) A WLAN indoor positioning floor. K-means Algorithm Based Criterion Softw 33 (12):114–117
7. Tran DN, Phan DD (2016) Human activities recognition in android smartphone using support vector machine. In: 2016 7th international conference on intelligent systems. Modelling and aimulation
8. Zhang Y, Ren T, Luo Y (2016) Research on human motion pattern recognition. Comput Meas Control of 2016, 24 Based Multisen Data fusion 1:303–307
9. Eluoushi M, Georgy J, Noureldin A et al (2016) Motion mode recognition for indoor pedestrian navigation using portable devices. IEEE Trans Instrum Meas 65(1):208–221

Accuracy Analysis of Spaceborne Heterologous Multi-view Stereo Positioning with SAR Image Substitute of CCD Image

Ying-ying Li, Hao Wu, En Long, Xueli Chang and Zhixin Li

Abstract Data source access conditions of traditional homologous CCD stereo are very harsh. With the greatly improvement of spaceborne SAR image resolution, in particular its all-time all-weather observation capability and height sensitive side-view feature, SAR has become an effective complement to optical remote sensing, thereby which greatly increase the possibility of stereo conformation. And heterologous stereo precision is expected to reach or exceed the homologous optical stereo. However, because its data source is not uniform and precision factors are multiple, a comprehensive error analysis law should be put forward to support if the heterologous stereoscopic would be put into application. But current studies could not present clear conclusion. In this article, first the equivalent relationship between SAR strict geometry model and the rational polynomial model (RPC) is derived. Through the whole link error propagation modeling, we simulate and evaluate the quantitative relationship between the positioning accuracy and influencing factors such as the source image angle, corresponding points matching accuracy. Then the theoretical accuracy and superiority achieved by heterogeneous stereoscopic can be analyzed. In-orbit data tests are conducted to fully verify our theory correctness, and the error rules under different data combinations are detailed analyzed, which will provide a reliable basis for the heterologous practical applications.

Keywords Heterologous multi-view · Stereo positioning · Rational polynomial model (RPC) · Error propagation

Y. Li (✉) · H. Wu · E. Long · Z. Li
Institute of Remote Sensing Information of Beijing, No.2, Qinghe Xiaoying East Road, Haidian, Beijing 100192, China
e-mail: 110836315@qq.com

X. Chang
State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430072, Hubei, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_62

1 Introduction

With the breakthrough of key measurement technologies such as positioning, navigation, attitude and time in the high-resolution earth observation satellite platform, the stereo positioning has become an important development trend in the field of remote sensing. Currently, a large number of high-resolution satellites in orbit can acquire massive repeatedly shot images of the global area in a short time. However, the stereo positioning potential of heterologous remote sensors is far from being fully developed. The joint processing of homologous data is still the main method to produce DEM (see [1–3]). However, these data acquisition conditions are very harsh and costly such as three-line stereo camera and InSAR, which make it difficult to be effectively applied in a wide range.

CCD cameras and SAR loads have become the mainstream remote sensors. Both of them have own advantages and disadvantages: CCD has high signal-to-noise ratio but limited stereo effects; SAR can work all-weather and all-day, and its unique side-view feature make it sensitive to height information. Especially with SAR image resolution significantly increased, SAR can well complement the optical remote sensing to achieve the auxiliary positioning purpose. The combination of SAR and CCD images for stereo positioning can make full use of two different imaging methods to achieve CCD high-plane positioning accuracy and SAR richness information. And compared with two-view stereo, multi-view has better stereo observation structure, which greatly enhances its stereo-imaging possibility. Therefore, in the previous study [4, 5], we have established a set of heterogenous multi-view stereo positioning scheme. By combining multiple CCDs and SARs with completely different imaging mechanisms, and any number of scales with overlapping regions, SAR remote sensing images into a unified three-dimensional processing system, and experiments show that it is expected to reach the traditional homologous CCD stereo positioning results.

At present, the error analysis theory of homologous optical stereo pair is very mature, and there are clear conclusions for the choice of image source, location accuracy and so on to guide the engineering application. However, the side-view imaging mechanism of SAR images is completely different from that of CCD images. As a result, the traditional theories of optical three-dimensional error analysis are completely unsuitable for heterogenous stereo. Therefore, the analysis of all-link error propagation must be conducted to quantitatively describe the relationship between the heterogeneous stereo accuracy and all possible influencing factors in order to provide a reliable theoretical basis for the practical application.

For example, how to make the choice of most contribution to the positioning accuracy in a large amount of data with different shooting angles, precision levels and load types, etc. The answer of these problems must be clear, cannot be blindly. However, there are few researches on stereotactic localization using active and passive remote sensing (see [6–9]), which are limited to a small amount of test data without error theory analysis, and image selection is completely random. The literature on SAR location errors is also mainly focused on the evaluation of plane

positioning accuracy and the equivalence with optical geometric models [10–12]. There is no relevant research about the law of heterogenous positioning errors.

Based on this, this paper starts with deducing the equivalence relation between SAR geometric model and rational polynomial model (RPC), and obtains the correspondence between the real side view angle of SAR and the incident angle of equivalent RPC. Then all-link error propagation modeling is set up to simulate and evaluate the quantitative relationship between the positioning accuracy and the influencing factors such as the image source angle, the tie-points matching accuracy, etc. We analyze the theoretical accuracy and superiority that can be achieved by the heterogenous stereo, and deduces a set of explicit data selection strategies, that is how to select effective data among a large number of images repeated shooting in the same region so as to maximize the final target area positioning accuracy. The correctness of the upper theoretical analysis is fully verified with a number of on-orbit data, and the error law under different data combinations has been detailed analyzed. All of these provide a reliable basis for practical applications of heterogeneous multi-view stereoscopic.

2 Derivation of RPC Equivalent Incident Angle of SAR

In our previous literature [4], we gave a detailed principle of stereo localization based on heterogeneous multi-view images of CCD and SAR. The brief introduction is as follows: ① Compensate the systematic errors of tight geometric model by means of calibration and other means. The high-precision internal and external orientation elements of CCD image are calibrated by use of calibration field. The SAR image delay system error is calibrated by use of external calibration equipment (such as corner reflectors); ② Using the general advantage of RFM model, the CCD and SAR image geometric models with completely different imaging mechanism are unified to be equivalent, which can facilitate the construction of any homologous or heterologous, two-view or multi-view stereoscopic positioning model, reduce the computational complexity and overcome the problem of non-convergence iteration; ③ Using the RFM stereo adjustment method based on the image compensation, through the connection point constraints of images and a certain number of control points, we can establish the image plane affine transformation model to correct the systematic errors to further improve the RFM positioning accuracy. Then a more accurate three-dimensional model can be established only using a small number of control points even without control. ④ Based on the affine transformation parameters and the RPC parameters of each image, we can obtain tie-point coordinates using the point match method and finally determine the space coordinates of the corresponding ground points through front intersection step.

Heterogeneous multi-view stereo is a scheme which unifies multi-CCD and SAR images into a three-dimensional processing. If still based on the respective imaging geometry models of these images, The condition of multi-image input will bring out

a series of problems such as the tie point matching, heterogenous stereo model building and so on. Therefore, it is the premise of the subsequent multi-view processing that all input image source should be unified into an equivalent modeling. RPC model is a very good fit of the geometric model which has a common advantage for multi-source image processing. It has been known that the CCD observation angle is basically coincident with its equivalent RPC light. If the relation between the SAR side view angle and its equivalent RPC light can be deduced, the error conclusion of the traditional optical stereo can be completely applied for heterogenous stereo.

Qin et al. [13] and Zhang Guo et al. [14] have analyzed in detail the accuracy of the RPC model to replace the space-borne SAR tight geometric imaging model with better than one-thousandth pixel, that is, the RPC model is in principle equivalent to the SAR geometric equation. No longer go into details here. The following is the derivation of the RPC equivalent incident angle for SAR. RPC essentially represents a ray of light. The basic step of solving the problem is that, based on the rigorous imaging model the projection ground coordinates on multiple elevation planes of any image point on the regular grid are computed; And fits the lines from the ground coordinates to their image coordinates to be the RPC beam pointing. Specifically, a point P, its height is h, its image coordinate is x_2. The satellite height is H, m_x is the distance resolution, R_0 is the near range slope distance. The geometric equation of point P is $(R_0 + x_2 m_x)^2 = (X_P - X_S)^2 + (H - h)^2$ the upper and lower elevation planes are uniformly layered by the interval Δl in the elevation direction. Let H-h be L, let the SAR load as the arc center and the distance $R_0 + x_2 m_x$ as the arc radius. Then the intersection points P1 and P2 between this arc and the upper and lower elevation surface are the projection points on the elevation surface of the same name of the P-like point. Both satisfy the geometric formula

$$\begin{cases} (R_0 + x_2 m_x)^2 = (X_{P1} - X_S)^2 + (L - \Delta l)^2 \\ (R_0 + x_2 m_x)^2 = (X_{P2} - X_S)^2 + (L + \Delta l)^2 \end{cases} \Rightarrow \begin{cases} X_{P1} = X_S + \sqrt{(R_0 + x_2 m_x)^2 - (L - \Delta l)^2} \\ X_{P2} = X_S + \sqrt{(R_0 + x_2 m_x)^2 - (L + \Delta l)^2} \end{cases} \tag{2.1}$$

Connect P1 and P2, the connection inclination is approximately equivalent to the beam pointing by RPC

$$\tan \gamma = \frac{|X_{P1} - X_{P2}|}{2\Delta l} = \frac{\sqrt{(R_0 + x_2 m_x)^2 - (L - \Delta l)^2} - \sqrt{(R_0 + x_2 m_x)^2 - (L + \Delta l)^2}}{2\Delta l} \tag{2.2}$$

Generally the SAR satellite height is very large relative to the image width, that is, the slope range at the ground point P is $R_0 + x_2 m_x \approx L / \tan \beta$, β is the SAR side view angle. Equation (2.2) becomes

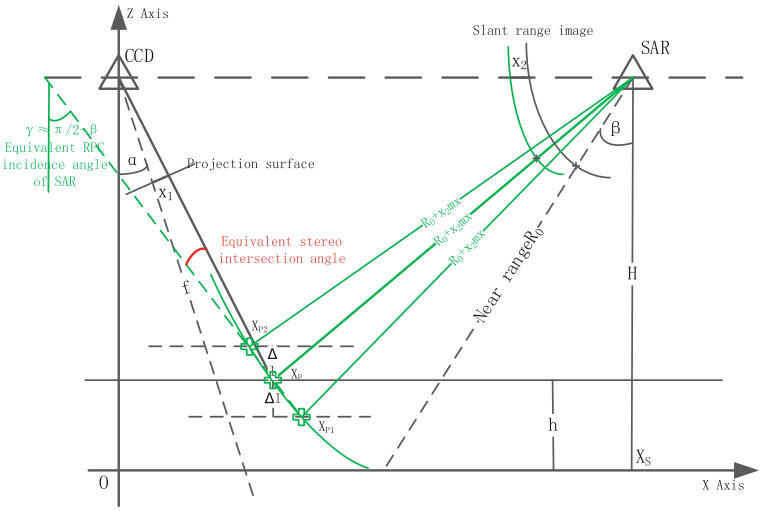


Fig. 1 The relationship between the side view angle of SAR and its equivalent RPC beam angle

$$\tan\gamma = \frac{\sqrt{(L/\tan\beta)^2-(L-\Delta l)^2}-\sqrt{(L/\tan\beta)^2-(L+\Delta l)^2}}{2\Delta l} \tag{2.3}$$

According to Eq. (2.3), draw the relationship between γ and β with Fig. 1, concluding that "the SAR equivalent incident angle of RPC is perpendicular to its side view angle", that is $\gamma \approx \pi/2 - \beta$, the data test result of Sect. 4 will also prove this. This is different from CCD because the equivalent perspective of RPC is same as its roll angle. According to the RPC equivalent incident angle of SAR, the equivalent stereo intersection angle of heterologous images can be calculated which can directly determine the elevation accuracy of stereo positioning. The intersection angle is suitable and the elevation accuracy is higher. Therefore, this conclusion is very important for the image source selection and the location accuracy prediction with SAR participation in stereo positioning.

3 Theoretical Derivation of Heterogeneous Stereo Positioning Error

3.1 Error Characteristic Modeling

The rigorous imaging geometric models of SAR and CCD are non-linear analytic and cannot be directly solved for positions and errors. Without loss of generality, a simplified plane coordinate system XOZ is established as shown in Fig. 2, and the

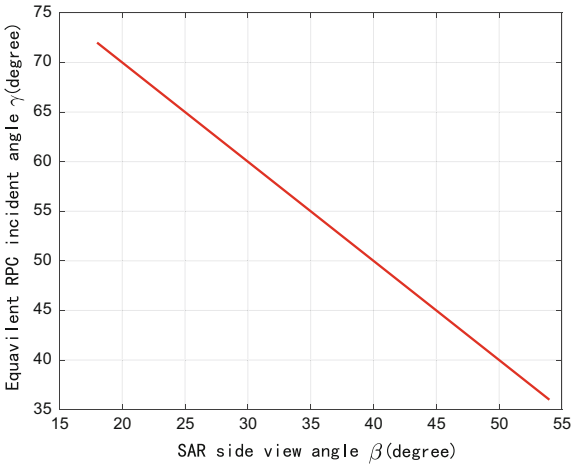


Fig. 2 The RPC solution of SAR and its equivalent three-dimensional intersection with the CCD

Y-direction components are all regarded as 0, and the subsequent analysis is conducted in this plane. Set a point P $(X_P, 0, h)$ on the ground level with elevation h. Its coordinate on the SAR slant image is x_2 and its coordinate on the center projection CCD image with the focal length f is x_1. In order to simplify the calculation, assuming that the height of both the SAR satellite and the CCD satellite are H. The left-hand side of the CCD satellite orbit shows the ground point P at α angle while the SAR satellite observes the β-angle at the side. The CCD position is $(0, 0, H)$ and the SAR position is $(X_S, 0, H)$. Neglecting the influence of satellite's own attitude changes such as pitch and yaw on the imaging geometry, the tight collinear equation of CCD becomes

$$x_1 = -f\frac{\cos\alpha X_P + \sin\alpha(h-H)}{-\sin\alpha X_P + \cos\alpha(h-H)} \Rightarrow X_P = \frac{x_1\cos\alpha + f\sin\alpha}{x_1\sin\alpha - f\cos\alpha}(h-H) \tag{3.1}$$

The SAR Range-Doppler model becomes $(R_P + x_2 m_x)^2 = (H/\cos\beta + x_2 m_x)^2 = (X_P - X_S)^2 + (h-H)^2$. When the SAR and CCD locate at different sides of the ground point, we can derive

$$X_P = X_S - \sqrt{(H/\cos\beta + x_2 m_x)^2 - (h-H)^2} \tag{3.2}$$

Equations (3.1) and (3.2) can be derived as

$$X_S - \sqrt{(H/\cos\beta + x_2 m_x)^2 - (h-H)^2} = \frac{x_1\cos\alpha + f\sin\alpha}{x_1\sin\alpha - f\cos\alpha}(h-H) \tag{3.3}$$

Then

$$\frac{x_1^2+f^2}{(x_1\sin\alpha-f\cos\alpha)^2}(h-H)^2-2X_S\frac{x_1\cos\alpha+f\sin\alpha}{x_1\sin\alpha-f\cos\alpha}(h-H)+X_S^2 - (H/\cos\beta+x_2m_x)^2 = 0 \tag{3.4}$$

The solution obtained is

$$h-H=\frac{X_S(x_1\cos\alpha+f\sin\alpha)(x_1\sin\alpha-f\cos\alpha)\pm(x_1\sin\alpha-f\cos\alpha)\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-(x_1^2+f^2)\left(X_S^2-(H/\cos\beta+x_2m_x)^2\right)}}{x_1^2+f^2} \tag{3.5}$$

We substitute the in-orbit data into the upper equation and conclude that it is significant to take this formula sign plus.

The effective range of SAR view area is generally 10–20 km, while the satellite altitude is generally above 500 km. According to that the view angle of SAR is generally 20°–60°, the distance between its near-range point and its subsatellite point is at least 180 km or more. So the width can be neglected and X_S can be approximated as

$$X_S\approx X_P+H\tan\beta=\frac{x_1\cos\alpha+f\sin\alpha}{f\cos\alpha-x_1\sin\alpha}H+H\tan\beta \tag{3.6}$$

For the satellite objects pre-designated to be involved in stereo, all the parameters f,α,β,m_x,X_S in Eq. (3.5) are known parameters. For any point, the accuracy of h is mainly determined by the matching error of the tie-points x1 and x2. If the coordinates of x1 and x2 are accurate, the calculated h is the true target height. Usually x1 and x2 have a number of pixel errors. According to the law of error propagation, the elevation error is given by

$$m_h^2=(\partial h/\partial x_1)^2m_{x1}^2+(\partial h/\partial x_2)^2m_{x2}^2 \tag{3.7}$$

Compute the partial derivative of x_2 from (3.5), the elevation error component obtained by SAR image matching error is

$$\frac{\partial h}{\partial x_2}=\frac{(x_1\sin\alpha-f\cos\alpha)(H/\cos\beta+x_2m_x)m_x}{\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-(x_1^2+f^2)\left(X_S^2-(H/\cos\beta+x_2m_x)^2\right)}} \tag{3.8}$$

Compute the partial derivative of x_1 and the elevation error component obtained by CCD image matching error is

$$\frac{\partial h}{\partial x_1}=\frac{2x_1\cdot\left(\frac{X_S(x_1\cos\alpha+f\sin\alpha)(x_1\sin\alpha-f\cos\alpha)+(x_1\sin\alpha-f\cos\alpha)}{\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-\left(x_1^2+f^2\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)}}\right)}{\left(x_1^2+f^2\right)^2}$$
$$\left(x_1^2+f^2\right)\cdot\left(\frac{\left(\begin{array}{r}X_S(x_1\sin2\alpha-f\cos2\alpha)+\\ X_S^2(x_1\sin2\alpha-f\cos2\alpha)(x_1\cos\alpha+f\sin\alpha)-\\ \left(2x_1^2\sin\alpha+f^2\sin\alpha-fx_1\cos\alpha\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)\end{array}\right)}{\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-\left(x_1^2+f^2\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)}}\right)- \tag{3.9}$$

Substituting (3.6) into (3.2) to get the target point plane coordinate X_P expression

$$X_P=\frac{X_S(x_1\cos\alpha+f\sin\alpha)^2+(x_1\cos\alpha+f\sin\alpha)}{\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-\left(x_1^2+f^2\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)}}{\left(x_1^2+f^2\right)} \tag{3.10}$$

Similarly we see that the plane error is

$$m_{Xp}^2=\left(\partial X_P/\partial x_1\right)^2m_{x1}^2+\left(\partial X_P/\partial x_2\right)^2m_{x2}^2 \tag{3.11}$$

Compute the partial derivatives of x_1 and x_2, we can obtain

$$\frac{\partial X_P}{\partial x_2}=\frac{(x_1\cos\alpha+f\sin\alpha)(H/\cos\beta+x_2m_x)m_x}{\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-\left(x_1^2+f^2\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)}} \tag{3.12}$$

$$\frac{\partial X_P}{\partial x_1}=\frac{2x_1\left(\frac{X_S(x_1\cos\alpha+f\sin\alpha)^2+(x_1\cos\alpha+f\sin\alpha)}{\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-\left(x_1^2+f^2\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)}}\right)}{\left(x_1^2+f^2\right)^2}$$
$$\left(x_1^2+f^2\right)\cdot\left(+\left(\frac{\left(\begin{array}{r}2X_S(x_1\cos\alpha+f\sin\alpha)\cos\alpha\\ 2\cos\alpha X_S^2(x_1\cos\alpha+f\sin\alpha)^2-\\ \left(x_1^2+f^2+x_1^2\cos\alpha+x_1f\sin\alpha\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)\end{array}\right)}{\sqrt{X_S^2(x_1\cos\alpha+f\sin\alpha)^2-\left(x_1^2+f^2\right)\left(X_S^2-\left(H/\cos\beta+x_2m_x\right)^2\right)}}\right)\right)- \tag{3.13}$$

3.2 *Intersection Angle of the Image Source*

Figure 3a draws the curve of elevation error component with β when α = 5°, 30°, Δx1 = 1 pixel. When SAR and CCD are at the same side of ground point, β is negative, and β is positive at the opposite side. The minimum value appears at β = 0°, which does not exist. Usually, β is valid in the range of 20° to 60°. That is to say, if only the CCD matching error is taken into consideration, SAR image with small incidence angle can contribute more stereo accuracy. In fact, SAR image with small incidence angle has poor range resolution, which affects the matching accuracy and leads to larger Δx2, resulting in larger elevation error. Therefore, there is a balance between the two. Figure 3b plots the elevation error component of Δx2 = 1 pixel as a function of β. When the SAR view angle is perpendicular to the CCD, the elevation error caused by the SAR matching error is the largest. When the SAR view angle coincides with the CCD, the elevation error caused by the SAR matching error is the smallest. For example, when α = 30, the SAR image with 30° incidence angle can contribute the best stereo accuracy.

The above conclusions are derived from the error propagation model of the rigorous geometric model, which is consistent with the conclusion in Sect. 2. When the SAR side view coincides with the CCD light, the equivalent RPC angle of the SAR is perpendicular to the CCD light. Then the equivalent stereo intersection angle at the target point achieves 90°. According to the traditional stereo image theory, the elevation error become minimum. On the contrary, the error become greater when the SAR and CCD is close to vertical.

Figure 3c, d plot the plane error component as a function of β when Δx1 and Δx2 are both 1 pixel respectively. Compared with the elevation error, the same matching error brings less plane error. When the CCD is vertical to the SAR incident angle, the plane error becomes maximum. If the CCD load is fixed in the opposite side, the plane error varies smoothly with different β. Especially smaller the CCD side angle is which means close to the down view, smaller the plane error will be affected by SAR. This will be verified in the following data experiment section.

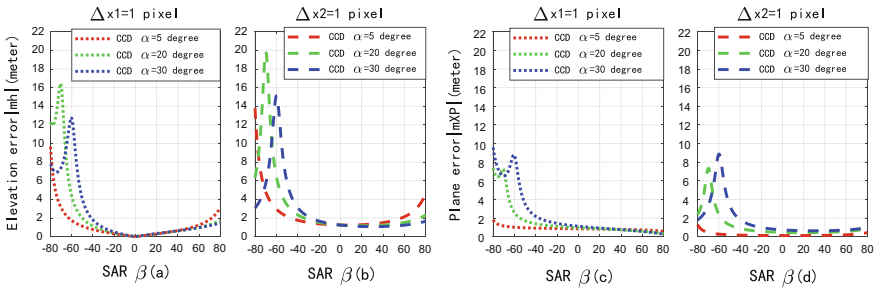


Fig. 3 Elevation error and plane error vary with the image source intersection angle

3.3 The Tie Point Matching Error

The essence of the human eye is the presence of left and right parallax, so the measurement accuracy of the tie point do great impact on the stereo elevation accuracy, which need to be considered. The Fig. 4 shows that the CCD scroll angle α is fixed at 10°, under different SAR side-views β, Δh varies with the CCD match point error $\Delta x1$ and the SAR match point error $\Delta x2$. We can draw a few conclusions:

① The elevation error increases with the increase of the matching error of the tie point; ② The elevation error brought by $\Delta x2$ is larger than that by $\Delta x1$, as shown in the figure when $\beta = -45°$, $\Delta x2 = 2$ when $\Delta h = 5$ m, $\Delta x1 = 2$ when $\Delta h \approx 2$ m, so it is very important to improve the SAR image matching accuracy; ③ For $\Delta x1$, theoretically when $\beta = 0°$ the error brought by the same $\Delta x1$ is smallest, The bigger the $|\beta|$, the larger the error caused by the same $\Delta x1$. In addition, $\Delta h > 0$ when the SAR and the CCD are on the same side, that is, the existence of $\Delta x1$ makes the calculated elevation larger than the true elevation. And the calculated elevation is smaller when the SAR and the CCD are on the different side; ④ For $\Delta x2$, the larger the intersection angle between SAR and CCD is, the smaller the equivalent intersection angle is, and the bigger Δh is caused by the same $\Delta x2$; on the contrary, the smaller the intersection angle is, the smaller Δh caused by $\Delta x2$; in addition, in any case $\Delta h < 0$, that is, the existence of $\Delta x2$ makes the calculated elevation smaller than the true elevation.

4 Heterologous Stereo Experiment with Multi-resolution Complex Combination

We currently have 1: 10000 and 1: 2000 DOM and DEM high precision reference data of Dengfeng area in Henan province. The main landform categories within the region are complete such as mountain, hills, plains, lakes and so on. We collected

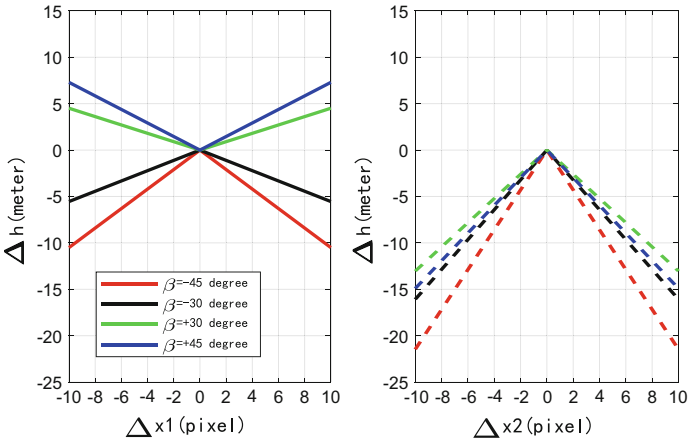


Fig. 4 The elevation error changes with the tie point matching error

multi-view CCD and SAR images within the scope of the reference data for heterogenous stereo positioning. The accuracy evaluation includes the plane errors in the X and Y directions and the elevation error. The basic parameters of the images are as follows. In order to compare the homology and heterologous multi-view test results on the same baseline, the comparative test of the group A and group B use the same points including the connection points, control points, check points.

Homology and heterologous stereo tests were first performed based on image group A. Four connection points were fixed, the number of control points and checkpoints is 15. The test using four control point distribution program, the remaining 11 points are used as checkpoints. We can summarize a few rules from the test results as the Fig. 5:

(1) Since the CCD satellites are all along descend track, the angle of the images captured by the same target repeatedly is very limited. For example, the angle between the two CCDs in the group A is only 4.28°. The stereo condition is poor and the elevation error is very large. The plane error can be well eliminated after 4-point control, but the elevation error does not improve. It proves that adding more control points does not help the elevation error if the stereo condition is not good.

(2) After adding the SAR image, their track and the viewing angles are greatly different. The equivalent stereo intersection angle increases from 4° to 51°, providing good stereo observation conditions. The elevation error is reduced from 29.017 to 3.039 m. It is proved that the SAR image contributes a great deal to the stereo localization. In addition, the plane positioning accuracy of three-view combination is the highest, reaching 1.802 m, to prove that the more the number of overlapping images, the higher the reliability of the adjustment results.

(3) Under the 4-point control scheme, by solving the parameters of the affine transformation model, the translation error due to the position and attitude of the sensor, the rotation and the proportional error caused by the spaceborne

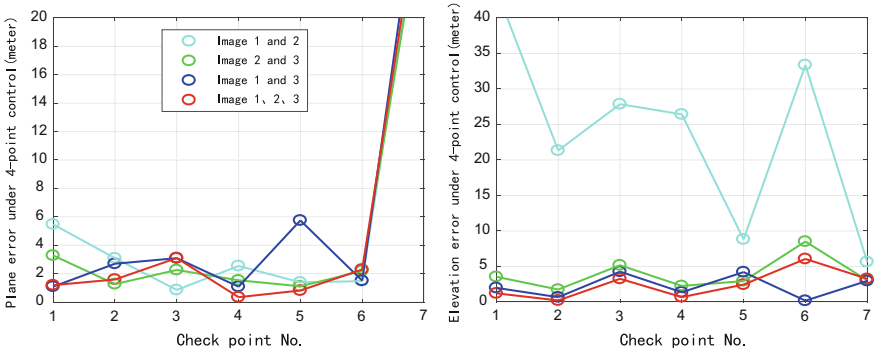


Fig. 5 Plane and elevation error of group A in 4-point control condition

Table 1 Basic parameters of the satellite images participating in the test

Group	Image	Load type	Acquisition time	Ascend or descend track/side view angle
Group A	1	CCD	2012-10	Descend track/left side view 6.75
	2	CCD	2013-03	Descend track/left side view 2.47
	3	SAR	2011-11	Ascend track/right side view 36.41
Group B	4	CCD	2014-12	Descend track/left side view 6.07
	5	CCD	2015-04	Descend track/left side view 28.17
	6	CCD	2015-01	Descend track/right side view 2.74
	7	SAR	2015-01	Ascend track/right side view 37.57
	8	SAR	2015-01	Descend track/right side view 35.97

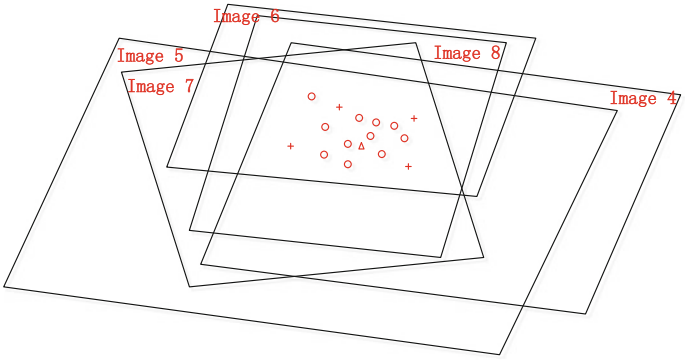


Fig. 6 The relative position of the images in group B

GPS and the inertial navigation system are absorbed. The plane and elevation error can reach 1.802 and 3.039 m, respectively, which can fully meet the 1:50000 topographic map stereo measurement requirements.

The more complex multi-resolution combined stereoscopic accuracy is further verified. Five images including CCD (0.8 m), CCD (1 m), SAR (3 m) are chosen as Group B in Table 1. Figure 6 shows the relative position of these images. Figure 7 shows the relative position of the satellite camera and the target image at the time when the five images are captured in a Cartesian coordinate system. The angle is from north to south. The solid lines of different colors are the real incidence angles of each image, and the dotted lines represent the equivalent perspective angle calculated from their RPC coefficients. It can be seen that the images 4, 5, and 6 are CCD images, and the roll angles are basically the same as the equivalent RPC angles. The images 7 and 8 are SAR images, and the roll angles and the equivalent RPC angles are substantially perpendicular, which is completely consistent with the theoretical derivation in the subsection "The RPC equivalent incident angle of SAR is perpendicular to its lateral viewing angle", and further proves the correctness of this conclusion.

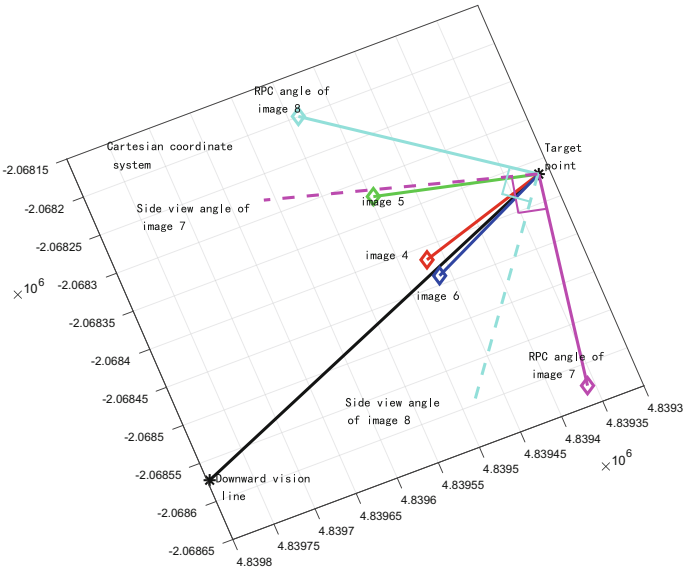


Fig. 7 Equivalent RPC incident light perspectives for each image

Table 2 shows the contrasts of the plane error and the elevation error for several typical combinations of 3-point control deployment. The heterogeneous stereo localization of this paper is that the heterogeneous SAR and CCD are transformed into the RFM model for stereo adjustment. The "equivalent stereo intersection angle" in Table 2 refers to the real intersection angle calculated using its RPC parameters after the relative orientation of all the images involved, in order to more clearly see the law. Summarize the rules as follows:

(1) The heterogenous stereo is the same as the homogenous stereo that the error decreases with the increase of the equivalent stereo intersection angle.
(2) The CCD is the central projection conformation, so its viewing angle is limited and the stereo effect is relatively insignificant. While the SAR side-view imaging mechanism makes that the larger the viewing angle, the higher the resolution. Large viewing angle itself is helpful to the stereo. For example, after adding SAR in group B2 and B8, the equivalent intersection angle increased from 25° to 46° and the elevation error decreased from 3.4 to 2.5 m. It embodies the heterologous advantage of adding SAR images to stereoscopic image sources.
(3) Compare the combination B6-B9, the stereo errors gradually decrease with the increase of the number of stereo images, which is consistent with the previous theoretical analysis and shows the advantages of multi-sources, even though the maximum intersection angle does not increase.
(4) Due to space limitations, this article does not list the single CCD and SAR image geometric accuracy. The experiment found that due to the limitations of platform design and imaging mechanism, the geometric accuracy of

Table 2 The plane error and the elevation error for group B under 3-point control conditions

Group No.	Image composition	Equivalent maximum stereo intersection angle (°)	Error at the check points (m)			
			X-direction	Y-direction	Plane	Elevation
B1	Homogenous and different side 4/6	7.16	1.533	2.179	2.664	18.227
B2	Homogenous and same side 4/5	25.5	1.962	2.436	3.128	3.311
B3	Homogenous and different side 5/6	32.6	1.711	3.166	3.598	3.621
B4	Homogenous 4/5/6	32.6	1.518	2.564	2.979	3.445
B5	Heterogenous and different side 5/8	20.8	5.286	3.020	6.088	5.643
B6	Heterogenous and different side 4/8	46.3	1.866	1.872	2.643	2.933
B7	Heterogenous and same side 6/8	53.4	1.682	2.500	3.013	2.945
B8	Heterogenous 4/5/8	46.3	2.207	2.243	3.146	2.569
B9	Heterogenous 4/5/6/8	53.4	1.595	2.358	2.847	2.416

single-view SAR is lower than the CCD. As shown in Table 2, the plane accuracy after stereo positioning is mainly determined by the CCD, which is consistent with the theoretical analysis in Sect. 3.1.

5 Conclusion

In view of the current lack of in-depth error analysis of the related studies at home and abroad, completely random choice of experimental images, it is difficult to turn the heterogeneous stereo positioning into practical. So based on the previously proposed heterogenous active-passive multi-view stereo positioning scheme, a clear heterogenous stereo positioning error law is derived from two aspects as error theory analysis and in-orbit data validation. Finally, the following conclusions are obtained: Under the premise of certain satellite parameters, the errors of the heterologous stereo positioning are mainly related to the equivalent stereo intersection angle, the matching accuracy of the tie point, the number of participating images, etc., but are different from the traditional optical stereo error theory, ① The equivalent RPC angle of SAR is perpendicular to its true side view angle. Therefore, when the SAR side view angle coincides with the CCD side swing angle, the equivalent stereo intersection angle reaches 90° and the elevation and

plane positioning errors are minimal. On the contrary, when SAR and CCD are vertical, the equivalent intersection angle is 0 and the error is the largest. ② A certain matching error has different influence on the positioning error under different intersection angles. While increasing the stereo intersection angle through image selection, the influence on image resolution and matching error should also be considered; ③ The larger equivalent stereo intersection angle calculated by all participating RPC parameters, the better the stereo condition, the smaller the elevation error. Even if the intersection angle does not increase, the positioning error will reduce as the number of images increases; ④The plane positioning accuracy of the heterologous multi-view stereoscopic is mainly determined by the CCD image, when the CCD perspective is smaller close to the under view, the plane error less affected by the SAR image.

The above conclusion provides a reliable basis for the practical application of heterogeneous multi-view stereo scheme. However, the promotion of heterogenous stereo positioning technology still need to solve the problem of automation. Among them, the automatic registration of multi-source remote sensing images is a key. With the universal advantages of RFM model, we will focus on the research of high-precision matching among the heterogenous multi-view images based on geometric constraints, which is a follow-up of a research focus.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Project No. 61501036) and the Opening Fund of Key Laboratory of Space Utilization, Chinese Academy of Sciences (LSU-2016-06-01).

References

1. Zhang Y, Zhang J (2003) Orientation of remote sensing image pairs from different orbits. Geomatics Inf Sci Wuhan Univ 28(5):521–524
2. Ding H, Yao A (2012) DEM generation and analysis using IKONOS stereo pairs. Sci Surveying Mapp 37(1):179–181
3. Dai J, Song W, Li J (2013) Stereoscopic positioning with multi-source optical high resolution satellite images. Sci Surveying Mapp 38(3):14–16
4. Li Y, Wu H, Sun X (2016) Research about stereo positioning using multi-source remote sensing images. In: China satellite navigation conference (CSNC) 2016 proceedings, vol 2
5. Li Y, Wu H, Chang X (2016) Research on stereo positioning with CCD and SAR multi-view image. Sci Surveying Mapp 41(11):137–141
6. Raggam J, Almer A (1990) Mathematical aspects of multi-sensor stereo mapping. IGARSS, Greenbelt, Maryland
7. Toutin T (2000) Stereo-mapping with SPOT-P and ERS-1 SAR images. Int J Remote Sens 21 (8):1657–1674
8. Jordi I, Alain G (2004) On the possibility of automatic multisensor image registration. IEEE Trans Geosci Remote Sens 42(10):2104–2120
9. Xing Shuai Xu, Qing He Yu (2009) Combined Stereo location among multi-sensor remote sensing images. Geomatics Inf Sci Wuhan Univ 34(5):522–526
10. Zhang G, Fei W (2010) Analysis and test of the substitutability of the RPC model for the rigorous sensor model of spaceborne SAR imagery. ACTA Geodaetica et Cartographics Sinica 39(3):264–270

11. Zhang G, Qin X (2013) Spaceborne SAR and InSAR data processing technology based on RPC model. SinoMaps, Beijing
12. You H, Ding C, Fu K (2007) SAR image localization using rigorus SAR Collinearity equation model. ACTA Geodaetica et Cartographics Sinica 36(2):158–162
13. Qin X, Zhang G, Li L (2006) The algorithm of computing RPC model's parameters for SAR image. J Chengdu Univ Technol (Science & Technology Edition) 33(4): 349–355
14. Zhang G, Fei W, Li Z (2010) Evaluation of the RPC model for spaceborne SAR Imagery. Photogram Eng Remote Sens 76(6):727–733

The PDR System Based on Improved QSF+ Map Matching Algorithm

Wenchao Zhang, Dongyan Wei, Peiwen Gong and Hong Yuan

Abstract The Zero-velocity Update (ZUPT) aided Extended Kalman Filter (EKF) is commonly used in the classical INS-based PDR system, which can effectively suppress the error growth of the inertial based pedestrian navigation systems. However, the system still suffers from the drift of heading error. The magnetic field is very useful to estimate the heading of the system, but the magnetic disturbance has a severely effect on the estimation. The Quasi-static magnetic Field (QSF) method was developed to estimate heading errors using magnetic field in perturbed environments, but the method may bring extra errors to system as the high false alarm probability of detecting the quasi-static field. In this paper, the improved QSF method is proposed to estimate the heading in the perturbed magnetic field. Also, the improved QSF method is combined with a compass filter, which can successfully extract the desired magnetic measurements and feedback them into the EKF to estimate the heading errors. At last, the iterative 2D map matching method is proposed to refine the trajectory of the PDR system, which can effectively suppress the long-term drift errors of the trajectory. The experiment result shows that the trial trajectory closed error length is 0.109%.

Keywords PDR system · ZUPT-aided EKF · QSF · Compass filter
Map matching

W. Zhang · P. Gong
University of Chinese Academy of Sciences, Beijing, China
e-mail: zhangwenchao@aoe.ac.cn

W. Zhang · D. Wei (✉) · P. Gong · H. Yuan
Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China
e-mail: weidongyan@aoe.ac.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_63

1 Introduction

The Global Navigation Satellite System (GNSS) is widely used as a basic method of pedestrian navigation positioning. However, in urban canyons and indoor environments, GNSS positioning systems cannot be used due to signal attenuation and interference. Indoor pedestrian positioning system is an effective means of indoor positioning, which is a good complement to the GNSS positioning system. Many scholars have proposed different methods for indoor pedestrian positioning system. Among these methods, the foot-mounted inertial measurement unit (IMU) positioning system has wide range of applications [1] because it does not require installation of infrastructure and can independently implement pedestrian positioning.

The disadvantage of IMU system is that its error accumulates with time. In pedestrian locating system, the zero velocity updating (ZUPT) algorithm and the extended Kalman filter algorithm (EKF) are often used to limit the IMU accumulation error [2]. The principle is that there will be a short zero-speed interval during the pedestrians gait cycles, in the interval, the position error can be corrected by using the correlation of velocity and position. However, because of the weak observability of velocity and attitude, the heading of the system will still diverge with time, resulting in a large error in the positioning result [3].

Earth's magnetic field is very useful for estimating the heading of the system, but the magnetic field disturbance can have a serious impact on the heading estimation of the pedestrian positioning system. The quasi-static magnetic field (QSF) method [4] can be used to estimate the heading in the disturbed magnetic field. However, this method may bring extra error to the system due to the higher probability of false alarm detection in quasi-static fields.

In this paper, the basic algorithm of pedestrian location mainly refers to the relevant contents in [2] and [3]. The main contributions of this paper are as follows: 1. An improved QSF algorithm combined with a multi-stage compass filter [5] is proposed to estimate the heading information in the disturbed magnetic field. 2. An iterative two-dimensional map matching algorithm is proposed to solve the problem of heading error caused by the low output frequency of reliable magnetic heading in disturbed environment. This map matching algorithm can ensure the reliability of the pedestrian positioning result. 3. A complete PDR positioning system is designed, and the positioning performance of the system has been verified by the experiments.

2 Overview Scheme of the PDR System

The pedestrian location algorithm designed in this paper is shown in Fig. 1. The bottom module is the IMU device (including three-axis accelerometer, three-axis gyroscope, three-axis magnetometer). The middle module contains four units:

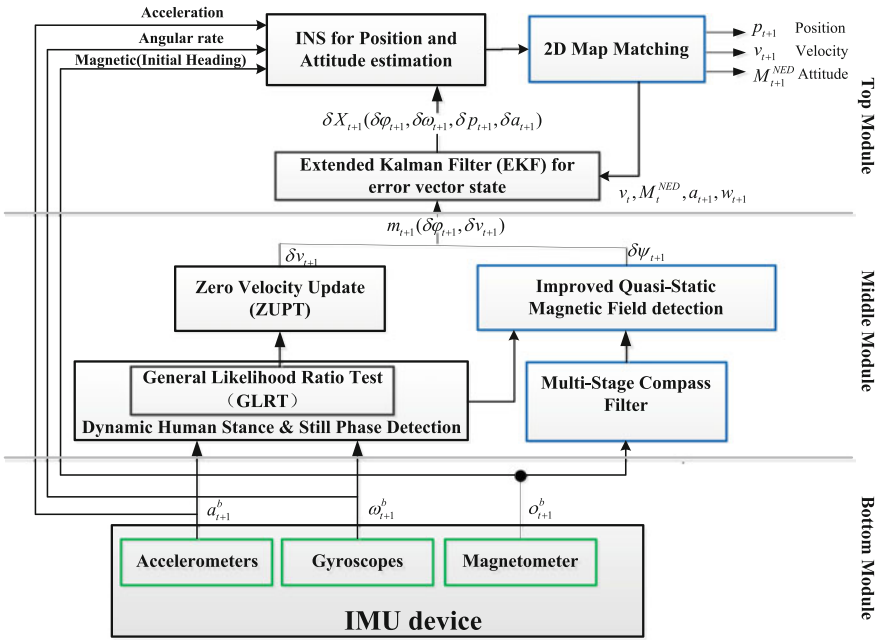


Fig. 1 The PDR scheme of improved QSF+ Map matching algorithm

the Generalized Likelihood Ratio Test (GLRT) unit is used to detect dynamic pedestrian static interval and assists the ZUPT algorithm unit. The multi-stage compass filter unit is used to preprocess magnetic interference data in order to assist the improved QSF algorithm unit to extract reliable magnetic heading; The top module contains three units: The EKF filter unit is used to estimate the pedestrian positioning system state error, the inertia solution unit is used to achieve inertial solution, the 2D map matching unit is used to match the final positioning result with the map in order to revise the final pedestrian loci.

3 The PDR System Analysis

3.1 Dynamic Human Stance and Still Phase Detection

The pedestrian walk cycle consists of two phases: the swing phase and the stance phase. In the swing phase, the foot does not touch the ground. Instead, the foot touches the ground during the stance phase.

The Generalized Likelihood Ratio Test (GLRT) proposed in [6] can effectively detect the stationary period (stance stage) during pedestrian walk. The principle is to use both accelerometer and gyroscope output values, denoted as:

$z_{t+1}^{a} \triangleq \{a_k\}_{k=t+1}^{t+W}$ and $z_{t+1}^{w} \triangleq \{w_k\}_{k=t+1}^{t+W}$ (W refers to the data window of zero speed detection) respectively. If pedestrians are at stance stage, then:

$$T(z_{t+1}^{a}, z_{t+1}^{w}) < \gamma \tag{1}$$

Here, $T\left(z_{t+1}^{a}, z_{t+1}^{w}\right)$ is the detector of the stance stage. γ is the set threshold value; the GLRT detector statistic can be denoted as follows:

$$T(z_{t+1}^{a}, z_{t+1}^{w}) = \frac{1}{W} \sum_{k=t+1}^{t+w} \left(\frac{1}{\sigma_a^2} \left\| a_k - g \frac{\bar{a}_{t+1}}{\|\bar{a}_{t+1}\|} \right\|^2 + \frac{1}{\sigma_w^2} \|w_k\|^2 \right) \tag{2}$$

Here, σ_a^2 and σ_w^2 denote the acceleration and angular rate measurement noise variance, respectively. And there is $\|a\| = a^T a$, in this equation $\overline{a}_{t+1}$ denotes the acceleration measurements mean.

3.2 Multi-stage Magnetic Data Filter

The multi-stage magnetic data filter algorithm is used to preprocess the magnetic data in the disturbed magnetic field. If the magnetic data does not satisfy any stage of the filter, then the magnetic data at that moment is removed. The main three stages are as follows:

Stage 1: Compare the total field strength of the magnetic sensor measurement at the current moment with the total field strength of the magnetic field calculated by using the international geomagnetic reference field model. If the difference between them exceeds the set threshold, the current magnetic sensor output data should be removed.

Stage 2: Calculating the difference between the change in the navigation heading $\Delta\varphi_{-}^{Gyro}$ computed from the gyroscope sensor measurement, and the change in the magnetic heading $\Delta\varphi_{-}^{Compass}$ computed from the magnetic sensor measurement over the predefined period of time $t_{current} - t_{previous}$. Then, if the value of $\Delta\varphi_{current-previous}^{Gyro-Compass}$ exceeds the set threshold, the current magnetic sensor output data should be removed.

$$\begin{aligned} \Delta\varphi_{-}^{Gyro} &= \varphi_{current}^{Gyro} - \varphi_{previous}^{Gyro} \\ \Delta\varphi_{-}^{Compass} &= \varphi_{current}^{Compass} - \varphi_{previous}^{Compass} \\ \Delta\varphi_{current-previous}^{Gyro-Compass} &= \varphi_{-}^{Gyro} - \varphi_{-}^{Compass} \end{aligned} \tag{3}$$

Stage 3: Same as the stage 2, calculating the difference between the change in the navigation heading $\Delta\varphi_{+}^{Gyro}$ and the change in the magnetic heading $\Delta\varphi_{+}^{Compass}$ over the predefined period of time $t_{next} - t_{previous}$. Then, if the value of $\Delta\varphi_{next-current}^{gyro-Compass}$

exceeds the set threshold, the current magnetic sensor output data should be removed.

$$\begin{aligned}\Delta\varphi_{+}^{Gyro} &= \varphi_{next}^{Gyro} - \varphi_{current}^{Gyro} \\ \Delta\varphi_{+}^{Compass} &= \varphi_{next}^{Compass} - \varphi_{current}^{Compass} \\ \Delta\varphi_{next-current}^{gyro-Compass} &= \varphi_{+}^{gyro} - \varphi_{+}^{Compass}\end{aligned} \tag{4}$$

3.3 The Improved QSF Algorithm

After the preprocessing of the multi-stage magnetic data filter in Sect. 3.2, the disturbed magnetic sensor output has been effectively removed. The traditional QSF algorithm detects the quasi-static magnetic field only using the total magnetic field strength. However, this method may bring extra error to the system due to the higher probability of false alarm detection in quasi-static fields [7]. In order to solve this problem, an improved QSF algorithm is proposed which utilizes the following four kinds of information:

(1) Whether the current moment is in the pedestrian stationary period Γ (0 denotes the stationary period);
(2) The total magnetic field strength B at the current moment computed from the magnetic sensor measurement;
(3) The current magnetic heading θ computed from the magnetic sensor measurement;
(4) The navigation heading φ computed from the gyroscope sensor measurement;

Using the above four kinds of information and the GLRT algorithm proposed in [6], the magnetic heading and the navigation heading in the k-th interval can be obtained:

$$\theta_k' = \frac{1}{W}\sum_{l\in\Phi_k}\theta_l \quad \varphi_k' = \frac{1}{W}\sum_{l\in\Phi_k}\varphi_l \tag{5}$$

Here, θ_l and φ_l represents the magnetic heading and the navigation heading at the l-th moment, respectively; W represents the size of the detection window; In this equation, Φ_k can be denoted as $\Phi_k = \{l \in N : k-(W-1)/2 \le l \le k+(W-1)/2\}$.

The measurement model is as follows:

$$y_k = s_k + v_k \tag{6}$$

Here, $s_k = \begin{bmatrix} s_k^\theta & s_k^\varphi & s_k^B \end{bmatrix}^T$, where, $s_k^\theta = \left|\theta_k' - \theta_{k-1}'\right|$, $s_k^\varphi = \left|\varphi_k' - \varphi_{k-1}'\right|$, $s_k^B = |\|B_k\| - \|B_{k-1}\||$; $v_k \begin{bmatrix} v_k^\theta & v_k^\varphi & v_k^B \end{bmatrix}^T$ denotes the measurement noise. Assuming v_k is white noise.

The Quasi-static magnetic field detection can be formalized as a binary hypothesis testing problems as follows:

$$\begin{aligned} &H_0 : \exists k \in \Omega_n \ s.t. \ \Gamma \neq 0 \ s_k^\theta \neq 0 \ or \ s_k^\varphi \neq 0 \ or \ s_k^B \neq 0 \\ &H_1 : \forall k \in \Omega_n \ then \ \Gamma = 0 \ and \ s_k^\theta = 0 \ and \ s_k^\varphi = 0 \ and \ s_k^B = 0 \end{aligned} \tag{7}$$

where, H_0 denotes the non-static field, H_1 denotes the quasi-static field. With the measurement model specified by (6), the PDFs for the hypothesis $H_i (i = 0, 1)$ are given by:

$$\begin{aligned} p(y; s_k, H_i) &= \prod_{k \in \Omega_n} p(y_k; s_k, H_i) \\ &= \prod_{k \in \Omega_n} p(y_k^\theta; s_k^\theta, H_i) p(y_k^\theta; s_k^\varphi, H_i) p(y_k^\theta; s_k^B, H_i) \end{aligned} \tag{8}$$

Under the hypothesis $H_1, s_k = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$, then, the PDF for H_1 is given by:

$$\begin{aligned} p(y; s_k, H_1) = &\frac{1}{\left(2\pi\sigma_\theta^2\right)^{3N/2}} \exp\left(-\frac{1}{2\sigma_\theta^2} \sum_{k \in \Omega_n} \left(y_k^\theta\right)^2 \right) \\ &\cdot \frac{1}{\left(2\pi\sigma_\varphi^2\right)^{3N/2}} \exp\left(-\frac{1}{2\sigma_\varphi^2} \sum_{k \in \Omega_n} \left(y_k^\varphi\right)^2 \right) \\ &\cdot \frac{1}{\left(2\pi\sigma_B^2\right)^{3N/2}} \exp\left(-\frac{1}{2\sigma_B^2} \sum_{k \in \Omega_n} \left(y_k^B\right)^2 \right) \end{aligned} \tag{9}$$

Under the hypothesis H_0, the PDF for H_0 can't be specified due to lack of knowledge about s_k. However, by replacing s_k with its Maximum Likelihood Estimates (MLEs), the GLRT algorithm can be applied [8].

$$\left. \frac{\partial \ln p(y; s_k, H_0)}{\partial s_k} \right|_{s_k = \hat{s}_k} = 0 \tag{10}$$

Then:

$$\hat{s}_k^\theta = \frac{1}{N} \sum_{k \in \Omega_n} y_k^\theta, \ \hat{s}_k^\varphi = \frac{1}{N} \sum_{k \in \Omega_n} y_k^\varphi, \ \hat{s}_k^B = \frac{1}{N} \sum_{k \in \Omega_n} y_k^B \tag{11}$$

The GLRT decides on H_1, if:

$$T_G = \frac{p(y; s_k, H_1)}{p(y; \hat{s}_k, H_0)} > \gamma \tag{12}$$

If combined the (12) with (9), (10) and (11), then, the GLRT becomes:

$$\begin{aligned} T_G = \exp\left(-\frac{1}{2N\sigma_\theta^2}\left(\sum_{k\in\Omega_n} y_k^\theta\right)^2\right) \cdot \exp\left(-\frac{1}{2N\sigma_\varphi^2}\left(\sum_{k\in\Omega_n} y_k^\varphi\right)^2\right) \\ \cdot \exp\left(-\frac{1}{2N\sigma_B^2}\left(\sum_{k\in\Omega_n} y_k^B\right)^2\right) \end{aligned} \tag{13}$$

Taking the natural logarithm on both sides of (13) and simplifying yields:

$$-\left(\frac{1}{2N\sigma_\theta^2}\left(\sum_{k\in\Omega_n} y_k^\theta\right)^2 + \frac{1}{2N\sigma_\varphi^2}\left(\sum_{k\in\Omega_n} y_k^\varphi\right)^2 + \frac{1}{2N\sigma_B^2}\left(\sum_{k\in\Omega_n} y_k^B\right)^2\right) > \ln\gamma \tag{14}$$

Let $\gamma' = -2\ln\gamma$, then the GLRT can be stated as: choose H_1 if

$$\frac{1}{N}\left(\frac{1}{\sigma_\theta^2}\left(\sum_{k\in\Omega_n} y_k^\theta\right)^2 + \frac{1}{\sigma_\varphi^2}\left(\sum_{k\in\Omega_n} y_k^\varphi\right)^2 + \frac{1}{\sigma_B^2}\left(\sum_{k\in\Omega_n} y_k^B\right)^2\right) < \gamma' \tag{15}$$

3.4 The Extended Kalman Filter Algorithm (EKF)

The IMU/EKF+ZUPT+ Improved QSF+ Map Matching algorithm is shown in Fig. 1. The error state vector of EKF algorithm at time $t+1$ is a 15-element vector [9]:

$$\delta X_{t+1}(\delta\varphi_{t+1}, \delta\omega_{t+1}^b, \delta p_{t+1}, \delta v_{t+1}, \delta a_{t+1}^b) \tag{16}$$

where $\delta\varphi_{t+1}$,$\delta\omega_{t+1}^b$, δp_{t+1}, δv_{t+1} and δa_{t+1}^b represent the EKF's estimated errors of attitude, angular rate, position, velocity and acceleration at time $t+1$, respectively. And its function corrects the INS's output values: the velocity, the position, and the attitude.

When incorporating the ZUPT and improved QSF algorithm, the measurement error vector m_{t+1} of the EKF algorithm is:

$$m_{t+1} = \left[\delta\psi_{t+1}, \delta v_{t+1}^b\right] \tag{17}$$

where, $\delta\psi_{t+1}$ and δv_{t+1} denote the bias error of yaw and the bias error velocity at time $t+1$, respectively. And the measurement matrix H is as follows:

$$H_{4\times15} = \begin{bmatrix} [001] & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{18}$$

3.5 The Iterative 2D Map Matching Algorithm

Using the multi-stage magnetic data filter in Sect. 3.2 and the improved QSF algorithm in Sect. 3.3, the reliable magnetic heading of the pedestrian walk in the disturbing magnetic field can be effectively extracted. However, this method may cause the output frequency of the magnetic heading too low to revise the navigation heading error computed from the gyroscope sensor measurement in time.

In order to ensure the stability of the algorithm, an iterative 2D map matching method is proposed to correct the pedestrian localization trajectory of the PDR system, which can effectively restrain long-term drift error of pedestrian loci. Figure 2 is the iterative 2D map matching algorithm flow.

Using the IMU/EKF+ZUPT+Improved QSF algorithm, the current pedestrian's position and heading can be obtained and stored in the algorithm memory for the accessing of the map matching algorithm. During the pedestrian movement, the positioning result within the last 20 m (which can be adjusted according to the actual situation) at the current moment is divided into different groups according to the heading value of each step. The schematic diagram is shown in Fig. 3.

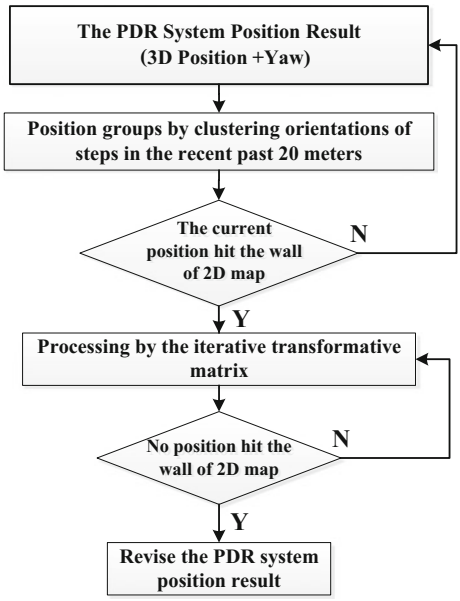


Fig. 2 Flow of map matching

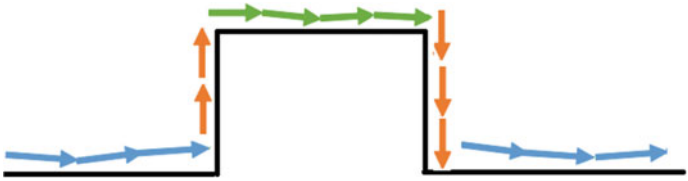


Fig. 3 Position groups by clustering orientations of steps

If the current pedestrian location results exceed the 2D map boundary, then the different groups of positioning results in the last 20 m are processed by different iterative rotation matrices, respectively, which include rotation and telescopic transformation, until all the positioning results in the last 20 m match the map very well.

4 The Experimental Analysis

We use the MTi-100 IMU device from Xsens. The IMU includes three orthogonally-oriented accelerometers, three gyroscopes, three magnetometers, and a barometer. In the experiment, the IMU device was fixed on the foot to collect pedestrian movement data.

4.1 The Experimental Route and Performance Evaluation Methods

The experimental route is the ring road of the park in Academy of Opto-electronics, Chinese Academy of Sciences, which is about 500 m. As shown in Fig. 4, there are 30 verification points (coordinates obtained by the DGNSS) are laid ahead in the test route for testing and verification. Also, there is a 2D map of the test route for map matching, which has been pre-existing in the algorithm memory.

Performance evaluation methods: During the test, pedestrian walks along the test route, when passing through the verification points, the test coordinates of the verification point are recorded by interacting with the computer. After the experiment, the coordinates error of each verification point is compared.

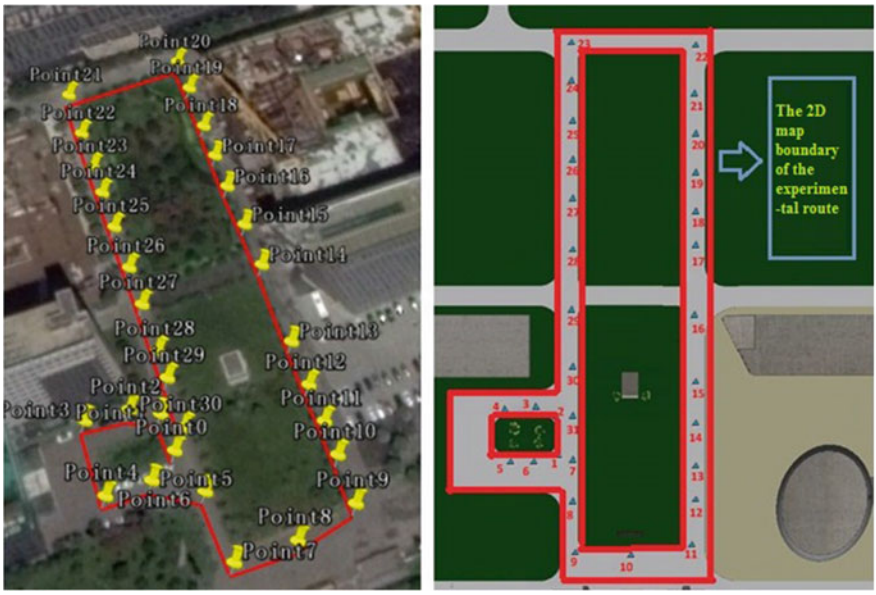


Fig. 4 The experiment route line and its 2D map boundary

4.2 Experimental Results

Using the Sect. 4.1 experimental route and methods, the experimental results are obtained as follows:

As shown in Fig. 5, the test result of the ZUPT+QSF algorithm is better than the test result of only using ZUPT algorithm, although its test trajectory still has obvious offset.

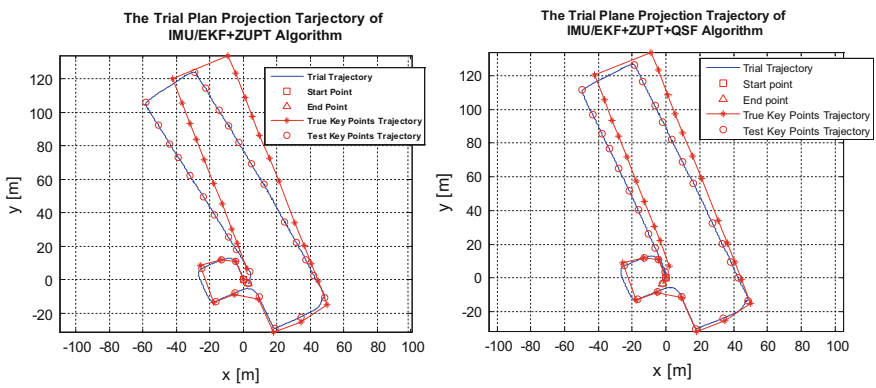


Fig. 5 The test result of IMU/EKF+ZUPT algorithm and IMU/EKF+ZUPT+QSF algorithm

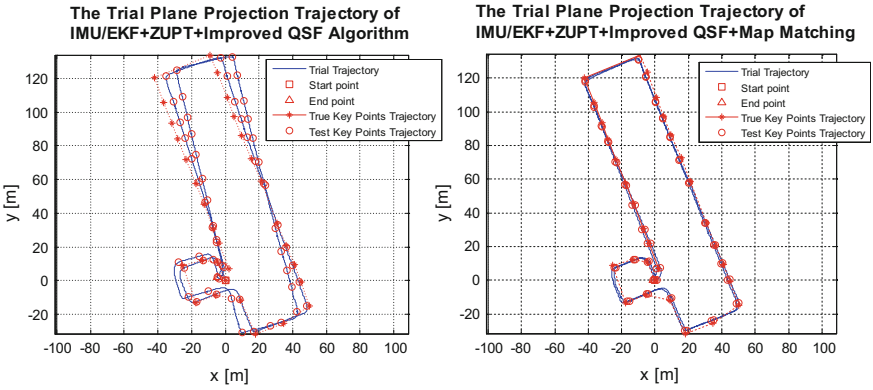


Fig. 6 The test result of IMU/EKF+ZUPT+Improved QSF algorithm and IMU/EKF+ZUPT+Improved QSF+ Map matching algorithm

As shown in the left picture of Fig. 6, the ZUPT+Improved QSF algorithm is used, the pedestrian walks along the test route for two times (total length is 1.0192 km). The test result is much better than the ZUPT+QSF algorithm, while its test trajectory still has slightly offset. The right picture of Fig. 6 shows the test result of ZUPT+Improved QSF+ Map Matching algorithm proposed in this paper. The total test length is 1.0192 km and the test trajectory coincides with the true trajectory very well. The left picture of Fig. 7 shows the position errors of the verification points, while the right picture is the CDF curve of the verification points' errors. The experimental results of the algorithm are shown in Table 1.

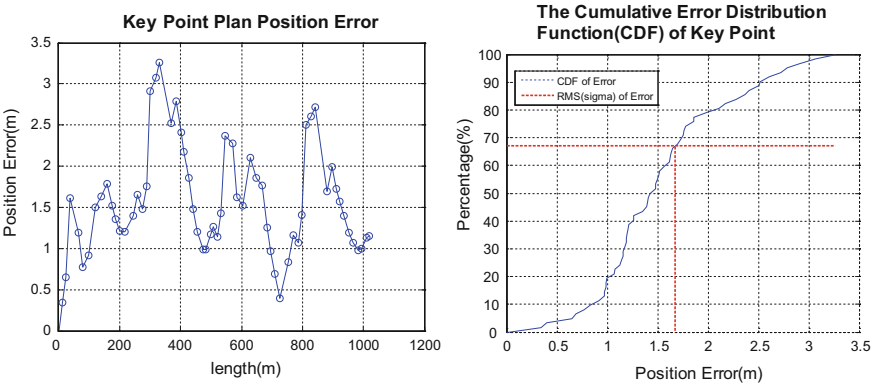


Fig. 7 The key points test error and CDF result of IMU/EKF+ZUPT+Improved QSF+ Map matching algorithm

Table 1 The assessment of ZUPT+Improved QSF+ Map matching

Algorithm type	Closed error (m)	Closed error/length (%)	RMS (m)
The proposed algorithm	1.1144	0.109	1.6495

5 Conclusion

In this paper, the PDR system based on improved QSF+ Map matching algorithm is proposed. An improved QSF algorithm combined a multi-stage compass filter is proposed to estimate the reliable heading of pedestrian in the disturbing magnetic field. Then, a two-dimensional map matching method is proposed to correct the trajectory and ensure the reliability of the PDR system, which can effectively restrain long-term drift error of pedestrian loci. Finally, the experiment shows that the outperformance of the algorithm proposed in this paper.

Acknowledgements The authors greatly thank reviewers' suggestions and comments to improve the manuscript. This work was supported by the National Key Research and Development Program of China (No. 2017YFC0803400).

References

1. Godha S, Lachapelle G, Cannon ME (2006) Integrated GPS/INS system for pedestrian navigation in a signal degraded environment. In: 19th international technical meeting of the satellite division, ION GNSS 2006. 26–29th September, 2006
2. Foxlin E (2005) Pedestrian tracking with shoe-mounted inertial sensors. IEEE Comput Graphics Appl 25(6):38–46
3. Nilsson JO, Skog I, Händel P (2010) Performance characterisation of foot-mounted ZUPT-aided INSs and other related systems. In: International conference on indoor positioning and indoor navigation (IPIN), IEEE, 2010, pp 1–7
4. Afzal MH, Renaudin V, Lachapelle G et al (2011) Use of earth's magnetic field for mitigating gyroscope errors regardless of magnetic perturbation. Sensors 12(6):11390–11414
5. Faulkner WT, Alwood R, Taylor DWA et al (2010) GPS-denied Pedestrian tracking in indoor environments using an IMU and magnetic compass. Proc Int Tech Meet Inst Navig Itm 1 (2):198–204
6. Skog I, Handel P (2010) Zero-velocity detection—an algorithm evaluation. IEEE Trans Biomed Eng 57(11):2657–2666
7. Ma M, Song Q, Li YH, et al (2016) A heading error estimation approach based on improved Quasi-static magnetic field detection. In: International conference on indoor positioning and indoor navigation (IPIN), IEEE, 2016, pp 1–8
8. Kay SM (1993) Fundamentals of statistical signal processing. Prentice Hall, PTR, New Jersey
9. Zhang W, Li X, Wei D et al (2017) A foot-mounted PDR system based on IMU/EKF+HMM +ZUPT+ZARU+HDR+ compass algorithm. In: International conference on indoor positioning and indoor navigation (IPIN), IEEE, 2017 pp 1–5

Part IV
PNT New Concept, New Methods and New Technology

Preliminary Study on Index Parameters of National Integrated PNT System

Qingjun Liu and Xiangjun Wu

Abstract The complexity of the national comprehensive PNT system architecture includes huge differences in system systems and technical means and diverse application scenarios. It is necessary to design a set of complete index parameters to describe it. The construction of the index system of the national integrated PNT system should be subject to the characteristics of ensuring that the system can provide positioning navigation and timing services to all kinds of users in a stable manner. This paper attempts to conduct a preliminary analysis of the national parameters of the integrated PNT system parameters, from the service index parameters and operating parameters of two indicators were combed. The parameters of the service index are from the perspective of the PNT system providing services, including nine indicators of service range, signal characteristics, accuracy, availability, integrity, reliability, user capacity, solution frequency and solution dimension. Operating parameters from the perspective of the system architecture and operation, including consistency, interoperability, adaptability, robustness, upgrades and scalability, maintainability of the six indicators. In this paper, the definition of each indicator is described. As a reference to the whole PNT system parameters, it is of certain reference significance to the framework of the future national integrated PNT system.

Keywords PNT system · Index · Parameters

1 Introduction

Time and space information is an important foundation for maintaining the normal operation of modern society. All technologies for obtaining space-time information are collectively referred to as Positioning Navigation Timing (PNT) technology.

Q. Liu (✉) · X. Wu
Beijing Institute of Spacecraft System Engineering, Beijing, China
e-mail: 7874348@qq.com

X. Wu
e-mail: wuxiangjun@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_64

Currently and in the next few decades, there is not yet a certain PNT technology and its corresponding PNT system can fully meet the application needs of all types of users. Here, the PNT system refers to the physical implementation of a specific PNT technology for the acquisition and transmission of space-time information, in the form of equipment, facilities, basic data, technical standards, etc. With the PNT technology evolution, the emergence of new PNT technologies and user requirements The growth of PNT systems will continue to evolve and be updated. Through the inheritance, integration and development of the existing PNT systems, it is a practical way to meet the needs of future PNT services by systematic means to meet the needs of users. On this basis, the concept of PNT system is put forward: National Integrated PNT (positioning, navigation, Timing system) is a national information infrastructure that serves national defense, economy and society under the unified organization and coordination at the national level, undertakes the tasks of establishing and maintaining national time and space standards, broadcasting and obtaining space and time information, and positioning and navigating time services and applications.

From the perspective of PNT capability development, PNT capability exists before the concept of PNT system. With the development of PNT system, it has been continuously strengthened. PNT system is the inheritance of development history of PNT capabilities, the integration of development reality and the development of the future optimization. At the same time, the composition of the national PNT system is not solidified and needs to be continuously updated and optimized based on reasonable user needs classification, a unified time and space benchmark, a complete technological innovation system and an efficient organizational management system. From the latest developments in the United States and other countries, the means of navigation are being developed from a single satellite navigation operator to a model based on satellite navigation, with multiple means of complementary integration and systematic development. The service coverage has expanded significantly from satellite navigation Mainly covering the surface and near space to the underwater, underground, electromagnetic interference environment, the physical development of the environment; navigation convenience, continuity, reliability, accuracy and other indicators significantly improved; navigation service cost-effective ratio increased significantly.

2 Index System Characteristics

The index system refers to a set of index parameters that are objectively related according to the requirements of different research purposes and the characteristics of the research object, and some indicators of the nature of the research object are scientifically classified and combined. For the national integrated PNT system, the system architecture is complicated, the systems and technical means contained in the system vary greatly, and the application scenarios and application modes are diverse. It is very necessary and difficult to design a set of complete technical

parameter to describe it. However, no matter how complex the system is, as a unified whole, its ultimate goal is to provide various users with positioning, navigation and timing services. All indicators that describe the PNT system should be subject to the characteristics that ensure that the system is able to provide a stable service. In general, the design of the index system for the national integrated PNT system needs to meet the following requirements.

(1) Purpose. The design objective of index system is the premise and foundation of existence of index system. The index system without design purpose is not necessary. The goal of PNT system is to realize space-based positioning and navigation service for all kinds of users. Its index system should be designed and serviced for this purpose.
(2) Theoretical. The design of the index system of the PNT system must be based on a certain theory and guiding ideology. The theory and guiding ideology for establishing the index system determine the selection of the index and the organic composition of the index system. For the construction of PNT system indicators, it not only reflects the architecture of the system and the characteristics of stable operation, but also with the various constituent PNT service system (such as satellite navigation system) are different.
(3) Scientific. The index system of PNT system is used to describe the construction and operation characteristics of the whole system, and science is the prerequisite and basis for the existence of index system. Relying on the different PNT service systems and technical means that make up the PNT system, when designing the index system, we must proceed from the actual situation. The designed indexes should be able to truly reflect the objective reality. At the same time, the designed indexes should conform to the scientific theory.
(4) Systematic. The index system of the PNT system is an organic whole consisting of different indicators describing its characteristics. The organic whole should have a distinct level and a close internal connection, and no indicator should be free from this system.

3 National PNT System Index Parameters

For the national integrated PNT system, the index system can be divided into two categories: system service index parameter package and system operation index parameter package. Among them, the system service index parameter package is sorted out from each system providing the PNT service directly and covers all service systems and means included in the PNT system. The system runtime indicator parameter package is described from the top level architecture of the PNT system and the overall operational characteristics (Fig. 1).

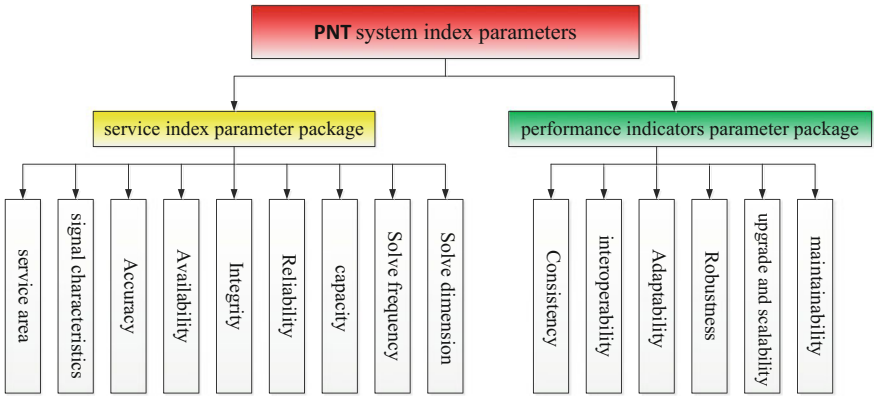


Fig. 1.1 National integrated PNT system index parameters

3.1 PNT System Service Index Parameter Package

Combined with the characteristics and application characteristics of each system providing PNT service, the system defines the following system parameters to make use of and limit the space signal of each independent PNT system according to the development status and future requirements of China's positioning and navigation timing system and technology Description. These parameters include a total of 9 services, signal characteristics, accuracy, availability, integrity, reliability, user capacity, solution frequency, and calculation dimensions. The specific definitions are as follows.

(1) Services area

The service area is the surface area or volume of space covered by the signal that the PNT system is capable of providing the specified navigation or positioning service accuracy. Factors that affect coverage include constellation geometry, signal power levels, atmospheric noise conditions, and other factors that affect signal usability.

(2) Signal characteristics

Spatial signal characteristics are mainly manifested in the power level, frequency, signal format, data rate, and other parameters or ways that can fully represent the user to obtain the PNT information.

(3) Accuracy

In navigation, the accuracy of the position obtained by the user receiver (hand-held, vehicle, aircraft, or device) at an estimated time or measured at a given time is consistent with the true position of the receiver at that moment.

Accuracy is a statistical property and it does not make sense for a particular PNT system to state only accuracy unless it provides both the uncertainty of location services. A detailed description of PNT system accuracy generally refers to one or more of the following definitions:

- Predictable accuracy: The position resolution of a PNT system can be plotted on the map, and the position resolution method and map must be based on the same ground reference.
- Repeatability: With the same PNT system, the same position accuracy can be achieved with one user using previously measured data.
- Relative Accuracy: The accuracy of a location that a user can measure relative to other users' gains in the same PNT system and at the same time.

(4) Availability

The availability of a PNT system is the percentage of time the system is available. Availability is an indication of the capabilities available in the system's coverage area. It is a common function of the physical environment characteristics and technical capabilities of service provisioning equipment.

(5) Reliability

The reliability of a PNT system is a function of the frequency of faults in the system. It describes the possibility that the system will be able to provide standard performance service capabilities for a given period of time under given operating conditions. Formally, reliability is the reverse description of the possibility of a system failure.

(6) Solve frequency

The solution frequency is defined as the number of times the system provides independent position resolution information or data points within a unit of time.

(7) Solve dimension

This feature defines whether the PNT system provides a linear, one-dimensional linear position, or two-dimensional or three-dimensional position solution information. The ability of the system to deliver the fourth dimension of information, such as time, is also included.

(8) System capacity

System capacity is the number of users a system can accommodate at the same time.

(9) Integrity

Integrity is a signal of space set by the standard as a service, providing the level of trust in the correctness of the information. Integrity includes the ability of the system to provide timely alerts to users when the system can not provide navigation services.

3.2 PNT System Performance Indicators Parameter Package

Judging from the whole process of national PNT system research, the evaluation method of national PNT system is significantly different from the traditional positioning navigation time-based system evaluation method. Because the service of national PNT system is flexibly combined according to user needs, PNT resource, application environment and its change, its PNT scheme is flexibly switched and recombined according to user needs, PNT resources, application environments and changes. Therefore, the quantitative indexes such as positioning accuracy, timing accuracy and speed accuracy are not suitable for evaluating the whole PNT system. To this end, we construct the parameter index package to describe the whole PNT system, including the six indicators of consistency, interoperability, adaptability, robustness, scalability and maintainability. Each indicator is graded to describe its degree of achievement of a certain characteristic.

(1) Consistency

Consistency is defined as the ability to submit information/data to all types of users in a consistent manner. Consistency is divided into five levels, namely:

A: Completely different benchmarks, standards, etc.;
Level 2: Limited access, multiple choice, and inconsistent "usage rules";
Level 3: Appropriate access, public benchmarks, different environmental models, etc.;
Level 4: Extensive use of coherence mechanisms;
Level 5: Public Reference Model, Environmental Model, Algorithm, Presentation Model, Protocol, Standard (Timing, Data, Grid).

(2) Interoperability

Interoperability refers to the ability of a system, unit or agency to provide (or receive) PNT services to (or from) other systems, units or agencies to meet mission requirements. Interoperability is divided into five levels, namely:

Level 1: Isolated operation, i.e. different sensors/sensors or electronic systems with different signals/resources;
Level 2: Specific Operations, Short-term Joint, Regional Cooperation
Three levels: coordinated operation, multi-functional electronic system, similar signal protocol;
Level 4: conventional joint operations, the recommended standards and habits;
Level 5: Seamless operation, common electronic system or shared signal.

(3) Adaptability

Adaptability Definition In response to changes in external conditions, adjust the elements of the system, without changing the system infrastructure capabilities. Changes may include tasks, emergencies, user needs and abilities, policies, hostilities, technology, threats and changes in the global environment. Adaptability is divided into five levels, namely:

Level 1: Significant changes to the system and/or take a long time;
Level 2: Need to change multiple components of the system (highly coupled, interdependent, fully tested);
Three levels: in the maintenance/development stage has a certain ability to adjust;
Level 4: Modular components, networked
Level 5: Rapid changes, with limited impact on the underlying system.

(4) Robustness

Robustness is defined as the ability of the PNT system to provide a continuous PNT solution under any condition (hostile behavior, environment, system internal failure) over a given period of time. Robustness can be understood as the ability to reconstruct the system and provide the ability to meet the needs and performance of the service under any conditions. Robustness is divided into five levels, namely:

Level 1: No backup if lost major ability
Level 2: Produces intermittent, unpredictable performance changes;
Level 3: rapid but predictable degradation of performance;
Level 4: Slow, long-term degeneration of function;
Level 5: PNT solution and backup capabilities at any time, any place, any condition, and the performance or ability of the system to be less susceptible to hostile actions and/or environmental conditions.

(5) Upgrade and scalability

Upgrading and scalability are defined as the impact on system services when systems or support systems that make up a PNT system need to be upgraded or expanded. When this effect is too high, it will affect the user's performance. Upgrade and scalability is divided into five levels, namely:

Level 1: Complicated—Large changes in the architecture of the PNT architecture and lead to prolonged interruptions or severe performance degradation of some PNT services;
Level 2: Complicated—Does not change the PNT architecture, but affects system services for a longer period of time or longer;
Level 3: Medium—requires some level of planning and attention, affecting users for a certain amount of time;
Level 4: Simpler—requires less planning and has limited impact on some users;
Level 5: Simple—ready access, smooth transition, no impact on existing users.

(6) Maintainability

Maintainability is defined as the ability to continuously maintain the required level of service. Maintainability Understand the input of manpower, equipment and consumables necessary to maintain service levels and service continuity. When this kind of investment is too high, it is not conducive to long-term and stable operation of the system. Maintainability is divided into five levels, namely:

Level 1: Heavy—Requires service providers to plan long-term, ongoing surveillance, large-scale investments, and full-scale procurement and implementation efforts;
Level 2: Difficulty—requires planning and attention, effective investment, serious implementation;
Level 3: Medium—requires some level of planning and attention, with providers and users sharing the investment and execution;
Level 4: simpler—requires little planning or continuous input, and enforces restrictions;
Level 5: Simple—Self-sustaining, self-motivated, with no implementation issues.

4 Summary

The National Integrated PNT system is a "system of systems" consisting of various systems that provide services, relevant government departments, and policies and regulations, and serves as a whole for future services. Different systems, sectors and agencies are undergoing different changes in their roles in the development vision, strategy and elements of the PNT system as a whole, and this change continues. In carrying out the overall design of the system framework and the key technology breakthrough comprehensive demonstration, the construction and demonstration of the index system is an indispensable part and should be carried out earlier. This paper attempts to explore the index system of the national integrated PNT system, combing from the parameters of service operation parameters and operating index parameters, which can serve as a reference for the national comprehensive PNT system demonstration.

References

1. Yang Y Time base construction and maintenance, "Hui world" high-end forum
2. Xie J (2016) National PNT system based on compass navigation satellite system. In: 2016 Annual academic conference on space electronics
3. Zhao D (2003) Change II navigation system time synchronization and timing research. Chinese Academy of Sciences master's degree thesis

4. Zhao X (2011) Long wave timing in the complex electromagnetic environment in the status and role. In: 2011 National Conference on time and frequency
5. Chen H, Bei Dou system changes in time and space standards and its response, navigation world
6. GJB6304-2008, 2000 China Geodetic Survey System
7. Jiang C, Micro-PNT development overview of the United States, the sixth annual meeting of China's satellite navigation academic
8. Zheng X (2014) Navigation, positioning and timing technology overview, navigation and positioning, and timing, vol 1, No 1, 2014
9. Liu C (2010) How to segment the user market—the latest definition of user domains, departments and applications in the US PNT system, Space Exploration, 2012 (10)
10. Federal Radio navigation Plan 2014
11. Zhang F, Ou M (2015) Overview of potential future PNT technology, Global Positioning System, December 2015, vol 40, No. 6
12. Jiang Z1, Xiong W American PNT policy analysis, Global Positioning Systems, vol 40, No 4

Optimization of Power Allocation Based on Particle Swarm Optimization in Wireless Location Network

Jinrui Lin, Guangxia Li, Shiwei Tian and Longlong Suo

Abstract Wireless location network combines data communication, contextual data collection and navigation. The purpose of this network is to determine the positions of agents based on measurements between nodes and nodes. In order to improve the positioning accuracy of wireless location network, the usual method is to increase the density of nodes, especially the density of anchor nodes, so as to optimize the topology of the network. However, the wireless location network node power is limited, especially the wireless sensor network. It is necessary to increase power consumption to meet the needs of a large number of nodes. However, it is required that a certain level of power consumption should be set in the wireless location network, and even require further reduction in power consumption in daily life. At the same time, we need to maintain a certain positioning accuracy, which requires the optimization of power allocation in the network, to make sure that power and positioning accuracy of the network can meet the needs of the actual living. In this paper, an optimization of power allocation based on particle swarm optimization in wireless location network is proposed to optimize the power allocation. The square position error bound is introduced as the evaluation standard of the network positioning accuracy. Through the power optimization, the positioning accuracy of the network is improved.

Keywords Wireless location network · Power optimization allocation
Square position error bound · Particle swarm optimization algorithm

1 Introduction

Wireless location network uses the characteristics of radio waves to determine the specific location of radio equipment in the network. People have proposed various wireless positioning methods, including the arrival time of signals, the arrival angle

J. Lin · G. Li (✉) · S. Tian · L. Suo
College of Communications Engineering, Army Engineering
University of PLA, Nanjing, China
e-mail: 13905177686@139.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_65

of signals, the intensity of signals and the phase of signals. In [1], the current several indoor positioning technologies are introduced, such as visual positioning, infrared positioning, pole positioning, ultrasonic positioning, WLAN positioning, RFID positioning, inertial navigation, geomagnetic localization, Bluetooth and ZigBee positioning and cellular network positioning.

Currently, Ultra-wideband (UWB) technology is being applied to wireless location networks. [2] points out that ultra-wideband technology can transmit data at the nanosecond level and below by transmitting ultra-narrow pulses, which can achieve GHz-level data bandwidth with low transmission power and no carrier. Because of its high bandwidth, theoretically based on Time of Arrived (TOA) or Time Difference Of Arrived (TDOA) method can achieve centimeter-level positioning. In [3] and [4], the development of ultra-wideband positioning technology, implementation principles and research prospects are reviewed. In [5], the time of arrival (TOA) algorithm is introduced in detail, and the indoor positioning performance of UWB is explored.

In [6], the Cramér–Rao bound in the satellite positioning scenario is introduced as the evaluation criterion of the positioning accuracy. In [7], the square error of position is used in wireless location network. In addition, [8] points out that for wireless wideband location networks, such as wireless sensor networks, resource allocation and optimization play a very important role in positioning accuracy because of the system resources are limited. In the current research of wireless location network, there are few researches about the optimal allocation when resources are limited.

This paper uses the square position error bound in [9] as the criterion for the positioning accuracy of wireless location networks, and focuses on the research of networks based on UWB technology. Base on the analysis of the impact of power allocation in wireless location network on positioning accuracy, the power allocation algorithm using particle swarm optimization theory is proposed.

2 Wireless Location Network

As shown in Fig. 1, wireless location network contains the anchors and agents. According to [7], agents cannot determine their position, and anchors obtain their positions through the GPS positioning system or other means. The agent can range with the anchor, and can also range with other agents. In fact, an agent needs at least three anchors to achieve positioning.

The distance between agent k and j is given by

$$d_{kj} = \left\| \mathbf{p_k} - \mathbf{p_j} \right\| \tag{1}$$

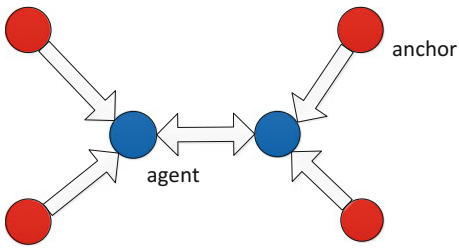


Fig. 1 Wireless location network

The angle from agent k and j is given by

$$\varphi_{kj} = \arctan\left(\frac{y_k - y_j}{x_k - x_j}\right) \tag{2}$$

According to [10], the square position error bound of agent k is given by

$$E\left\{\|\hat{\mathbf{p}}_\mathbf{k} - \mathbf{p}_\mathbf{k}\|^2\right\} \geq p(\mathbf{p}_\mathbf{k}) = tr\left\{\mathbf{J}_e^{-1}(\mathbf{p}_\mathbf{k})\right\} \tag{3}$$

It has been shown in [10] that the network EFIM of N_a agents in a cooperative localization network can be written as

$$\mathbf{J}_e^G = \begin{bmatrix} \mathbf{J}_{1,1} & \cdots & \mathbf{J}_{1,N_a} \\ \vdots & \mathbf{J}_{i,j} & \vdots \\ \mathbf{J}_{N_a,1} & \cdots & \mathbf{J}_{N_a,N_a} \end{bmatrix} \tag{4}$$

where

$$i = j, \mathbf{J}_{ij} = \mathbf{J}_e^A(\mathbf{p}_\mathbf{i}) + \sum_{k \neq i} \mathbf{C}_{ik} \tag{5}$$

$$i \neq j, \mathbf{J}_{ij} = -\mathbf{C}_{ij} \tag{6}$$

$\mathbf{J}_e^A(\mathbf{p}_k)$ and $\mathbf{C}_{kj}$ are the ranging information of agent k obtained from all N_b anchors and agent j, expressed as

$$\mathbf{J}_e^A(\mathbf{p}_\mathbf{k}) = \sum_{j \in N_b} \lambda_{kj} \mathbf{q}_{kj} \mathbf{q}_{kj}^T \tag{7}$$

$$\mathbf{C}_{kj} = \mathbf{C}_{jk} = \left(\lambda_{kj} + \lambda_{jk}\right) \mathbf{q}_{kj} \mathbf{q}_{kj}^T \tag{8}$$

where $\mathbf{q}_{kj} = \left[\cos \varphi_{kj}, \sin \varphi_{kj}\right]^T$

Moreover, the range information intensity can be given by

$$\lambda_{kj} = \xi_{kj} \frac{p_k \beta_k^2}{d_{kj}^{\delta}} \tag{9}$$

where ξ_{kj} is the range channel gain between agent k and j. p_k is the normalized power and β_k is the effective bandwidth allocated to agent k.

According to [11], the total SPEB of all agents can be obtained as

$$\sum_{1}^{N_a} p(\mathbf{p_k}) = tr\left\{ \left(\mathbf{J}_e^G \right)^{-1} \right\} \tag{10}$$

With the square position error bound (SPEB), the positioning accuracy of the network can be evaluated. Research on the power resources in wireless location network mainly refers to the function of the square position error bound and the transmit power. By constraining the power allocation of agents in the network, it is further explored how to optimize the power allocation and increase the network positioning accuracy.

In fact, the optimal power allocation problem in this paper can be modeled as follows

$$\begin{aligned} &\min \sum_{k \in N_a} \sum_{j \in N_a \cup N_b} p\left(\mathbf{p_k}, \{p_j\}\right) \\ &s.t. \quad p_j \le p_{peak} \quad \forall j = 1, 2, \ldots N_a \cup N_b \\ &\qquad \sum_{j \in N_a \cup N_b} p_j \le p_{total} \end{aligned} \tag{11}$$

where p_j is the normalized power of each agent in the network.

3 Algorithm Design

Particle swarm optimization algorithm is initialized to a group of random particles (random solution), and then find the optimal solution through iteration. In each iteration, the particle updates itself by keeping track of the two extremes; the first one is the optimal solution found by the particle itself, which is called the individual extreme; the other extreme is the optimal one found so far for the entire population, this extreme is the global maximum. Alternatively, instead of using the entire population but only one of them as a particle's neighbor, the extreme value in all neighbors is the local extreme.

Based on the wireless location network, the power allocation algorithm designed in this paper is as follows.

Step 1: Initialize the parameters of Particle swarm optimization. $c_1 = 1.4962$, $c_2 = 1.4962$, $w = 0.7268$.

Step 2: According to the above derivation, calculate the square position error bound of each agent in the wireless location network. Each agent is located with all other nodes, including the anchor and the agent. In this process, the normalized transmit power of the node needs to be limited according to the constraints.

Step 3: Calculate the square position error bound of the whole wireless location network, and get the function of the square position error bound of the wireless location network and the normalized transmitting power of the agents as the objective function of particle swarm optimization.

Step 4: Substituting the objective function and constraints into the particle swarm optimization algorithm for optimal solution. All vectors to be optimized form the dimension of power, i.e.

$$\mathbf{p}_j = (p_1, p_2, \ldots p_j) \tag{12}$$

In order to obtain the global optimal solution of power, we set the algorithm precision as 10^{-6}, the maximum number of iterations is 250. All the algorithm parameters are given in Table 1. N represents the number of particles. c_1 and c_2 mean the learning factors of Particle swarm optimization. w represents inertia weight of Particle swarm optimization.

4 Simulation Analysis

The typical ranging signals are UWB signal (CM1 in the IEEE 802.15.4a channel model). And the pulse width of 2 ns Gaussian pulse signal (occupied bandwidth of 3.1–3.6 GHz) as the transmit waveform. Other parameters are given in Table 2.

Set up such a wireless location network, there are some agents in the network (label) randomly distributed in a square area, the square area is U [0,10] * [0,10]. Also set 4 anchors (base station). The total power is normalized to unity and the power peak for a single agent is 0.4. According to Table 2, the range channel gain is 10^{15}. The path attenuation factor is 2.

Table 1 Algorithm parameters

Parameters	Value
N	30
C_1	1.4962
C_2	1.4962
w	0.7268
Algorithm precision	10^{-6}
Maximum number of iterations	250

Table 2 UWB signal parameters

Parameters	Value
Bandwidth (W)	500 MHz
Wavelength (v)	0.1 m
Ratio of DP ($\bar{\gamma}_{DP}$)	0.2
Path overlap factor ($\bar{\chi}_{kj}$)	0.32
Power spectral density (PSD)	41.3 dBm/MHz
Noise power spectral density (N_0)	−174 dBm/Hz

In order to compare the impact of the transmit power of the anchors and the agents on the positioning accuracy of the network, different transmit power is used for the anchor and the agent. In this case, when the agent and the anchor use different transmit power, the square position error bound of the network with respect to the transmit power can be obtained, as shown in Fig. 2.

It can be seen from Fig. 2 that the influence of power of anchors and agents on the square position error bound of network is different. Obviously, when the transmit power of the anchor and the agent reach the maximum, the square position error bound of the location network can reach the minimum, and the network's positioning accuracy reaches the highest. However, it is not clearly from Fig. 2 that the anchor and the agent have different influences on the square position error bound of the positioning network. In order to know more clearly the different influences of these two on the network positioning accuracy, a location network with 4 agents and 4 anchors will be used in simulation. The location of anchors and agents is shown in Fig. 3. Red dots represent anchors, while blue dots represent agents.

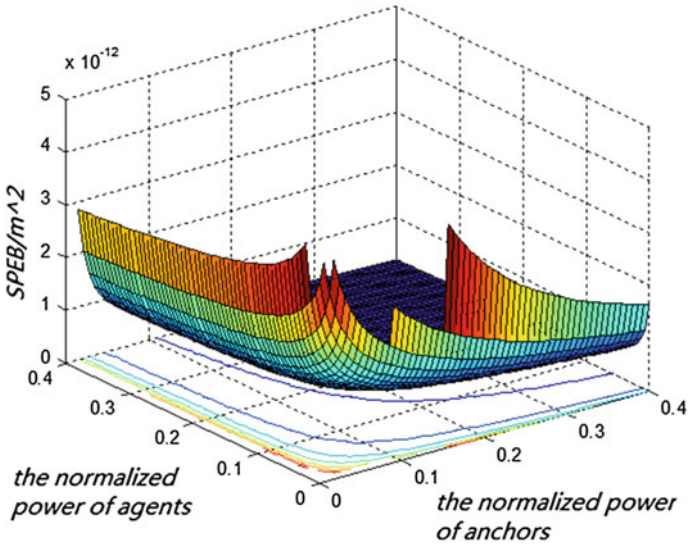


Fig. 2 Effect of signal transmit power on positioning accuracy

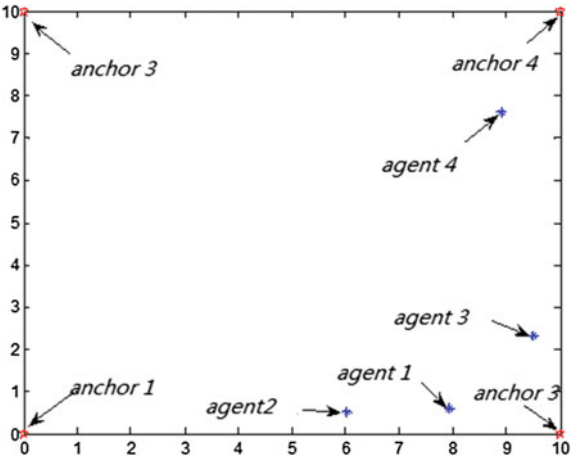


Fig. 3 Location of anchors and agents

Table 3 Optimization results

Optimized object	SPEB (m^2)
Anchor	2.968038260317849e−13
Agent	2.682957452642523e−13

First, the agents of the network use the same transmission power, and the anchors use the algorithm in this paper to optimize the power allocation. Then the corresponding power allocation value of the anchors can be obtained and the square position error bound of the network can be obtained. Then, the anchors of the network use the same transmission power, and the agents use the algorithm in this paper to optimize the power allocation, then the corresponding power allocation value of the agents can be obtained and the square position error bound of the network can be obtained. In this way, the influence of different power distributions on the network positioning accuracy can be compared by comparing the square position error bound of the two. It can be seen that the square position error bound of the network obtained by optimizing the power of the agents is smaller than the gain of the optimized anchors. According to Table 3, It can be seen that the power of the optimized agents is more conducive to improving the network positioning accuracy.

Figure 4 shows the power allocation of anchors. It can be seen that the power allocated by anchor 1 is the lowest and the power allocated by anchor 4 is the highest. As can be seen from Fig. 3, the anchor 1 and the agent 1 and 2 approximate the geometric relationship in a straight line. In this way, the anchor 1 contributes little to improving the positioning accuracy of the two agents, so the power allocated to the anchor node 1 is also reasonable. Anchor 2 and 4 have a similar power allocation because all the agents are more concentrated in the location between these two anchors so the power allocated to anchor 2 and 4 is similar.

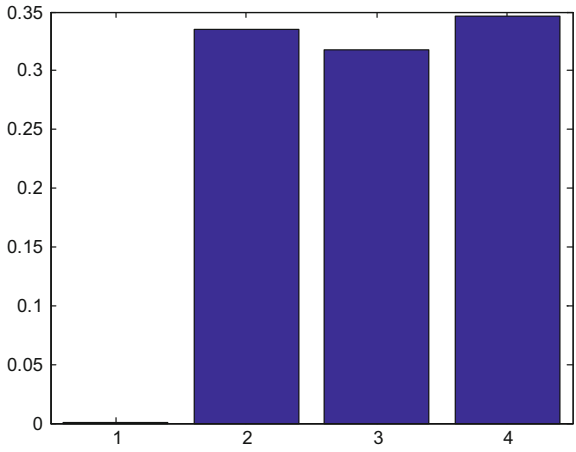


Fig. 4 Power allocation of anchors

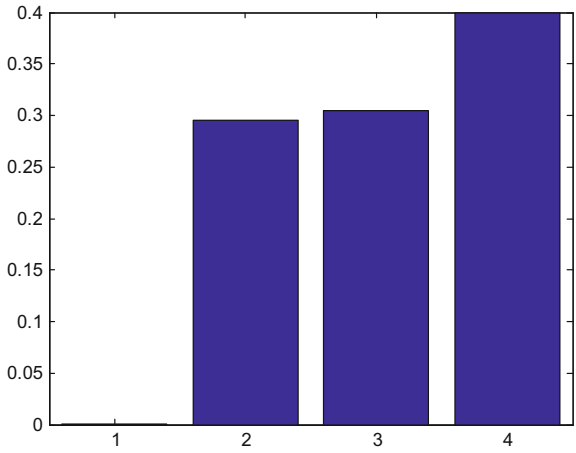


Fig. 5 Power allocation of agents

Figure 5 shows the power allocation of the agents. It can be seen that the power allocated by the agent 1 is the lowest and the power allocated by the agent 4 is the highest. In fact, since the distance between agent 1 and 2, agent 1 and 3 is relatively close, the agent 1 is at an intermediate position between the agent 2 and 3, the transmission power required for ranging between agent 1 and them can be lowered, and the agent 4 is far away from other agents, the transmission power required for ranging between the agent 4 and other agents is obviously higher.

In this paper, UWB ranging technology is used to locate the agents of wireless location network. The square position error bound is introduced as the standard of positioning accuracy of wireless location network. The function of the transmit power of ranging signal and the square position error bound of network is established as the objective function of optimization and the necessary constraints are added. Particle swarm optimization is used to iteratively solve the optimization

objective function to get the optimized power configuration solution of the network. The optimization results show that by optimizing the transmission power of network nodes, the positioning accuracy can be significantly improved. At the same time, the transmit power of the anchor and the agent has different impact on the positioning accuracy, which needs to be optimized according to the actual situation. Furthermore, when the anchor and the agent of the location network perform power optimization at the same time, the positioning accuracy of the network will be greatly improved.

References

1. Ping L, Liu D, Qian J (2017) A review of indoor location techniques and applications. Positioning Time Track 4(3):2–9
2. Pang Y, Qiao J (2005) Discussion on UWB wireless location technology. Telecommun Lett 5 (11):13–15
3. Tong K, Tian S, Li G (2014) Applications of UWB in wireless location technology. In: Annual Meeting of China Satellite Navigation
4. Tong K, Zhou X, Li G et al (2015) Application of UWB in wireless location technology. J Navig Positioning 5(1):10–14
5. Yang D, Tang X, Li B et al (2015) Research overview of indoor positioning technology based on ultra-wideband. Positioning Syst 40(5):34–40
6. Zhu X, Hei Y, Yu Q et al (2009) Introduction of pulsed ultra-wideband positioning technology. China Sci Inf Sci 9(10):1112–1124
7. Penna F, Caceres MA, Wymeersch H (2010) Cramer-Rao Bound for hybrid GNSS-terrestrial cooperative positioning. IEEE Commun Lett 14(11):1005–1007
8. Chen J, Dai W, Shen Y et al (2016) Power management for cooperative localization: a game theoretical approach. IEEE Trans Signal Process 64(24):6517–6532
9. Liu L (2008) RDM positioning based on UWB pulse signal. Harbin Institute of Technology, Harbin
10. Wymeersch H, Lien J, Win MZ (2009) Cooperative localization in wireless networks. Proc IEEE 97(2):427–450
11. Li B, Xia W, Guo Q et al (2015) Optimized power distribution scheme based on lower square error of square position. Comput Appl Softw 5(6):140–143

Error Compensation of Geomagnetic Field Measurement Used in Geomagnetic Navigation

Binfeng Yang, Chi Li, Junmin Xu and Huan Sun

Abstract Getting accurate geomagnetic field value is the prerequisite and foundation for the application of geomagnetic navigation. However, the measuring precision of geomagnetic field is badly influenced by the coupling effect between the field of earth and interference errors. The magnetic compensation has become the key factor which restrict the progress of geomagnetic navigation. To overcome the problem that environmental magnetic field will influence the error compensation result of geomagnetic field measurement, the geomagnetic field was simulated by ANSYS. The parameterized geomagnetic measurement compensation model was constructed by analyzing the hard iron and soft iron errors. Those parameters of the model were estimated using least squares based algorithm. Simulation results show that this compensation method is easy to be used and which can achieved higher error compensation ability.

Keywords Geomagnetic navigation · Geomagnetic field measurement
Least squares · Error compensation

1 Introduction

Geomagnetic field is a natural resources of earth, which provides the coordinate system for aviation, aerospace and navigation. Geomagnetic navigation has been used for a long time in history; for instance, the compass was used to locate in China and Colombo discovered the new world with compass. Geomagnetic information used in flight navigation has the advantages of passive, non-radiation and anti-interference, so it can be used in any weather and any location to supply position and attitude information, which has become one of the research focus in navigation [1–3].

B. Yang (✉) · C. Li · J. Xu · H. Sun
Information and Navigation College, Air Force Engineering University,
Xi'an 710077, China
e-mail: bf_yang@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_66

The principle of geomagnetic navigation relies on matching the data achieved by magnetometers with the geomagnetic reference map. As the geomagnetic field is vector, then, a vector field in any place of the world is different from the others. As a result, navigation and position will be realized when the geomagnetic field is measured.

Strapdown three-axis magnetometers are used for measuring the strength and direction of magnetic fields. However, the output of magnetometers is influenced and corrupted by errors [4–7], which can be divided into two categories, one kind of error is caused by the configuration and material of the magnetometers itself, such as constant bias, scale factor deviations and nonorthogonality of the sensor axes. This kind of error is called the instrument error. The other error is the interference error, which includes two classes: hard iron interferences and soft iron interferences. The hard iron error is generated by residual magnetism of ferromagnetic material, it adds a constant bias to the magnetometers output. The soft iron interference interacts with the earth's magnetic field and generates bias that changes as the vehicle's orientation relative to the local magnetic field changes. As instrument error can be eliminated by calibration of magnetometers, then, only the calibration of vehicle's interference magnetic field is discussed in this paper.

Stated thus, getting accurate geomagnetic field value is the prerequisite and foundation for the application of geomagnetic navigation. However, the measuring precision of geomagnetic field is badly influenced by the coupling effect between the field of earth and interference errors. The magnetic compensation has become the key factor which restrict the progress of geomagnetic navigation.

There are two ways to carry out the calibration of magnetometers. The first method is kalman filter, including unscented kalman filter (UKF) and extended kalman filter (EKF), which need to linearize the equation, and the estimation is repeated until convergence is achieved. This method requires a good initial guess values, which may lead to divergence if the initial parameter is not correctly selected. The other method is ellipse fitting, this calibration method is based on the fact that the locus of error-free measurements from two perpendicularly mounted magnetometers is a circle or a sphere if a three-axis magnetometers is used. The radius of the circle is equal to the magnitude of the horizontal component of the local earth's magnetic field vector. Hard iron error shifts the origin of the circle, and soft iron error changes the circle into a ellipse. Then, the geomagnetic measurement error can be calibrated by using the relation between the parameters of ellipse and coefficient of the vehicle's magnetic field. However, this algorithm is complicated in parameter estimation.

To overcome the above problems, the influence of hard iron errors and soft iron errors for geomagnetic measurement was analyzed firstly, and the geomagnetic field was simulated by ANSYS, then, the parameterized geomagnetic measurement compensation model was constructed and least squares based algorithm was used in this paper. Simulation results show that this compensation method is easy to be used and which can achieve higher error compensation ability.

The rest of the paper is organized as follows: Sect. 2 analyzes the influence of hard iron and soft iron errors for the measurement of geomagnetic field. Section 3

presents the parameter estimation method based on least squares algorithm. Geomagnetic field is designed and modeled using finite element software ANSYS in Sect. 4. The analysis of calibration ability is shown in Sect. 5. Finally, The conclusions are outlined in Sect. 6.

2 Magnetometer Measurement Errors

The magnetometer is installed in navigation carriers to measure magnetic field, and the carrier is usually made of metal, so it may be magnetized. Meanwhile, cutting magnetic field lines for the movement of metal carriers will bring the induced magnetic field. Hard iron errors and soft iron errors of the carrier mostly affect the measurement results of the magnetometer [8, 9].

Hard iron error is generated by the residual magnetism of ferromagnetic materials on the carriers [10]. The common hard iron materials include permanent magnet and high-carbon steel. Because the hard iron materials have large coercive force, the magnetic field generated by magnetizing goes slowly with time's change. The magnetometer is usually strapdown with carrier so that the coordinate system of carrier and magnetometer can be coincided with each other. As a result, no matter how the carrier's attitudes change, hard iron error's projection component in three coordinate axes is not changed. Usually, hard iron error can be represented by H_h.

$$H_h = \begin{bmatrix} H_{hx} & H_{hy} & H_{hz} \end{bmatrix}^T \tag{1}$$

Soft iron error comes from carrier's soft magnetic materiel which is magnetized by geomagnetic field. Owing to small coercive force, the soft iron errors change with the size of geomagnetic field and carrier's attitude, and soft iron error can be represented by H_s.

$$H_s = DH_e = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} H_e \tag{2}$$

In the above formula, D represents coefficient matrix of induced magnetic field. When carrier is fixed, the elements of this matrix is constant. He represents real value of local geomagnetic field.

Based on theory analysis, when geomagnetic measurement is not affected by errors, its measurement track is a circle in two-dimensional measurement. Since hard iron error is fixed, it makes the track of geomagnetic measurement away from the center of a circle, the offset is the value of hard iron error. Soft iron errors also affect geomagnetic measurement. As a result, the track becomes ellipse rather than circle. The principle of three-dimensional measurement is the same as two-dimensional, so the track is changed from ellipse to ellipsoid under the

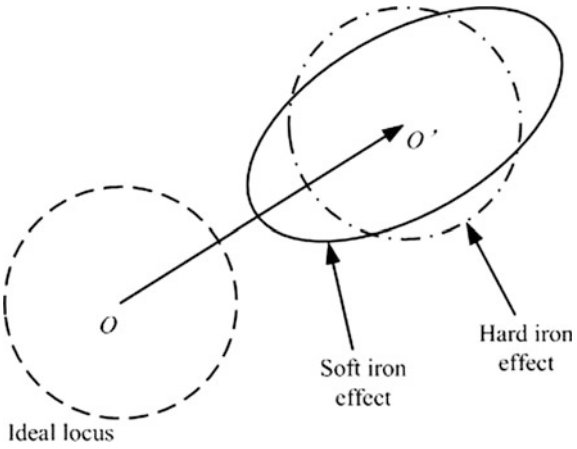


Fig. 1 Hard and soft iron errors on geomagnetic field measurement locus in 2D

influence of hard and soft iron errors. The effect of hard and soft iron errors on geomagnetic field measurement is shown in Fig. 1.

Considering the effect of hard and soft iron errors, the output of magnetometer on geomagnetic field measurement can be expressed as:

$$H_m = H_e + H_h + H_s \tag{3}$$

From Eqs. (1) to (3), we can get the equation as follows:

$$\begin{aligned} H_m = (I_{3\times3} + D)H_e + H_h = KH_e = \begin{bmatrix} K_1 & K_2 & K_3 \\ K_4 & K_5 & K_6 \\ K_7 & K_8 & K_9 \end{bmatrix} H_e \\ + \begin{bmatrix} H_{hx} \\ H_{hy} \\ H_{hz} \end{bmatrix} \end{aligned} \tag{4}$$

From Eq. (4), it can be concluded that nine parameters of coefficient matrix and three parameters of hard iron errors must be solved if we want obtain the true value of geomagnetic field from the measured value of magnetometer. Hence, the magnetometer must be measured under four attitudes at least in order to solve twelve error compensation coefficients.

3 Parameter Estimation Based on Least Squares

To solve twelve unknown parameters in Eq. (4), we can transform and simplify the equation and get formula AX = B. Then, these parameters can be estimated by the least squares algorithm [11]. The solution is $X_{LS} = \left(A^T A\right)^{-1} A^T b$. Where,

$$A = \begin{bmatrix} 1 & 0 & 0 & H_{ex} & 0 & 0 & H_{ey} & 0 & 0 & H_{ez} & 0 & 0 \\ 0 & 1 & 0 & 0 & H_{ex} & 0 & 0 & H_{ey} & 0 & 0 & H_{ez} & 0 \\ 0 & 0 & 1 & 0 & 0 & H_{ex} & 0 & 0 & H_{ey} & 0 & 0 & H_{ez} \end{bmatrix} \quad (5)$$

$$X = [\, H_{hx} \quad H_{hy} \quad H_{hz} \quad k_1 \quad k_4 \quad k_7 \quad k_2 \quad k_5 \quad k_8 \quad k_3 \quad k_6 \quad k_9 \,]^T \quad (6)$$

$$B = [\, H_{mx} \quad H_{my} \quad H_{mz} \,]^T \quad (7)$$

In the above formula, H_{ex}, H_{ey}, H_{ez} are the measured value of geomagnetic field without interference. H_{mx}, H_{my}, H_{mz} are the measured value with interference.

4 Simulation of Geomagnetic Field

According to the analysis of error compensation, several measurement results must be obtained with different attitudes of magnetometer to estimate the parameters in error compensation model. Actually, the magnetometer is placed in open side to measure geomagnetic field. However, the results are always interfered by surrounding environment. An alternative method is that Helmholtz coils are used to generate simulated geomagnetic field. Then, these coils are putted into an cylinder where other magnetic field is shielded. However, this method costs lot. In this paper, the finite element analysis is performed to establish the models of geomagnetic field and simulation is performed by using ANSYS, hard and soft magnetic materials are added into the model to simulate the interference source. As a result, the magnetic field that generated by interference can be obtained accurately, thus, it increases the estimation precision of parameter estimation.

Helmholtz coils consist of a pair of coaxial parallel placed coils with the same size, the distance of two coils is equal to the radius of Helmholtz coils. As the constant current is applied to the coils respectively, a large-scale uniform magnetic field will be generated in the center of Helmholtz coils. This is the reason that Helmholtz coils can be used to simulate geomagnetic field.

The simulations are conducted by using ANSYS. Three pairs of Helmholtz coils are needed to generate three directions of geomagnetic field. The simulation model is shown in Fig. 2, the inner, the medial and the outer coils are used to generate the Y, X, Z component of geomagnetic field respectively. The magnetometer is located in the center of the model to measure three components of geomagnetic field, a metal material plate is put at the bottom of the magnetometer to simulate the source of interference. In the simulation model, the Helmholtz coils are much bigger than magnetometer and the interference source, so, its can not be seen clearly in Fig. 2. The parameters of the Helmholtz coils are shown in Table 1.

Figure 3 shows the magnetic field distribution on the axis of three pairs of Helmholtz coils. It can be seen that a uniform magnetic field is generated in the center of three pairs of Helmholtz coils. The magnetometer and the interference

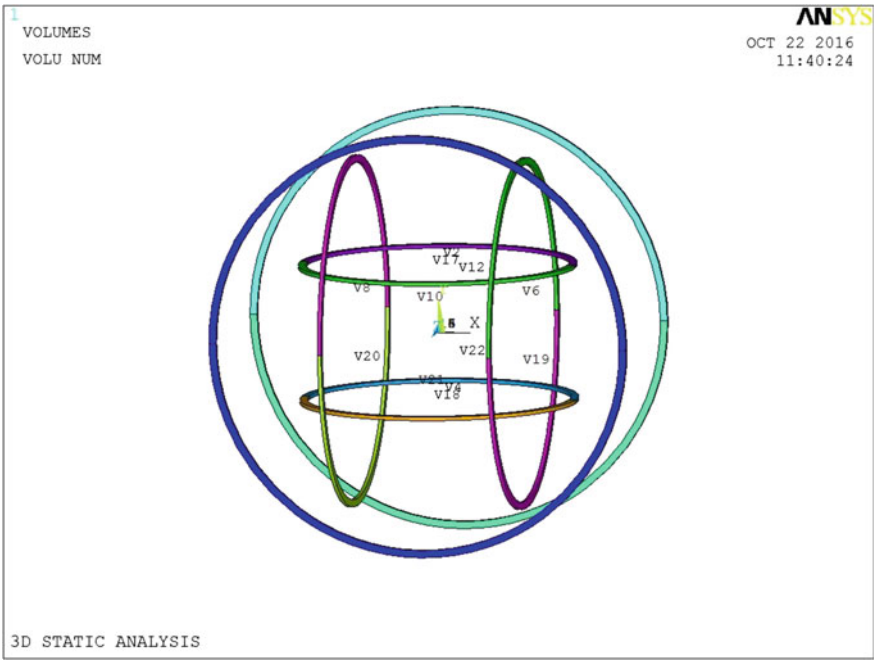


Fig. 2 The geomagnetic field simulation model by Helmholtz coils

Table 1 The parameters of the Helmholtz coils

	Inner coils	Medial coils	Outer coils
Inner radius/mm	180	230	280
Outer radius/mm	190	240	290
Turns	100	100	100
Current/A	0.05	0.05	0.05

source are just located in this uniform field zone, so, the simulation model can be used to simulate geomagnetic field.

5 The Simulation Analysis of Compensation Algorithm

In simulation, the Helmholtz coils is rotated with the center and magnetometer keep still to obtain the measurement magnetic field with interference at different attitudes,, then, a simulation is performed with a 10° rotation of coils. The interference source used in the simulation is iron material, the relative permeability is 100, the resistivity is 2.65×10^{-8} Ωm, and the size is 20 mm × 20 mm × 1 mm.

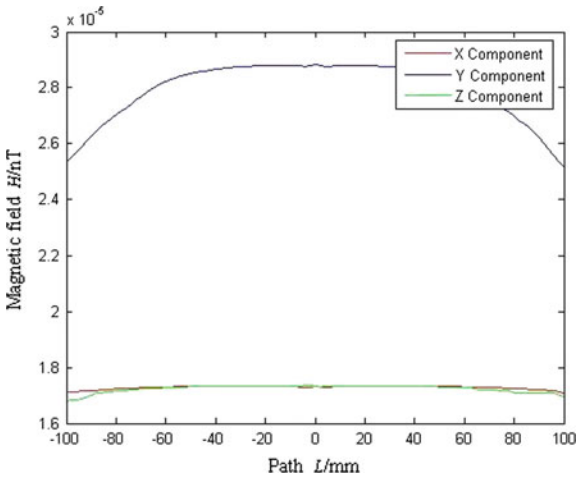


Fig. 3 The magnetic field distribution of three pairs of Helmholtz coils on the axis

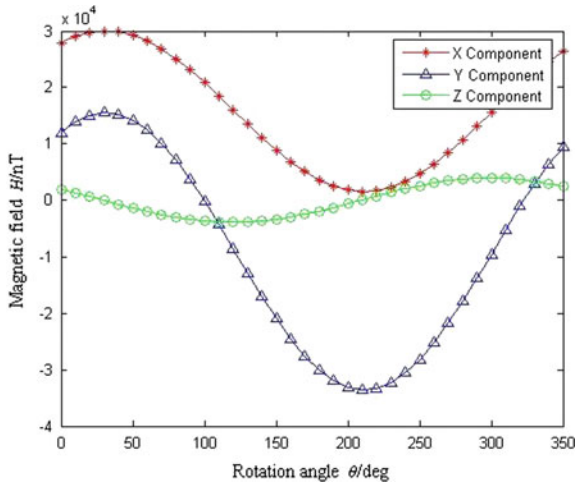


Fig. 4 Geomagnetic field component

Figures 4 and 5 show the component of X, Y, Z and the total value of the geomagnetic field measured by magnetometer with interference respectively. From these figures, it can be found that the measured value of geomagnetic field is affected greatly by interference. The total value of geomagnetic field changes a lot with the spinning attitude of magnetometer. The variation range can reach nearly 16,000 nT, which can't be accepted in the precision measurement of geomagnetic field.

According to Eqs. (5) and (7), we need measure the values of geomagnetic field under four spinning attitudes at least with and without errors to compensate magnetic field. Then, two simulations are performed respectively when the magnetometer rotates per 60°, so, six sets of measured values with and without errors

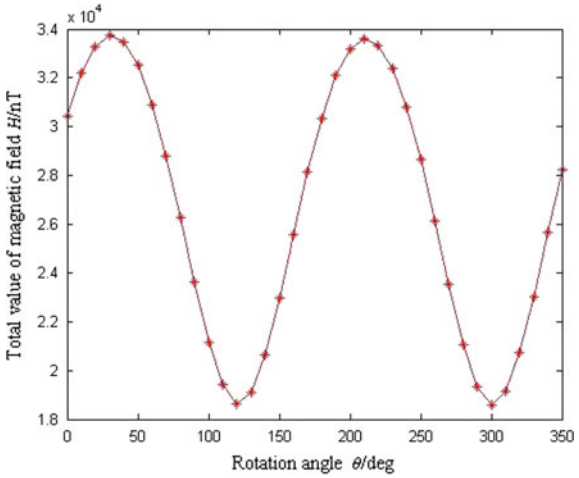


Fig. 5 The total value of geomagnetic field

can be obtained when the material of interference source is set as iron and air in the model respectively. As shown in Table 2, H_{mx}, H_{my}, H_{mz}, H_{ex}, H_{ey}, H_{ez} are the measured value of geomagnetic field with and without errors respectively.

Next, we can use least squares to solve the Eq. (6), where,

$$K = \begin{bmatrix} -3.6073 & 6.9694 & 0.0013526 \\ -4.613 & 11.099 & -0.001984 \\ 1.172 & -0.71669 & 1.1716 \end{bmatrix} \begin{bmatrix} H_{hx} \\ H_{hy} \\ H_{hz} \end{bmatrix} = \begin{bmatrix} 30077 \\ 11624 \\ -2994 \end{bmatrix}$$

After the parameters are estimated, the geomagnetic field with errors can be compensated. Figure 6 shows the three-dimensional locus of geomagnetic field before and after the error compensation. It can be seen that the measured value of magnetometer deviates seriously from the true value of geomagnetic field without compensation. However, the calibrated value of geomagnetic field coincides with the true value very well after compensation, which indicates that the compensation algorithm proposed in this paper can compensate the interfered geomagnetic field.

To illustrate the problem clearly, Fig. 7 shows the two-dimensional locus of magnetometer. It can be seen that the locus of measured value of magnetometer is a

Table 2 The measured value of geomagnetic field with and without errors

Rotation/deg	0	60	120	180	240	300
H_{ex}/nT	3315	3336	1868	434	405	1846
H_{ey}/nT	1405	1460	−1048	−3577	−3630	−1137
H_{ez}/nT	1736	−1617	−3333	−1732	1625	3360
H_{mx}/nT	27,919	28,210	16,032	3581	3317	15,496
H_{my}/nT	11943	12,413	−8607	30,055	−30,566	9507
H_{mz}/nT	1916	−2024	−3959	−1951	1984	3922

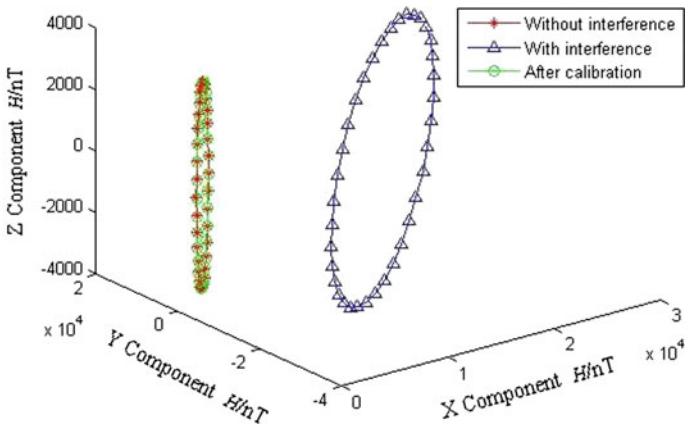
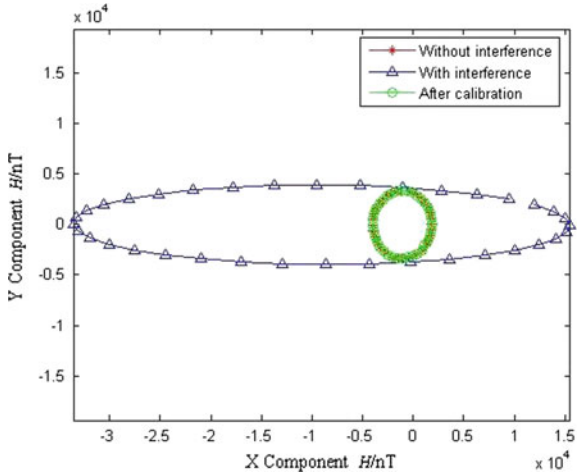


Fig. 6 3D locus of magnetometer

Fig. 7 2D locus of magnetometer



circle, whose radius is the total value of geomagnetic field without interference. The locus will change to an ellipse whose center is away from the center of a circle with interference, which indicates the magnetometer is affected by hard and soft iron errors at the same time. The locus of the calibrated value becomes a circle again. The result of compensation consists with the theory analysis, which also confirms that the method adopted by the paper is correct.

Figure 8 shows the total geomagnetic field values before and after compensation. It can be seen that great errors arise when the measurement of geomagnetic field is interfered. However, the measured value nearly coincides with the true value no matter how the magnetometer spins and the attitudes change after error compensation.

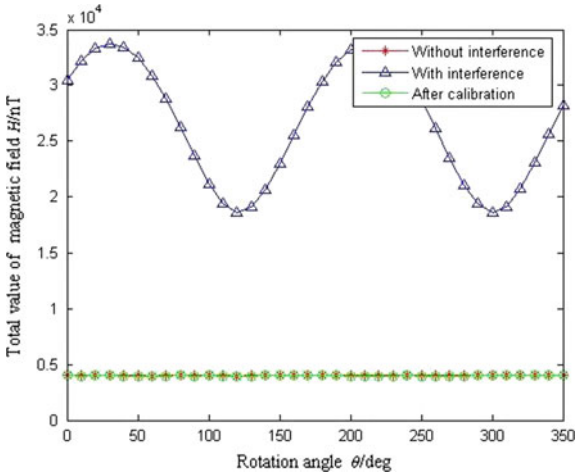


Fig. 8 Geomagnetic field measurement results with and without error compensation

Table 3 Errors analysis

Rotation (°)	180	190	290	300	310	320
Total field without interference/nT	3998.7	3999.5	4000.1	3999.2	3998.8	3999.4
Total field with interference/nT	30,331	32,094	19,346	18,598	19,123	20,731
Total field after compensation/nT	3996.6	3999.8	3997	3996.8	3999.2	4000.6
Errors before calibration/nT	26,332.3	28,094.5	15,345.9	14,598.8	15,124.2	16,731.6
Errors after calibration/nT	−2.1	0.3	−3.1	−2.4	0.4	1.2

To analyze the result of compensation quantitatively, six sets data is selected for analysis, and the results are shown in Table 3.

According to Table 3, the measurement errors were decreased from more 20,000 nT to several nT with calibration, thus the precision of measurement increased significantly.

6 Conclusions

To overcome the problem that environmental magnetic field will influence the measurement result of geomagnetic field, a geomagnetic measurement compensation algorithm based on the analysis of hard and soft iron errors is proposed in this

paper. Then, the parameters of compensation model are estimated based on least squares on the foundation of the geomagnetic field simulation by using ANSYS. Simulation results show that the compensation algorithm is simple and easy to be used, at the same time, which can achieve higher compensation ability. The research results can build the foundation for the precision measurement and matching of geomagnetic navigation.

Acknowledgements This research was financially supported by the National Natural Science Foundation of China (Grant No. 51577191), Shaanxi Province Science Foundation (Grant No. 2015Jm5147).

References

1. Mohamadabadi K, Coillot C, Hillion M (2013) New compensation method for cross-axis effect for three-axis AMR sensors. IEEE Sens J 13(4):1355–1361
2. Leinweber HK, Russell CT, Torkar K (2012) In-flight calibration of the spin axis offset of a fluxgate magnetometer with an electron drift instrument. Meas Sci Technol 23:1–13
3. Zhang Q, Li J, Chen DX et al (2014) Method and experiment for compensating the interferential magnetic field in underwater vehicle. Measurement 47:651–657
4. Pang HF, Luo ST, Zhang Q et al (2013) Calibration of a fluxgate magnetometer array and its application in magnetic object localization. Meas Sci Technol 24:1–8
5. Wang XL, Zhang Q, Li HN (2014) An autonomous navigation scheme based on starlight, geomagnetic and gyros with information fusion for small satellites. Acta Astronaut 94: 708–717
6. Pang HF, Zhang Q, Li J et al (2014) Improvement of vector compensation method for vehicle magnetic distortion field. J Magn Magn Mat 353:1–5
7. Wu ZT, Wu YX, Hu XP et al (2013) Calibration of three-axis magnetometer using stretching particle swarm optimization algorithm. IEEE Trans Instrum Meas 62(2):281–292
8. Wu ZT, Hu XP, Wu MP et al (2013) Attitude-independent magnetometer calibration for marine magnetic surveys: regularization issue. J Geophys Eng 10:1–9
9. Wu Z T, Wu Y X, Hu X P et al (2012) Calibration of strapdown three-axis magnetometer measurement error compensation of geomagnetic field based on total least squares. ACTA Armamentarii 33(10):1202–1209 (in Chinese)
10. Li J, Pan M C, Luo S T, Luo F L(2013) Study on the application of semi-parametric model in vehicle interferential magnetic field compensation. Chin J Sci Instrum 34(9):2147–2151 (in Chinese)
11. Zhao GR, Ji SK, Yang F, Lin KQ (2012) Magnetic signal detection error analysis and compensation algorithm. Syst Simul Technol 8(2):153–157 (in Chinese)

Research on Barometric Altimeter Assisting GNSS Navigation for Urban Environments

Han Guo, Dongyan Wei, Guangzhou Ouyang and Qifeng Lai

Abstract In the satellite navigation system, the altitude accuracy of an regular receiver is usually poor. This paper proposed a method that uses satellite positioning assisting differential barometric altimetry under the condition without the data of pressure base stations. It also proposed using barometric altimetry assisting satellite positioning when there're fewer satellites. And we discussed the method of RAIM algorithm with differential barometric altimetry. The result shows: the improved method of utilizing the differential barometric altimetry to assist satellite positioning can locate which there are 3 visible satellites. And it can also improve positioning accuracy when there are fewer visible satellites and the detection rate of the single-satellite fault is improved.

Keywords Differential barometric altimetry · Satellite positioning RAIM

1 Introduction

In the satellite navigation system, affected by constellation configuration, the vertical accuracy of the single point positioning of a regular receiver is usually poor. Especially in typical urban complex environments, the satellite signal is susceptible to occlusion, multipath and other factors which cause part of the satellite signal abnormal even unusable.

Based on the principle that atmospheric pressure varies with altitude, barometric pressure can be used to measure altitude. Currently there are two main ways:

H. Guo
University of Chinese Academy of Sciences, No. 19(A) Yuquan Road, Shijingshan District, Beijing 100049, China

H. Guo · D. Wei (✉) · G. Ouyang · Q. Lai
Academy of Opto-Electronics, Chinese Academy of Sciences, No. 9, Deng Zhuang South Road, Haidian District, Beijing 100094, China
e-mail: weidongyan@aoe.ac.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_67

(1) Use international standard atmospheric pressure values for barometric altimetry; (2) Establish a pressure reference station, use the station's pressure data and the user's barometric pressure data for differential barometric altimetry [1]. However, the establishment of pressure reference station cost much and transmitting data is inconvenient and the used range of distances is limited. So how to achieve the differential barometric altimetry without establishing a pressure base station is one of the questions we are going to explore.

The method of satellite positioning assisted by barometric altimetry is mainly used: (1) Use altitude combined with receiver observation information to locate when there are 3 visible satellites [6–8]; (2) Use altitude combined with receiver observation information to improve positioning accuracy when there are more than 3 visible satellites but constellation configuration is poor; (3) Use altitude combined with satellite integrity monitoring (RAIM) to identify the availability of satellite observations.

References [1, 2] proposed to use the international standard atmospheric pressure data to establish a model for barometric altimetry. This method is accessible to data, but due to the changing trend of atmospheric pressure in different areas is different, the vertical accuracy measured directly according to the international standard atmospheric pressure model is poor and fixed point error drift is up to meters or even hundreds of meters [5]. References [3–5] all proposed that use differential barometric altimetry to measure altitude to improve vertical accuracy. But building pressure base stations and transmitting data cost much and because the location of base stations is fixed, it limits the method's use range. Literature [4] proposed to transmit air pressure data through mobile communication base stations, but it still needs to install barometric sensors on stations. The relevant data is transmitted to the mobile terminal through the base station transmission link.

Literature [6–8] proposed that they combined altitude and receiver observation information with ellipsoid model for positioning and satellite integrity monitoring. However, the ellipsoidal model algorithm itself has model error.

Based on the above questions, we proposed a method of inversely predicting the air pressure at the reference point through receiver observations over a period time to do differential barometric altimetry. Combine altitude with GPS and through establishing a virtual satellite at geocentric, we solve the positioning problem when there are fewer visible satellites. And through the method of receiver positioning results feedback correcting altitude, reduce the error drift of satellite positioning assisted by differential pressure altimetry. And solve the problem of satellite integrity monitoring by using the deviation of position solution's altitude and pressure altitude.

2 Differential Barometric Altimetry Principle

Atmospheric pressure values obtained with the international standard atmospheric model are relative to sea level. Promotion to the concept of "difference", that is to set a reference point, the height of the user's position is equal to the sum of the height of the known reference point and the calculated height difference. Based on the ideal gas state equation, a differential barometric altimetry formula can be derived from the international standard atmospheric pressure [5]:

$$h = h_0 + 18410\left(1 + \frac{t_m}{273.15}\right)\lg\frac{P_0}{P} \tag{1}$$

where h_0 is the height of the reference point, P_0 is the pressure value of the reference point, t_m is the average Celsius between the reference point P_0 and the measurement point P, $t_m = (t_0 + t_R)/2$, t_0 is the temperature of the reference point and t_R is the temperature of the measurement point.

3 Analysis of the Key Factors Affecting Differential Barometric Altimetry

Differential barometric altimetry requires using barometric sensors to measure temperature and pressure, so it's influenced by the surrounding more. In this chapter, we focus on the height accuracy of differential pressure altimetry influenced by the environment of barometer and the condition at high-speeds. Based on the differential barometric altimetry principle, we use a professional calibrated VAISALA barometer for experimental verification.

We first study the case of the same speed. The barometer was placed in three states: in the car, outside the car (no pressure regulator's environment), outside the car (with pressure regulator's environment). Test the accuracy of differential barometric altimetry influenced by the environment of barometer. Pressure regulator is used to stabilize the rapidly changing airflow around the barometer under dynamic conditions. Table 1 is the mean error, variance and mean square error of differential barometric altitude error when a barometer is measured in different

Table 1 Mean error, variance and mean square error of differential barometric altitude error of the barometer measured in different environments

	Outside the car (with pressure regulator)	Outside the car (no pressure regulator)	In the car
Mean error/m	2.497	3.090	2.796
Variance	3.853	7.361	7.710
Mean square error	0.077	0.100	0.096

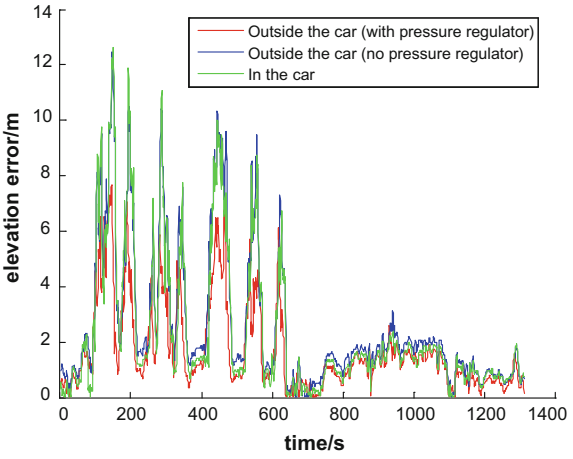


Fig. 1 Differential barometric altitude error of the barometer measured in different environment

environments. Figure 1 is the accuracy of differential barometric altimetry when a barometer is measured in different environments.

From Fig. 1 and Table 1, it can be intuitively perceived that differential pressure altitude precision of barometer with pressure regulator outside the car is superior to that of the other two environments.

Second, we study differential pressure altitude accuracy at different speeds. Set the speed to 4 segments, they are below 60, 65–70, 75–80 and 85–90 km/h, respectively. We choose the pressure sensor outside the car (with pressure regulator). Its altitude error is shown in Fig. 2, the corresponding mean altitude error and variance shown in Table 2.

From the chart we can see, in the low-speed driving state (60 km/h speed or less), altitude error is maintained within 2 m, and meet the needs of altitude accuracy. However, in the high-speed driving state, altitude error significantly

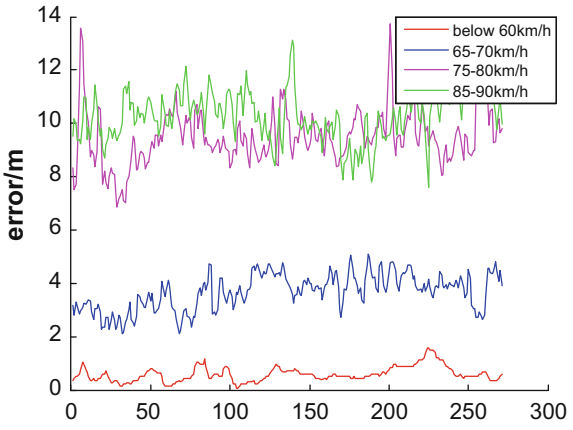


Fig. 2 Differential barometric altitude error at different speeds

Table 2 Mean error and variance of the differential barometric altitude error at different speeds

Speed (km/h)	<60	65–70	75–80	85–90
Mean error/m	1.51	3.78	9.3	10.36
Variance	2.13	0.98	1.02	1.07

increased. Mainly because in high-speed state, air pressure around the sensor changes rapidly so that it is impossible to accurately calculate the altitude. So this method needs to be further optimized for using high-speed moving speeds more than 60 km/h.

4 Satellite Positioning Assisted by Differential Barometric Altimetry

4.1 Principle of Satellite Positioning Assisted by Altitude

The traditional GPS positioning solution is based on pseudo range positioning principle. Due to the specific city environment, such as urban canyons, tunnels, jungle and tall buildings shelter, the receiver can only receive messages sent by a smaller number of satellites so additional external altitude information plays a key role in positioning. When using differential barometric altimetry aids positioning, equivalent to placing a virtual satellite fs in the center of the earth. Satellite coordinates are set to (0 0 0), the user's estimated position coordinates in earth-centered earth-fixed coordinate system (ECEF) are (x_u y_u z_u). The observation equation is:

$$\begin{cases} \rho_{c1} = \sqrt{(x_1 - x_u)^2 + (y_1 - y_u)^2 + (z_1 - z_u)^2} + b \\ \rho_{c2} = \sqrt{(x_2 - x_u)^2 + (y_2 - y_u)^2 + (z_2 - z_u)^2} + b \\ \cdots \\ \rho_{cj} = \sqrt{(x_j - x_u)^2 + (y_j - y_u)^2 + (z_j - z_u)^2} + b \\ \cdots \\ \rho_{cn} = \sqrt{(x_n - x_u)^2 + (y_n - y_u)^2 + (z_n - z_u)^2} + b \\ \rho_{fs} = \sqrt{x_u^2 + y_u^2 + z_u^2} + 0 \end{cases} \tag{2}$$

In Eq. 2, (x_j y_j z_j) represents the position coordinates of satellite j under ECEF. ρ_{cj} is the distance from the user to the satellite j through the pseudo range correction. b is the equivalent distance generated by user's clock error. n is the number of observation satellites. Equation 2 is linearized using Newton's iteration and least-squares method. The geometric matrix G is shown as Eq. 3:

$$\begin{bmatrix} \frac{-(x_1-x_u)}{r_1} & \frac{-(y_1-y_u)}{r_1} & \frac{-(z_1-z_u)}{r_1} & 1 \\ \frac{-(x_2-x_u)}{r_2} & \frac{-(y_2-y_u)}{r_2} & \frac{-(z_2-z_u)}{r_2} & 1 \\ & \cdots & & \\ \frac{-(x_j-x_u)}{r_j} & \frac{-(y_j-y_u)}{r_j} & \frac{-(z_j-z_u)}{r_j} & 1 \\ & \cdots & & \\ \frac{-(x_n-x_u)}{r_n} & \frac{-(y_n-y_u)}{r_n} & \frac{-(z_n-z_u)}{r_n} & 1 \\ \frac{x_u}{r_{fs}} & \frac{y_u}{r_{fs}} & \frac{z_u}{r_{fs}} & 0 \end{bmatrix} \tag{3}$$

The intermediate variable is:

$$\begin{cases} r_j = \sqrt{(x_j - x_u)^2 + (y_j - y_u)^2 + (z_j - z_u)^2} \\ r_{fs} = \sqrt{x_u^2 + y_u^2 + z_u^2} \end{cases} \tag{4}$$

Pseudo range measurement error matrix is:

$$\begin{cases} \triangle\rho_j = \rho_{cj} - r_j - b \\ \triangle\rho_{fs} = h - h_u \end{cases} \quad j = 1, 2, 3 \tag{5}$$

where h is the altitude of differential barometric altimetry and h_u is the altitude value in the geodetic coordinates (B_u, L_u, h_u) which is the user's estimated position (x_u, y_u, z_u) converted to. The least square estimate of the user location error X is:

$$\Delta X = (G^T G)^{-1} G^T \Delta\rho \tag{6}$$

where $\Delta X = [\Delta x\ \Delta y\ \Delta z\ \Delta b]^T$ is the offset vector of the location and clock error of the client and $\Delta\rho = [\Delta\rho_1\ \Delta\rho_2\ \ldots\ \Delta\rho_j\ \ldots\ \Delta\rho_n\ \Delta\rho_{fs}]^T$ is the ranging error vector.

4.2 Improved Method of Satellite Positioning Assisted by Differential Barometric Altimetry

By the differential barometric altimetry principle of the second chapter, we can see that differential barometric altimetry requires air pressure and altitude for a known reference point. But most of the time, we cannot build a barometric reference station. This section therefore solves the issue of differential barometric altimetry without a barometric reference station.

First, we know from Eq. 1 when we have known the initial point of height h_0 and pressure P_0 over a period of time, and the average height h, the average temperature t_m during this time, the air pressure P_a at the average height h_a can be calculated. The formula is as follows:

$$P_a = \frac{P_0}{10^{\left(\frac{h_a - h_0}{18410 + 67.4t_a}\right)}} \tag{7}$$

Use (h_a, P_a) as the reference point (h_0, P_0) for differential barometric altimetry in the next period of time.

When we could not get a high-precision altitude for some time, (h_a, P_a) calculated in the previous period is taken as the reference pressure-altitude. The real-time altitude information h is obtained from the user's pressure sensor by formula 1. Use this altitude to assist satellite positioning.

But due to the changing atmospheric pressure over time, positioning results will drift. Therefore, we need to correct the real-time altitude h by using the discrepancy between the location aided by altitude and the positioning solved by the receiver on the last moment.

Specific flow charts are shown in Figs. 3 and 4.

Specific steps are as follows:

1. We obtained the initial position information (x_0, y_0, z_0) over a period of time under the ECEF by the receiver, converted into the location (b_0, l_0, h_0) in geodetic coordinate system. The corresponding pressure P_0 was obtained by the mobile pressure sensor.

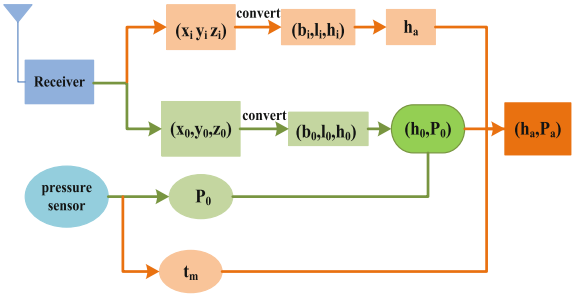


Fig. 3 Flow chart of establishing the altitude-pressure reference point

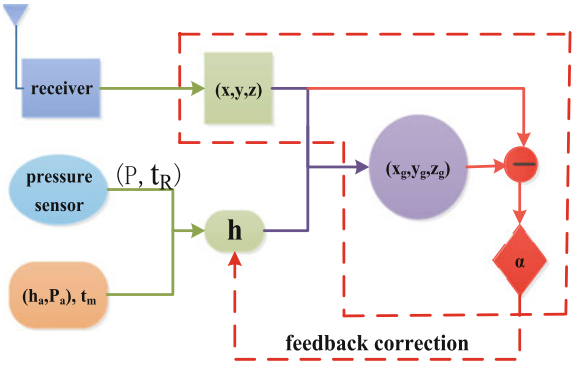


Fig. 4 Flow chart of satellite positioning assisted by differential barometric altimetry

Table 3 Correction coefficient table

Deviation ΔL/m	<1.0	1.0–1.5	1.5–2.0	2.0–2.5	2.5–3.0	>3.0
Correction coefficient α	0.5	0.6	0.7	0.8	0.9	1

2. Obtain the average height h_a by the location (b_i, l_i, h_i) which the location (x_i, y_i, z_i) converts into and the average temperature t_m during this time.
3. The pressure value P_a corresponding to the average height h_a is calculated by the formula (7). Use them as the initial altitude—pressure reference point at the next average height h_a.
4. In the next period of time, through the real-time pressure P, temperature t_R and the reference point (h_a, P_a) calculated over the previous period, get real-time pressure height h_p according to formula (1), where tm = (t_a + t_R)/2.
5. The correction coefficient of the real-time altitude h is determined by the deviation between the positioning aided by altitude and the positioning solved by the receiver on the last moment. The deviation ΔL is calculated as shown in Eq. 8 and altitude calculation is shown in Eq. 9. The coefficient is determined by Table 3. When the deviation is more than 2 m, set the current time corrected h and its corresponding P as the initial altitude—pressure reference point in the next period time.
6. Combine the real-time receiver position with real-time altitude h, position (x_g, y_g, z_g) aided by altitude is calculated by Eq. (6).

$$\triangle L = \sqrt{(x_{g_{i-1}} - x_{i-1})^2 + (y_{g_{i-1}} - y_{i-1})^2 + (z_{g_{i-1}} - z_{i-1})^2} \tag{8}$$

(x_{gi-1}, y_{gi-1}, z_{gi-1}) is the position aided by altitude at the previous moment, and (x_{i-1}, y_{i-1}, z_{i-1}) is the position solved by receivers at the last moment.

$$h = \alpha h_p + (1 - \alpha) h_i \tag{9}$$

h_p represents the improved altitude value at the current moment, h_i represents the altitude value solved by receiver and α is the correction coefficient.

4.3 Data Results and Analysis

Select Academy of OPTP-Electronics, Chinese Academy of Science (AOE) as the experimental site. We use Novatel GPS receivers to do long-time single-point positioning to get the exact coordinates. Choose any 4 satellites to simulate the "half the sky" bad scene.

Figure 5 shows that in "half sky" scene, under the condition that position dilution of precision (PDOP) is 6.7261, the error of the position calculated by a

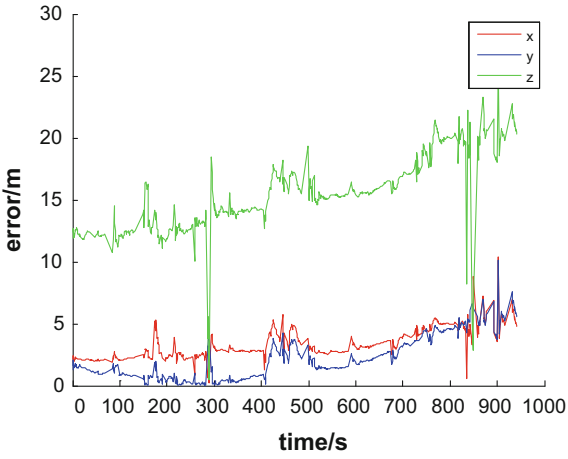


Fig. 5 Error of the position calculated by a receiver deviating from the real position

receiver deviates from the real position. The experiment time is 15 min. Figure 6 shows the error of the position calculated by the method of satellite positioning assisted by improved differential barometric altimetry deviating from the real position. Calculate the average altitude h_a every 1 min, corresponding mean error, variance as shown in Table 4. Figure 7 shows the error of the altitude calculated by the improved method deviating from the real receiver altitude.

As shown in the above figure and table, differential barometric altimetry can be performed without a barometric reference station through an improved method and it can assist satellite positioning to improve positioning accuracy.

The time is generally 1–5 min, when a receiver crosses tall buildings under the condition of 3 visible satellites. Use the improved method without feedback correction to solve three satellites positioning in short time. Select AOE as the

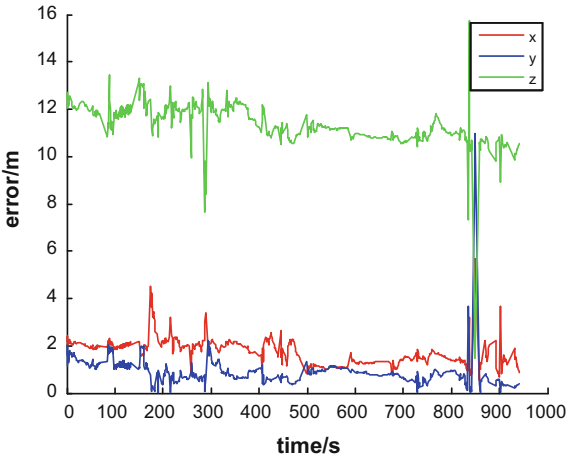


Fig. 6 Error of the position calculated by the method of satellite positioning assisted by improved differential barometric altimetry deviating from the real position

Table 4 Mean error and variance of the position result of the two solutions

	The positioning error solved by receivers		Positioning error assisted by improved differential barometric altimetry	
	Mean error/m	Variance	Mean error/m	Variance
x	1.7755	1.2971	1.1072	0.2680
y	6.2607	2.8000	5.2998	0.4123
z	6.5356	5.5408	4.6911	0.7574

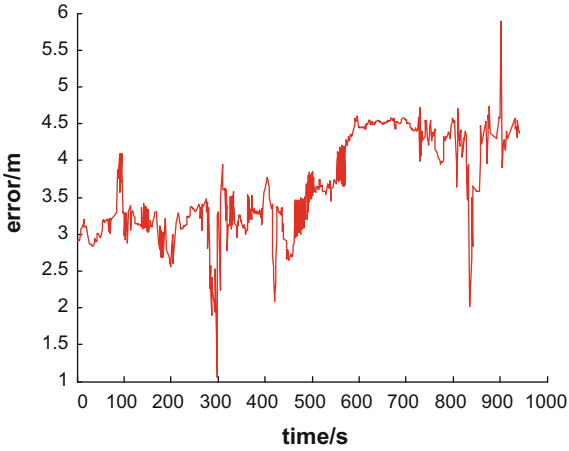


Fig. 7 Error of the altitude calculated by improved differential barometric altimetry deviating from the real receiver altitude

experimental site. When the number of visible satellites is more, we set altitude-pressure reference points. When the number of visible satellites is three, we use the (h_a, P_a) in the recent period time as the altitude-pressure reference point for differential barometric altimetry. Figure 8 shows the error of three satellite positioning assisted by improved differential barometric altimetry.

From Fig. 9 we know that the improved differential barometric altimetry method can be used to assist the three satellite positioning.

5 Satellite Integrity Monitoring Assisted by Differential Barometric Altimetry

5.1 Principle of Satellite Integrity Monitoring Assisted by Altitude

One of the most commonly used in traditional satellite integrity monitoring algorithms (RAIM) is the pseudo range residual detection method. At least 5 satellites

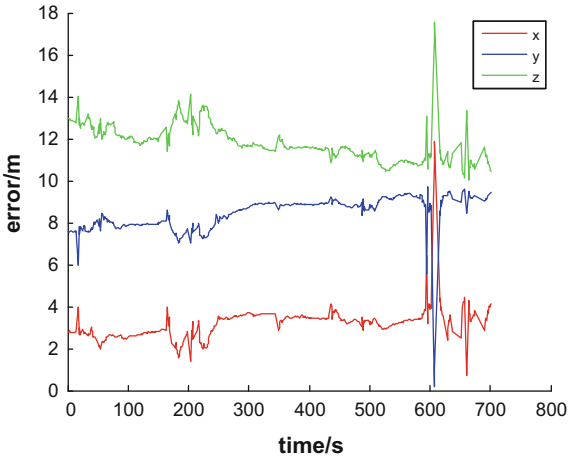


Fig. 8 Error of three satellite positioning assisted by improved differential barometric altimetry

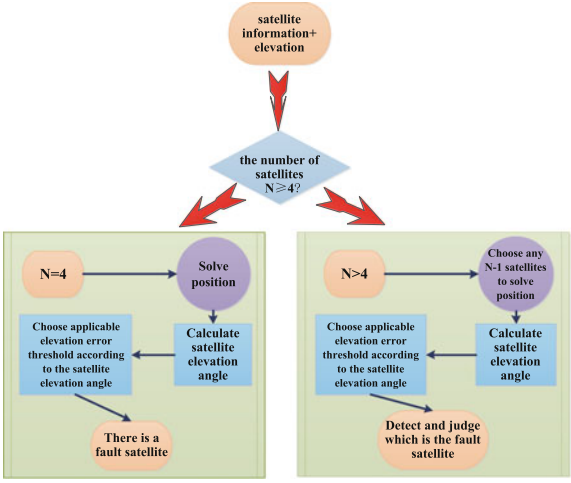


Fig. 9 Flow chart of RAIM algorithm assisted by differential barometric altimetry

are required for fault detection and the number of satellites is at least 6 to troubleshoot. In this article, due to the addition of known altitude information, the fault of satellites can be detected when there are 4 satellites and of single-satellite can be excluded when there are 5 satellites. The process of RAIM algorithm aided by altitude is shown as Fig. 9:

Specific steps are as follows:

1. After getting the satellite position and differential barometric altimetry, firstly determine the number of satellites. Only when the number of satellites is more than or equal to 4, integrity monitoring can be performed.
2. If the number of satellites is only four, first we will solve the position to get a set of altitude positioning. If the environmental conditions are good, we will get N

satellites and then we will use any N-1 satellites for solving position, N positioning results will be obtained.
3. Because satellite integrity monitoring aided by altitude is not sensitive to low-elevation satellites, firstly determine the elevation angle of the satellite and secondly select the appropriate altitude error threshold Δh within a range of satellite elevation.
4. The deviation of all the obtained altitude positioning results from the measured differential pressure altitude is compared with the threshold value Δh. When the number of satellites is four, if the deviation exceeds Δh, it will be determined that there is a satellite fault. When the number of satellite is more than four, if N-1 groups of altitude deviation exceed Δh, it will be determined that there is a single-satellite fault.
5. If the fault of satellite is detected, then the non-faulty satellite will be the satellites whose altitude deviations are within the range of Δh.

The RAIM algorithm aided by altitude lies in the consistency of satellite measurements. If there is a satellite fault, pseudo-range residual will be inconsistent so that the positioning deviation will be greater. This is the basic principle of satellite integrity monitoring aided by altitude.

5.2 *Experimental Results and Data Analysis*

Experimental site is still selected in AOE and use Novatel GPS receiver for a long time single point positioning to get the exact coordinates. Choose one satellite from N satellites every time, adding a certain deviation to its corresponding pseudo range. Altitude error deviation threshold Δh was selected as 3 m. Fault detection rate is shown in Tables 5, 6, 7 and 8:

1. Due to the satellite integrity monitoring aided by altitude requiring a higher precision, the fault detection rate has a high dependency on the satellite elevation. When the elevation is greater than 30°,fault detection rate has been improved comparing with the elevation greater than 10°.

Table 5 The fault detection rate of satellite elevation greater than 10° and adding 50 m deviation to the range

Adding 50 m deviation, satellite elevation greater than 10°, Δh = 3 m				
The number of satellites	5	6	7	8
Fault detection rate (%)	40	55	67	80

Table 6 The fault detection rate of satellite elevation greater than 10° and adding 100 m deviation to the range

Adding 100 m deviation, satellite elevation greater than 10°, Δh = 3 m				
The number of satellites	5	6	7	8
Fault detection rate (%)	55	63	72	83

Table 7 The fault detection rate of satellite elevation greater than 30° and adding 50 m deviation to the range

Adding 50 m deviation,satellite elevation greater than 30°, Δh = 3 m				
The number of satellites	5	6	7	8
Fault detection rate (%)	60	70	76	85

Table 8 The fault detection rate of satellite elevation greater than 30° and adding 100 m deviation to the range

Adding 100 m deviation,satellite elevation greater than 30°, Δh = 3 m				
The number of satellites	5	6	7	8
Fault detection rate (%)	70	75	80	87

2. When there is a larger pseudo range deviation, the detection rate is greater than the detection rate with smaller pseudo-range deviation. This is because when the observed noise is similar to the pseudo-range deviation, RAIM algorithm is difficult to distinguish the faulty satellite. Therefore, the RAIM algorithm is more effective for the detection of faulty satellites with large pseudo range deviation.
3. As the number of satellites increases, the altitude will be more accurate. This also increases the detection rate of the RAIM algorithm with altitude assistance.

Acknowledgements Project Y70B13A1BY supported by The Innovation Program of Academy of Opto-Electronics (AOE), Chinese Academy of Science (CAS).

References

1. Sun J (2008) Combination navigation and positioning technology of GPS and high-precision air pressure altimeter. D Nanjing University of Aeronautics and Astronautics, Nanjing
2. Hu Z, Zhang L (2012) Performance analysis of differential barometric altimetry in indoor positioning systems. J Chin J Sens Actuators 25(10):1463–1467
3. Wang X (2015) Design and realization of barometric altimeter in navigation receiver. D Chinese Academy of Sciences University, Beijing
4. Du X, Pei J, Zhang L, Liu C (2013) Differential pressure altimetry method based on ground mobile communication base station. J J Beijing Univ Aeronaut Astronaut 39(1):83–88
5. Hu Z (2014) Study on key techniques of navigation system assisted by differential pressure altimetry and other sensing information. D Chinese Academy of Sciences University, Beijing
6. Hu Y, Cao K, Li B, Zhu Y (2013) Research on RAIM algorithm with additional altitude information. J Comput Simul 30(12):61–64
7. Gao F (2015) Research on Improvement of RAIM algorithm in harsh conditions. D'Xidian University, Xi'an
8. Wu P (2013) Research on positioning technology of satellite navigation receiver assisted by altitude and clock error. D National University of Defense Technology, Hunan

Simulation on the Qualitative Relationship Between Carrier Movement and Eddy Current Interfering Magnetic Field in Geomagnetic Navigation

Chi Li, Binfeng Yang, Junmin Xu and Huan Sun

Abstract The premise of using geomagnetic measurement to realize high precision navigation is that magnetometer can obtain geomagnetic field data accurately. However, the moving carriers, including missiles and aircraft, whose attitudes will continue to change during the flight and eddy current magnetic field generated by carrier moving also changes. Therefore, it is of great significance to study the intrinsic relationship between them to eliminate the influence of eddy current magnetic field. In this paper, COMSOL Multiphysics was used to establish the model of eddy current magnetic field, and it was studied that the influence of carrier on geomagnetic measurement under different motion conditions. Firstly, the influence of the characteristics of the permanent magnet on the uniform magnetic field is analyzed. Secondly, the distribution of the carrier's eddy current magnetic field under the rotation condition is studied. Finally, the relationship is considered between the eddy current field and the carrier speed under the condition of swing. The simulation results clarify the relationship between eddy magnetic field and the carrier movement mode. At the same time, it provides an important theoretical basis for the geomagnetic measurement error compensation algorithm.

Keywords Geomagnetic navigation · Ferromagnetic material · High speed motion Eddy magnetic field

1 Introduction

Geomagnetic navigation provides the of navigation service by matching the geomagnetic data measured by magnetometer with the previously prepared geomagnetic map [1, 2]. The earth's magnetic field is very stable and contains rich volume

C. Li (✉) · B. Yang · J. Xu · H. Sun
Information and Navigation College, Air Force Engineering University, Xi'an 710077, China
e-mail: Nacl_1208@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_68

of magnetic characteristics, such as the total intensity and magnetic field intensity gradient, inclination and declination. These characteristics and geographical position has a corresponding relation, so the geomagnetic navigation is a passive, no radiation, anti-jamming, autonomous navigation mode all day long time, and has wide application prospect in the military field [3–5].

One of the key technologies to realize geomagnetic navigation is to measure the geomagnetic field data accurately [6–8]. The geomagnetic navigation is mainly applied to Underwater Unmanned Underwater Vehicles, unmanned aircraft carrier is closely related to the and the military field [9], and the extensive use of ferromagnetic materials in the carrier, resulting inevitably carrier magnetic interference in the measurement process. The interference magnetic field has a great influence on the geomagnetic survey data. If we do not study the interference field and make error compensation, it is impossible to obtain accurate geomagnetic data matching [10, 11]. Therefore, it is very important to study the interference magnetic field of the carrier for geomagnetic measurement.

In geomagnetic measurement, the interference of the measurement mainly includes instrument error, hard iron error, soft iron error and eddy current magnetic field interference. The instrument error can be eliminated by calibrating the magnetometer in advance, and it has nothing to do with the carrier and the geomagnetic field. Hard and soft iron error, and eddy current magnetic field is related to the carrier, is unique in the geomagnetic measurement [12]. In the study of carrier magnetic interference, the scholars at home and abroad focus on the hard and soft iron error for land vehicles and underwater vehicles. For instance, the Hu Xiaoping studied the error compensation of unmanned underwater vehicles, the instrument error, hard and soft iron error and random error of integrated modeling, but did not consider the eddy current magnetic field interference [13–15]. When studying moving targets with low velocity in ground and underwater, the smaller eddy current magnetic field can be ignored. However, in high speed aircraft carrier, the effect of the eddy current magnetic field is not clear.

Therefore, this paper mainly studies the eddy current interference magnetic field characteristics of high speed moving carrier. The relationship between eddy magnetic field and velocity, carrier position and direction are simulated and analysed so as to guide the installation of the magnetometer and the compensation of the interference field of the carrier.

2 Mechanism of Carrier Interference Magnetic Field

Carrier interference magnetic field includes hard iron errors, soft iron errors and eddy current magnetic field. Hard iron errors are caused by the carrier of the remanence of the ferromagnetic material. Due to the hard magnetic materials with high coercivity, it is considered to be a constant value. Soft iron errors are produced by magnetization of soft magnetic material of carrier. Because soft magnetic material has low coercive force, when the external magnetic field changes, the soft

magnetic interference also changes, so the soft iron errors are closely related to the movement attitude of the carrier and the geomagnetic field [16].

The eddy current magnetic field is the magnetic field produced by the metal conductor cutting the magnetic field line when the navigation carrier moves in the geomagnetic field. If the influence of eddy current magnetic field is not made compensation, it will seriously affect the real-time output data of magnetometer, make the navigation error bigger, even not to provide accurate navigation and positioning for the aircraft.

3 Establishment of Simulation Model and Analysis of Uniform Magnetic Field Characteristics

The geomagnetic field is globally distributed and stable, so the geomagnetic vector can be regarded as a function varying with position. In the simulation model, the influence of the motion state of the carrier on the measurement of geomagnetic field is qualitatively studied. Here, the geomagnetic field can be considered as a locally uniform magnetic field [17].

3-D simulation model of cylindrical carrier moving in a uniform magnetic field is established, as shown in Figs. 1 and 2. Uniform magnetic field generated by a permanent magnet is used to simulate the geomagnetic field in the model.

The outer layer of the model is a cylindrical air field, and the model is confined inside air domain, which is the condition for finite element analysis. The U type structure in the middle is a device to produce uniform magnetic field, inside which there are two cylinders. The outer cylinder is a rotating domain established, by which the simulation model is divided into two parts. The outside of the cylinder is

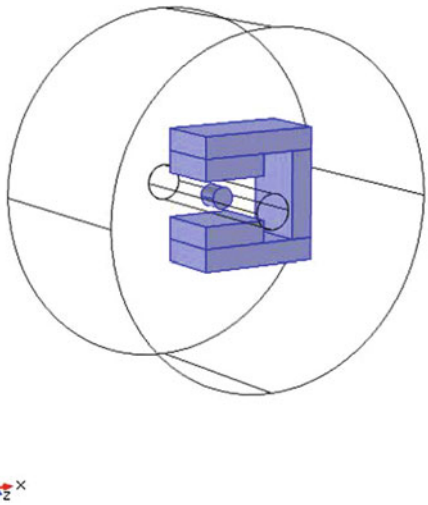


Fig. 1 3D diagram of the model

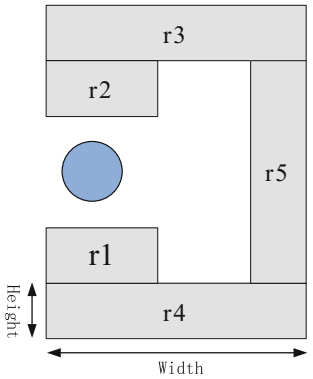


Fig. 2 2D diagram of the model

a stationary domain and the inside is a rotating domain. Inner cylinder is solid copper carrier that will move in the rotating domain. The initial size of the model is shown in Table 1.

The uniform magnetic field is generated by U type structure, in which r5 is a permanent magnet, and the initial remanence is 50,000 nT. r1, r2, r3, r4 are yokes, and their function is to guide the propagation of the magnetic field in the magnetic circuit.

In order to verify the permanent magnet at the center of the magnetic field, in the static state of the carrier for steady state analysis of magnetic field distribution between the magnetic yoke, the results are shown in Fig. 3, it can be seen from the figure, the uniform magnetic field distribution between the magnetic yoke, and stronger than the surrounding magnetic field, permanent magnet that has produced the magnetic field.

In order to further verify the quantitative relationship between the remanence of permanent magnet and uniform field strength, the distribution curves of the magnetic field on the path parallel to the axis of the carrier inside the yoke are obtained under different external magnetic fields, and the results are shown in Fig. 4. It can be seen that the magnetic field is uniformly distributed along the length of the yoke and gradually decreases at both ends. When the remanence of the permanent magnet is 40,000, 50,000 and 60,000 nT, the magnetic induction intensity becomes 2858, 3572 and 4287 nT on the path respectively. The ratio of the uniform field intensity is similar to that of the remanence of the permanent magnet, which both

Table 1 Model size

	Air domain	Rotating domain	Carrier	r1	r2	r3	r4	r5
Radius (m)	0.08	0.008	0.005	/	/	/	/	/
Height (m)	0.09	0.09	0.01	0.01	0.01	0.01	0.01	0.04
Width (m)	/	/	/	0.03	0.03	0.055	0.055	0.01
Length (m)	/	/	/	0.03	0.03	0.03	0.03	0.03

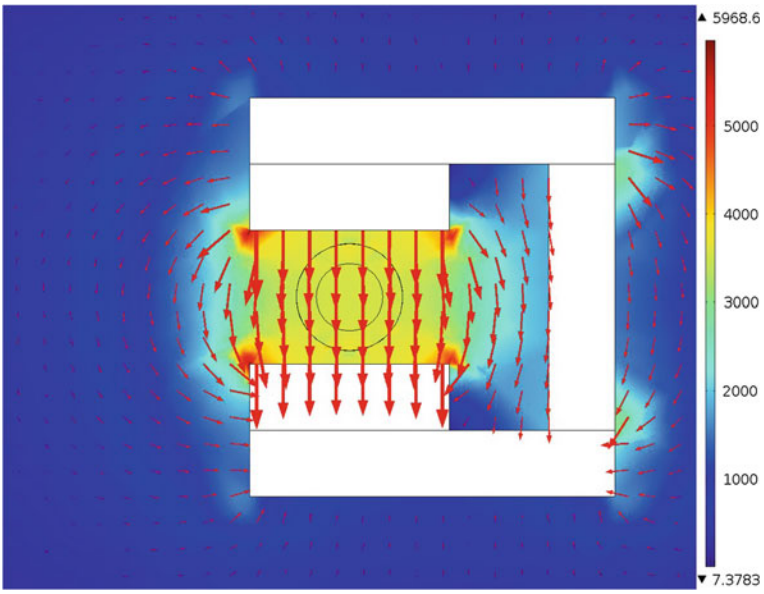


Fig. 3 Simulation on uniform field

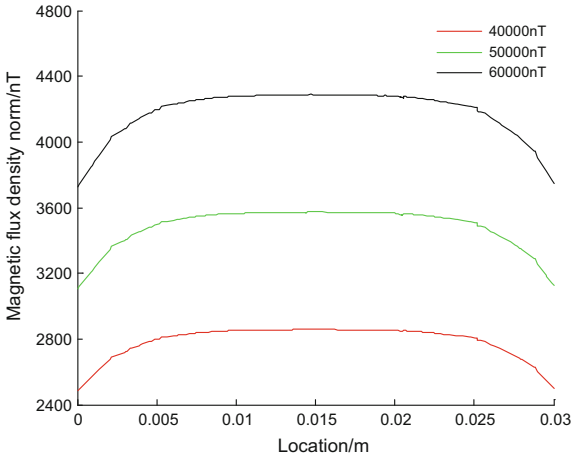


Fig. 4 Uniform field produced by permanent magnets of different strength

are 4:5:6. Therefore, when the material and structure are unchanged, uniform field intensity is proportional to the remanence of permanent magnet. By changing the remanence of the permanent magnet, we can produce different intensity of magnetic field.

4 Influence of Carrier Motion on Eddy Magnetic Field

When the model can produce a uniform magnetic field, the carrier is placed in the uniform magnetic field to study the motion characteristics. The motion of the carrier mainly includes two ways of rotation and swing, so the eddy current magnetic field is discussed under these two modes of motion.

4.1 Carrier Rotates at Uniform Speed

The external uniform magnetic field moves along the −Y axis, and the carrier rotates uniformly around the Z axis, as shown in Fig. 5.

When rotating at constant speed, the velocity remains constant in one cycle. When the rotational speed is 2 Hz, the current distribution on the carrier surface is observed, as shown in Fig. 6, and it can be seen from the figure that the eddy exists on the surface of the carrier. In one cycle, the eddy flows along the plane parallel to the YOZ in clockwise direction, and the direction remains unchanged.

In order to qualitatively analyze the relationship between the rotational speed of the carrier and the eddy intensity, the speed is set as 2 and 40 Hz respectively, and the transient analysis of the model is carried out. The distribution of the eddy intensity at different speeds is obtained. The results are shown in Figs. 7 and 8. It can be seen from the diagram that the eddy intensity remains unchanged at different times. However, at the same moment, the eddy in the middle of the carrier is larger, the two sides smaller. When the speed increases from 2 to 40 Hz, the eddy intensity has obvious enhancement. The reason is that when the carrier speed increases, the speed of the carrier cutting the magnetic line will also speed up. According to Faraday's law of electromagnetic induction, the induced EMF will be increased, leading to the induced eddy enhancement.

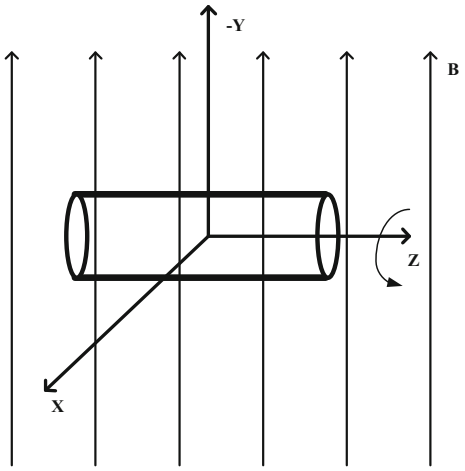


Fig. 5 Uniform rotation of carrier

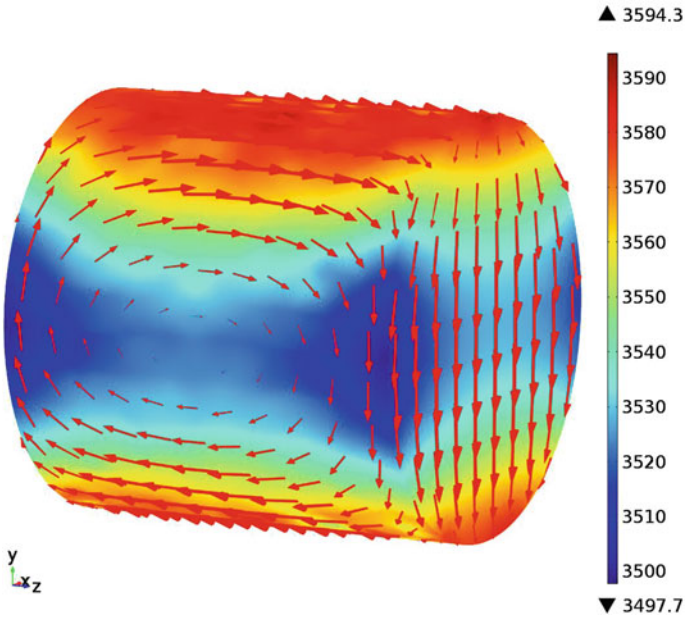
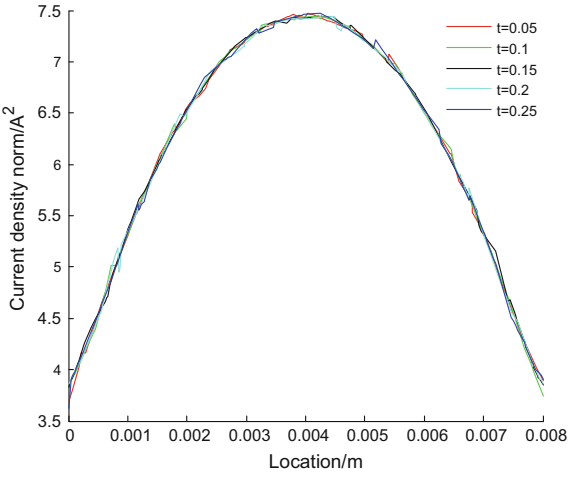


Fig. 6 Eddy current distribution on the surface of carriers

Fig. 7 Eddy current strength when ω = 2 Hz



Based on the analysis of the relationship between the eddy current and the rotating speed of the carrier, it is necessary to further clarify the relationship between the magnetic field strength and the rotational speed of the carrier. The simulation results show that when the carrier speed is small, the eddy current is very weak, so the eddy current magnetic field strength is relatively small. Thus it is not

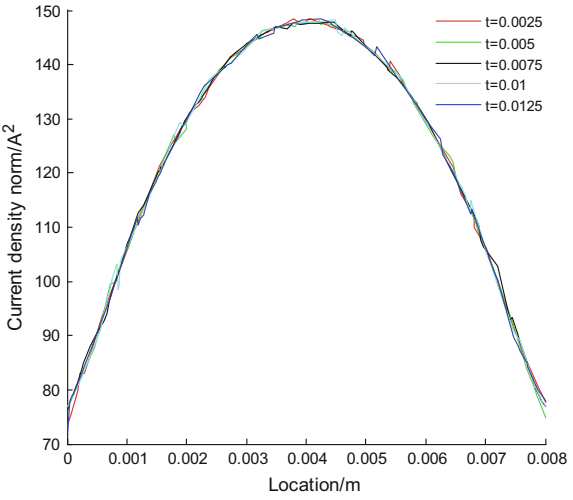


Fig. 8 Eddy current strength when ω = 40 Hz

intuitive to study the relationship between the eddy magnetic field and the speed directly. The theoretical analysis shows that the total magnetic field around the carrier is the sum of the external magnetic field and the eddy current field, and the additional uniform field strength remains unchanged in the simulation. Therefore, the influence of the rotating speed on the eddy current magnetic field can be analyzed by studying the variation of the total magnetic field around the carrier. The 5 sets of rotational frequencies in the simulation are 2, 10, 20, 30 and 40 Hz respectively. The total magnetic field intensity curve on the axial path of the carrier is obtained by simulation, as shown in Fig. 9.

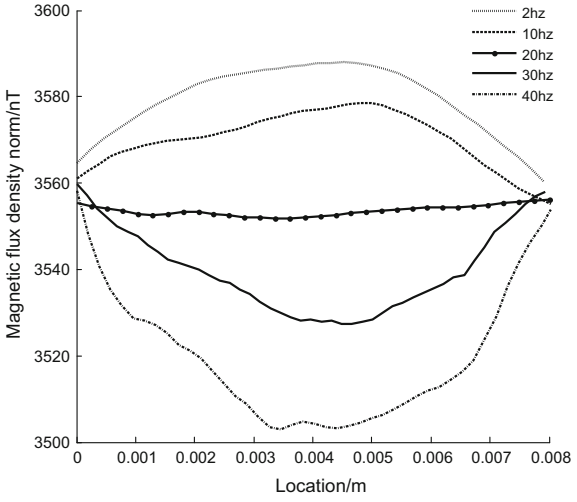


Fig. 9 Total magnetic induction intensity

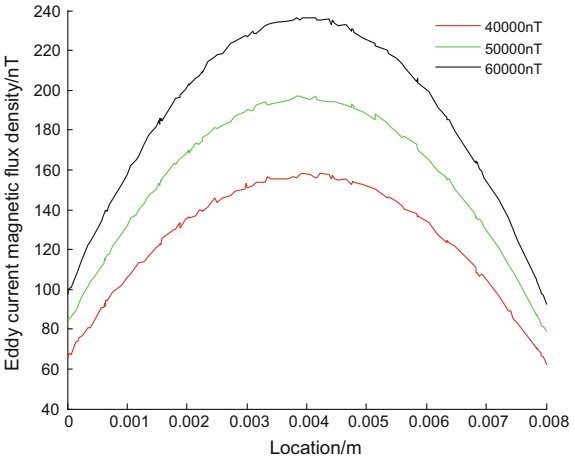


Fig. 10 Eddy field induced by permanent magnets of different strength

It can be seen from the Fig. 9 that with the increase of the speed, the total magnetic field intensity decreases obviously. According to Lenz's law, The magnetic field generated by the induced current always impede the change of the flux caused by the induced current. That is to say, the eddy current magnetic field hinders the change of magnetic flux. When the speed is small, the eddy current magnetic field is weak, and the obstruction is not obvious. But when the velocity increases, the eddy current field becomes stronger and stronger, and the blocking effect becomes stronger and stronger, which leads to the decrease of the total magnetic induction intensity of the carrier.

In order to study the influence of the external magnetic field on the eddy magnetic field, the rotational speed of the carrier is set as 100 Hz, and the remanence of the permanent magnet is 40,000, 50,000 and 60,000 nT respectively. The variation curve of the eddy current field along the axial direction is obtained, as shown in Fig. 10. It can be seen from the figure, the strength of eddy magnetic field were 156, 195, 234 nT, which is approximately equal to 4:5:6. It is the same as the ratio of the remanence of the permanent magnet. The above simulation analysis has shown that the uniform magnetic field is proportional to the remanence strength of the permanent magnet. Therefore, there is a linear relationship between the eddy magnetic field and the strength of the external uniform magnetic field.

4.2 Carrier Swings

In the state of swing, the velocity and direction of the carrier changes in one cycle. In the simulation, the swing curve is set as sine function. The carrier starts from the balance position and swings around the Z axis, as shown in Fig. 11. The swing

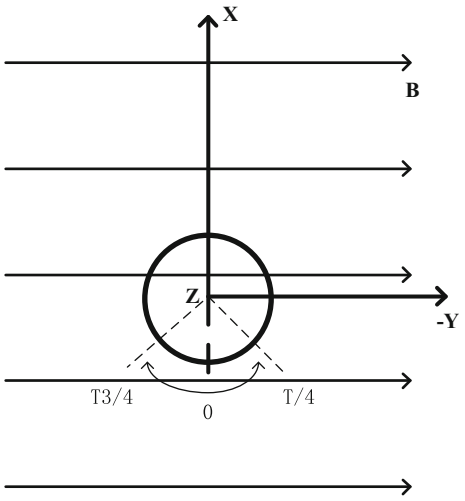


Fig. 11 Carrier oscillates

angle frequency is 628 rad/s, and the period is 0.01 s. The amplitude is $\pi^2/30$, and the initial phase is 0°. The direction of the uniform magnetic field is along the −Y axis.

According to the right-hand rule, when the carrier moves forward along the Y axis, the eddy is always distributed on the plane parallel to the YOZ, and the eddy current moves along the carrier surface clockwise. When along the −Y axis, it flows along the surface of the carrier in an anticlockwise manner. In order to verify the correctness of the model, transient analysis of model is calculated to observe the eddy current distribution of carrier in t = 0.002 s and t = 0.005 s, and the results are shown in Figs. 12 and 13. The size of arrows indicates eddy intensity. We can obtain two conclusions from the figure. Firstly, the flow direction of the induced eddy is closely related to the direction of the carrier wobble. The carrier swing along the Y axis when t = 0.002 s and along the −Y axis when t = 0.005 s, so the induction eddy current flow is reversed, which is consistent with the theory analysis. It also verified the correctness of the simulation model. Secondly, the intensity of the induced eddy current is closely related to the velocity of the carrier. When t = 0.002 s, carrier is about to swing to the maximum amplitude, so the speed is significantly less than the speed when t = 0.005 s. Because the carrier just swing to the equilibrium position when t = 0.005 s, where speed is the maximum value. It can be seen from the figure, the induced eddy current in the equilibrium position is obviously stronger than that in other positions, which is consistent with the conclusion from the rotating carrier.

The eddy current intensity is qualitatively analyzed in a period, and the distribution curve of the eddy intensity at different times in the cycle is obtained, as shown in Fig. 14. It can be seen from the figure that the maximum value of the eddy appears at the t = 0.005 and t = 0.01. The carrier swing just after the equilibrium position, and the velocity is maximum. So the eddy current is the strongest. In other

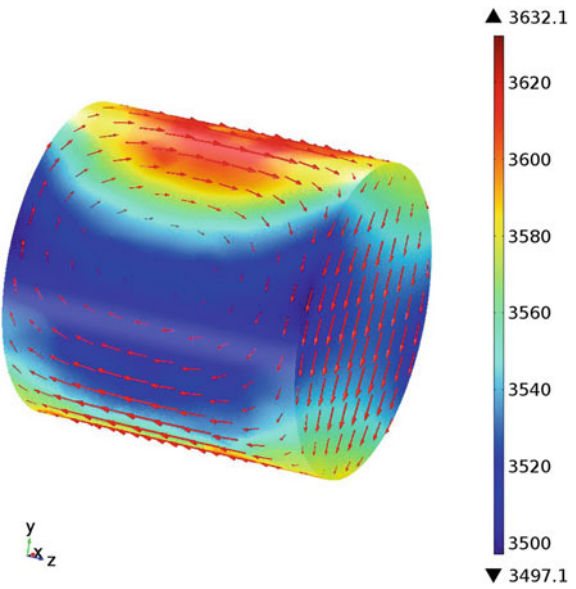


Fig. 12 Eddy distribution when t = 0.002 s

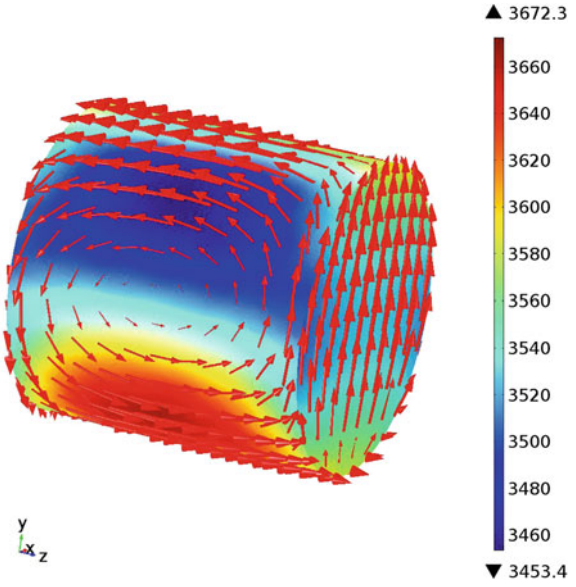


Fig. 13 Eddy distribution when t = 0.005 s

locations, with the decrease of the swing speed of the carrier, the induction eddy current is reduced gradually.

In order to further verify the relationship between the eddy current magnetic field and the velocity of the carrier, total magnetic field intensity curves at 3 different

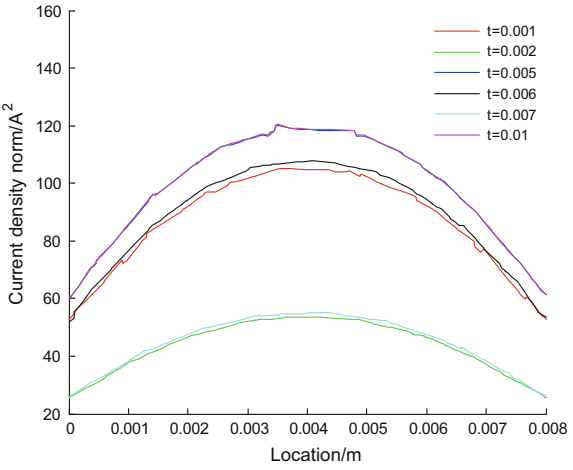


Fig. 14 Eddy intensity at different times

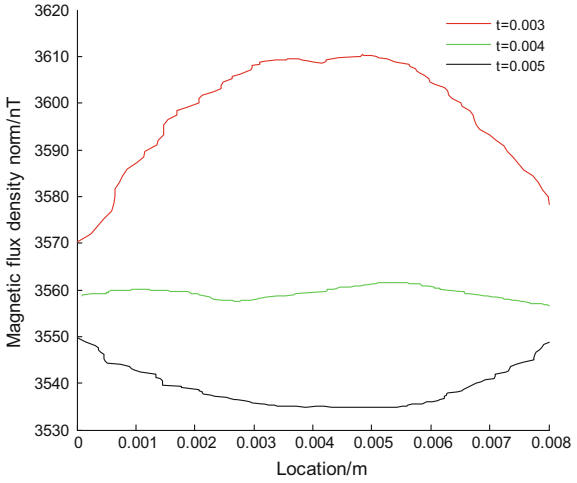


Fig. 15 Total magnetic induction intensity at different moments

times in the 1/4 period of oscillating are simulated, which is shown in Fig. 15. It can be seen from the figure that its position moves from near the maximum swing to the equilibrium position when the carrier swings from t = 0.003 s to t = 0.005 s, so the swing speed will gradually increase. The magnetic field of the induced eddy will gradually increase with the increase of the velocity of the carrier, which leads to the decrease of the total magnetic field strength. This is consistent with the conclusions drawn from previous simulation.

In summary, the following conclusions can be drawn: the eddy current magnetic field increases with the increase of the carrier velocity, and the distribution of the magnetic field presents a strong middle and weak side distribution, while it is proportional to the magnitude of the external magnetic field. The research shows

that the eddy current magnetic field cannot be ignored when the carrier is moving at high speed. In order to achieve high precision geomagnetic navigation, the eddy current magnetic field must be eliminated and corrected. Meanwhile, the installation position of the magnetometer must be optimized. It should be installed on both sides to reduce its impact on the eddy current magnetic field as far as possible.

5 Conclusions

Aiming at the problem of eddy magnetic field in the carrier interference field, the simulation model of eddy current interference field is built by using COMSOL Multiphysics simulation software. First, the characteristics of the uniform field of permanent magnets used to simulate geomagnetic field are studied, and then the distribution of the eddy current magnetic field of the carrier under uniform rotation and swing is simulated. The influence of the eddy current interference field on the geomagnetic measurement is clarified. The simulation results show that the eddy current magnetic field increases with the increase of the carrier velocity, and the eddy current magnetic field mainly exists in the middle of the carrier. The simulation results provide theoretical basis for the error modeling and the magnetic field correction technology in geomagnetic navigation.

References

1. Li X, Zhang H, Zhang S (2013) Development status of geomagnetic navigation technology. Aerodyn Missile J 10:80–83
2. Chen Z, Tian F, Sun J (2016) Review of key technologies in geomagnetic navigation. Geomat Spat Inf Technol 9(1):16–18
3. Guo C, Hu Z, Zhang S et al (2009) A survey of geomagnetic navigation. J Astronaut 30 (4):1314–1319
4. Zhou J, Ge Z, Shi G et al (2008) Development and key technologies of geomagnetic navigation. J Astronaut 29(5):1467–1472
5. Wang LJ (2014) Magnetic sensors for navigation applications: an overview. J Navig 67 (2):263–275
6. Ge Z, Liu S, Li G et al (2017) Error model of geomagnetic-field measurement and extended Kalman-filter based compensation method. PLoS One 12(4):e0173962
7. Wei JF (2017) Key technology status of geomagnetic matching navigation. Electron Qual
8. Sun XJ, Kou J, Zhang XN et al (2016) Research progress in geomagnetic navigation. Navig Control
9. Yu ZT, Junwei L, Ji SK (2014) A compensation method for the vehicle 3-D magnetic field based on ellipsoid constraint. J Harbin Eng Univ 35(6):731–734
10. Guo C, Li A, Hong C et al (2011) Algorithm for geomagnetic navigation and its validity evaluation. In: IEEE international conference on computer science and automation engineering. IEEE, pp 573–577
11. Wang Y, Zhang JY, Zhang DW (2015) Error analysis and algorithm validation of geomagnetic sensor. Appl Mech Mater 742(2):21–26

12. Li J (2013) Study on the characteristics and compensation method of vehicle interferential magnetic field in geomagnetic measurement. National University of Defense Technology
13. Wu Z (2013) Research on geomagnetic measurement error compensation for underwater geomagnetic navigation. National University of Defense Technology
14. Zhang Q, Li J, Chen DX et al (2014) Method and experiment for compensating the interferential magnetic field in underwater vehicle. Measurement 47(1):651–657
15. Wang P (2014) Study on the suitability of underwater geomagnetic navigation. National University of Defense Technology
16. Gebre-Egziabher D, Elkaim GH, Powell JD et al (2001) A non-linear, two-step estimation algorithm for calibrating solid-state strapdown magnetometers. In: International St Petersburg conference on navigation systems, pp 28–30
17. Lv Z, He H, Zhang J et al (2014) Magnetic simulation device for simulating geomagnetic field. Electron Opt Control 21(6):76–80

Further Analysis on Cramer-Rao Bound for Cooperative Positioning Network

Siming Li, Jing Lv, Shiwei Tian, Longlong Suo and Yaning Liu

Abstract Cooperative positioning is expected to be a newly canonical technology to achieve highly accurate location awareness. In this paper, we present a further analysis for Cramer-Rao Bound (CRB) from the perspective of quantity of nodes, network topology and ranging error distribution. We propose and certify a lower limit of CRB and succinct topology design rules for cooperative network. Compared to the case with Gauss noise, the CRB of non-Gaussian ranging error model is developed and calculated. Ultimately, by simulation results, our conclusion yields important insights into the location system design fulfilling certain accuracy requirements.

Keywords Cooperative positioning · Ranging error · CRB

1 Introduction

Cooperative positioning is an emerging paradigm that circumvents the needs for high-power, high-density anchor deployment, and offers additional localization accuracy by enabling the agents to help each other in estimating their positions [1, 2]. The availability of such relative positional information is of great importance in many fields, such as parking garage, package distribution and security rescue. However, precise determination of agent positions is a challenging task, especially when it is influenced by topology and ranging error model.

Previous work introduces the CRB in terms of localization accuracy. It is a classical result of mathematical statistics which basically gives the lower bound on the covariance of all statistical estimators for a parameter. The bound is given in

S. Li · J. Lv (✉) · S. Tian · L. Suo · Y. Liu
Institute of Communication Engineering, Army Engineering University of PLA, Nanjing, China
e-mail: 13305147766@189.cn

S. Li
e-mail: siming.li.cn@outlook.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_69

terms of the Fisher information (FI) or Equivalent FI [3]. A general expression for the CRB of wireless cooperative networks has been derived in the literature [3–5]. Also, the upper limit of the CRB is confirmed to infinite large [6, 7]. Hybrid schemes using more than one type of position-determining methods are proved to reduce the CRB and improve position accuracy [6]. However, this result is calculated based on Gaussian ranging error. The experiments show the ranging error may not obey Gaussian distribution [8, 9].

Compared with the above existing work, the main contributions of this paper are summarized as follows:

- A closed-form expression of lower limit for CRB is derived. When the error probability distribution function (pdf) is very informative, the lower limit can be calculated.
- According to aforementioned lower bound, several nodes configuration rules for 2-dimensional cooperative network are proposed further.
- We numerically analyze CRB using non-Gaussian ranging error models and compare with Gaussian case. The impact of these models on CRB is verified.

The rest of this paper is organized as follows. Section 2 describes a cooperative positioning system and Fisher Information Matrix (FIM). In Sect. 3 we derive the lower limit of CRB, propose the several nodes distribution rules, and compare CRB in Gaussian and non-Gaussian ranging error. We show numerical simulation in Sect. 4, and complete this paper by some conclusions in Sect. 5.

2 System Setup and FIM

Consider a $N+1$ nodes in two-dimensional network, of which $N(N \geq 2)$ cooperative nodes with known position as $P_i = [x_i, y_i]^{\mathrm{T}}$, $i = 0, \ldots, N-1$ and one agent with unknown position as $P = [x, y]^{\mathrm{T}}$. ρ_i is a peer-to-peer range measurement between the ith node and the unknown agent.

$$\rho_i = \|P_i - P\| + \varepsilon_i \tag{1}$$

where

$$d_i = \|P_i - P\| = \sqrt{(x_i - x)^2 + (y_i - y)^2} \tag{2}$$

is the physically direct distance between the transmitter and the receiver and ε_i is the distance ranging (estimation) error variable with mean value μ_ε and variance σ_ε^2. Note that both ranging error are relative with measurement environment but independent of numerical order of the nodes. Thus, (1) has

$$\rho_i = d_i + \varepsilon \tag{3}$$

and let $f(\varepsilon)$ be the pdf of ε. Focus on one single agent $P = [x, y]^{\mathrm{T}}$, the PDF of the ith range measurement ρ_i can be expressed as

$$f(\rho_i, d_i) = f(\rho_i - \|P_i - P\|) = f(\rho_i - d_i) \tag{4}$$

As the range measurements which occur in different nodes are mutual independent, the joint PDF of the N measurements should be

$$p(\boldsymbol{\rho}; \mathbf{P}) = \prod_{i=0}^{N-1} f(\rho_i - d_i) \tag{5}$$

where $\boldsymbol{\rho} = [\rho_0, \ldots, \rho_{N-1}]^{\mathrm{T}}$ and $\mathbf{P} = [P_0, \ldots, P_{N-1}]^{\mathrm{T}}$ are measurement vector and position vector, respectively.

According to the notion of FIM, we have

$$\mathbf{I}(P) = \begin{bmatrix} E\left[\left(\frac{\partial \ln p(\boldsymbol{\rho};\mathbf{P})}{\partial x}\right)^2\right] & E\left[\frac{\partial \ln p(\boldsymbol{\rho};\mathbf{P})}{\partial x}\frac{\partial \ln p(\boldsymbol{\rho};\mathbf{P})}{\partial y}\right] \\ E\left[\frac{\partial \ln p(\boldsymbol{\rho};\mathbf{P})}{\partial y}\frac{\partial \ln p(\boldsymbol{\rho};\mathbf{P})}{\partial x}\right] & E\left[\left(\frac{\partial \ln p(\boldsymbol{\rho};\mathbf{P})}{\partial y}\right)^2\right] \end{bmatrix} \tag{6}$$

Then, [10] has made further efforts to facilitate $\mathbf{I}(P)$.

$$\mathbf{I}(P) = \sum_{i=0}^{N-1} E\left[\left(\frac{\partial \ln f(\rho_i - d_i)}{\partial d_i}\right)^2\right] \cdot \begin{bmatrix} \left(\frac{x-x_i}{d_i}\right)^2 & \frac{x-x_i}{d_i}\frac{y-y_i}{d_i} \\ \frac{x-x_i}{d_i}\frac{y-y_i}{d_i} & \left(\frac{y-y_i}{d_i}\right)^2 \end{bmatrix} \tag{7}$$

The right-side matrix has been nominated as ranging direction matrix (RDM) in [3].

$$\begin{aligned} \mathbf{R}(\phi_i) &= \begin{bmatrix} \left(\frac{x-x_i}{d_i}\right)^2 & \frac{x-x_i}{d_i}\frac{y-y_i}{d_i} \\ \frac{x-x_i}{d_i}\frac{y-y_i}{d_i} & \left(\frac{y-y_i}{d_i}\right)^2 \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\phi_i & \cos\phi_i \sin\phi_i \\ \cos\phi_i \sin\phi_i & \sin^2\phi_i \end{bmatrix} \end{aligned} \tag{8}$$

where

$$\phi_i = \tan^{-1}\frac{y - y_i}{x - x_i} \tag{9}$$

denotes the angle from the unknown agent to node i. Let the part corresponding to error distributions in (7) to

$$E\left[\left(\frac{\partial \ln f(\rho_i - d_i)}{\partial d_i}\right)^2\right] = g(f(\varepsilon)) \tag{10}$$

Then, (7) is formulated as

$$\mathbf{I}(P) = \sum_{i=0}^{N-1} g(f(\varepsilon)) \cdot \mathbf{R}(\phi_i) \tag{11}$$

Now, it is obvious that $\mathbf{I}(P)$ is relative to nodes number, RDM and ranging error distribution.

3 Derivation and Analysis of CRB

The Cramer-Rao inequality gives a lower bound for the error variance of any unbiased estimate of some unknown parameter [11]. Therefore, it is an appropriate reference for the location accuracy. In this section, we present the analysis of the location accuracy to illustrate the influence factors.

3.1 The Lower Limit of CRB

Noting the invertibility of FIM, we discuss CRB in 2 situations.

Case 1: If FIM is not inverse, $\det(\mathbf{I}(P)) = 0 \Rightarrow \sigma_{CRB} \to \infty$. That is

$$\forall i, \forall j > i \quad \phi_i = \phi_j \, or \, \phi_i - \phi_j = \pi \tag{12}$$

Case 2: If FIM is inverse, we have

$$\begin{aligned} \det(\mathbf{I}(P)) =& g^2(f(\varepsilon)) \cdot \left[\sum_{i=0}^{N-1} \cos^2 \phi_i \sum_{i=0}^{N-1} \sin^2 \phi_i - \left(\sum_{i=0}^{N-1} \cos \phi_i \sin \phi_i\right)^2\right] \\ =& g^2(f(\varepsilon)) \cdot \sum_{i=0}^{N-1} \sum_{j>i}^{N-1} \sin^2(\phi_i - \phi_j) \end{aligned} \tag{13}$$

and

$$\begin{aligned}\sigma_{CRB} &= \sqrt{\mathbf{I}_{11}^{-1}(P) + \mathbf{I}_{22}^{-1}(P)} \\ &= \sqrt{tr\left[\mathbf{I}^{-1}(P)\right]} \\ &= \sqrt{\frac{1}{\lambda_1} + \frac{1}{\lambda_2}}\end{aligned} \tag{14}$$

where $\mathbf{I}_{11}^{-1}(P)$ and $\mathbf{I}_{22}^{-1}(P)$ are elements in the 2×2 matrix $\mathbf{I}^{-1}(P)$, λ_1, λ_2 are eigenvalues of $\mathbf{I}(P)$ and $\frac{1}{\lambda_1}$, $\frac{1}{\lambda_2}$ are eigenvalues of inversion of $\mathbf{I}(P)$. Besides,

$$\begin{aligned}tr[\mathbf{I}(P)] &= g(f(\varepsilon)) \cdot N \\ &= \lambda_1 + \lambda_2\end{aligned} \tag{15}$$

$$\det(\mathbf{I}(P)) = \lambda_1 \cdot \lambda_2 \tag{16}$$

Hence, (16) can be induced to be positive to prove that λ_1 and λ_2 are greater than 0. Then,

$$\sigma_{CRB} = \sqrt{\frac{1}{\lambda_1} + \frac{1}{\lambda_2}} \geq \sqrt{\frac{2}{\sqrt{\lambda_1 \lambda_2}}} \tag{17}$$

If and only if $\frac{1}{\lambda_1} = \frac{1}{\lambda_2}$, i.e.

$$\lambda_1 = \lambda_2 = \frac{g(f(\varepsilon)) \cdot N}{2} \tag{18}$$

the equation will be established. Then, we get the minimum value of CRB.

$$\sigma_{CRB_{\min}} = \sqrt{\frac{2}{\sqrt{\lambda_1 \lambda_2}}} = \frac{2}{\sqrt{g(f(\varepsilon)) \cdot N}} \tag{19}$$

Equation (19) gives the optimal CRB theoretically for any topology structure and therefore localization algorithm design in practice is supposed to draw near $\sigma_{CRB_{\min}}$.

3.2 Topology Design Rules Based on CRB

Equation (11) has shown an underlying connection between CRB and ϕ_i, and such connection will be further demonstrated.

Rule 1: The worst nodes topology is linear distribution as Fig. 1a.

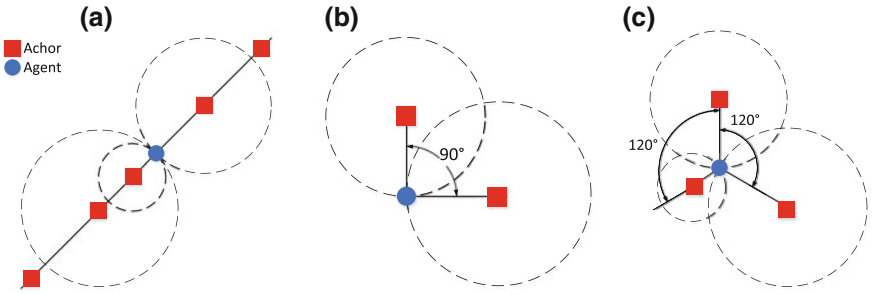


Fig. 1 Topology structure. **a** Example of Rule 1. **b** Example of Rule 2. **c** Example of Rule 3

It is intuitive that (13) sets a compact example when CRB arrives infinity.

Rule 2: The perfect case in 2 cooperative nodes is $\pi/2$ included angle, depicted in Fig. 1b.

We assume that 2 nodes cooperate to determine corresponding unknown coordinates. Combining (13) and (19), this can be written as

$$\sin^2(\phi_0 - \phi_1) = 1 \tag{20}$$

Rule 3: For $N \geq 3$ nodes, we can find a best case is uniform distribution Fig. 1c. In this condition, we obtain

$$\sum_{i=0}^{N-1} \sum_{j>i}^{N-1} \sin^2(\phi_i - \phi_j) = \frac{N^2}{4} \tag{21}$$

and we can find a solution of (21), given by

$$\forall i,\ \phi_i - \phi_{i+1} = \frac{2\pi}{N} \tag{22}$$

3.3 *Ranging Error Model*

It is well known that ranging error model typically hinges on cluttered environments, where multipath, LOS blockage, and excess propagation delays through materials lead to positively biased range measurements [12, 13]. Therefore, researching the positioning accuracy under different error models exhibits great necessity. Gaussian model [6], Exponential model and Asymmetric double exponential model (ADE) [8] used in this paper are listed in Table 1. The parameter λ_P, λ_N are positive and related to the measurement environment, which provide control over the shape of the model.

Table 1 Three ranging error model

Ranging error model	Probability distribution function	Mean	Variance
Gaussian model	$f_G(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma_G^2}} \exp\left(-\frac{(\varepsilon-\mu_G)^2}{2\sigma_G^2}\right)$	μ_G	σ_G^2
Exponential model	$f_E(\varepsilon) = \begin{cases} \lambda \exp(-\lambda\varepsilon), & \varepsilon > 0 \\ 0, & \varepsilon \leq 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
ADE	$f_{ADE}(\varepsilon) = \begin{cases} \frac{1}{\lambda_P + \lambda_N} \exp(-\frac{\varepsilon}{\lambda_P}), & \varepsilon \geq 0 \\ \frac{1}{\lambda_P + \lambda_N} \exp(\frac{\varepsilon}{\lambda_N}), & \varepsilon \leq 0 \end{cases}$	$\lambda_P - \lambda_N$	$\lambda_P^2 + \lambda_N^2$

Through (10), we have

$$g_G = \frac{1}{\sigma_G^2} \tag{23}$$

$$g_E = \lambda^2 \tag{24}$$

$$g_{ADE} = \frac{1}{\lambda_P \lambda_N} \tag{25}$$

It is observed that both Gaussian and Exponential case are inversely proportional to error variance. When assuming that the same mean value and variance, that is $\mu_G = \lambda_P - \lambda_N, \sigma_G = \lambda_P^2 + \lambda_N^2$, (23) can be simplified as

$$g_G = \frac{1}{\sigma_G^2} = \frac{1}{\lambda_P^2 + \lambda_N^2} \tag{26}$$

Since λ_P, λ_N are greater than 0 and $\lambda_P^2 + \lambda_N^2 \geq \lambda_P \lambda_N$ is always established, we have

$$g_G \leq g_{ADE} \tag{27}$$

Noting that the corresponding FIM $\mathbf{I}_G(P)$ and $\mathbf{I}_{ADE}(P)$ are determined by respective g_G and g_{ADE} for the one specific topology, we have

$$\mathbf{I}_G(P) \preceq \mathbf{I}_{ADE}(P) \tag{28}$$

Equation (28) means that $\mathbf{I}_{ADE}(P) - \mathbf{I}_G(P)$ is positive semidefinite [14]. Hence, (28) shows realistic Asymmetric double exponential model is superior to traditionally general Gaussian model in terms of positioning accuracy. Numerical simulation will be given in next section.

4 Numerical Example

The above results allow to compute $\mathbf{I}(P)$ for a given network configuration or error distribution and to express the σ_{CRB} by inverting itself.

Here, we adopt the experiments data from National Institute of Standards and Technology [15] and set $g(f(\varepsilon)) = 1$ for purpose of comparison within diverse nodes distributions. Figure 2 shows CRB values for several nodes distribution, where General 1 and General 2 are respectively set up two different topology structures, such as for 2 nodes they are $\pi/4$ and $2\pi/3$ included angle. It depicts the fundamental limits that CRB for any nodes structure is between $\sigma_{CRB_{\min}}$ and infinity. In addition, the optimal CRB for 2 nodes is $\pi/2$ and for $N \geq 3$ is uniform.

The Asymmetric double exponential model presented in [8] provides a realistic and flexible approach to statistically characterizing ranging errors experienced in typical indoor situation. The performance of it was simulated in Fig. 3, compared to

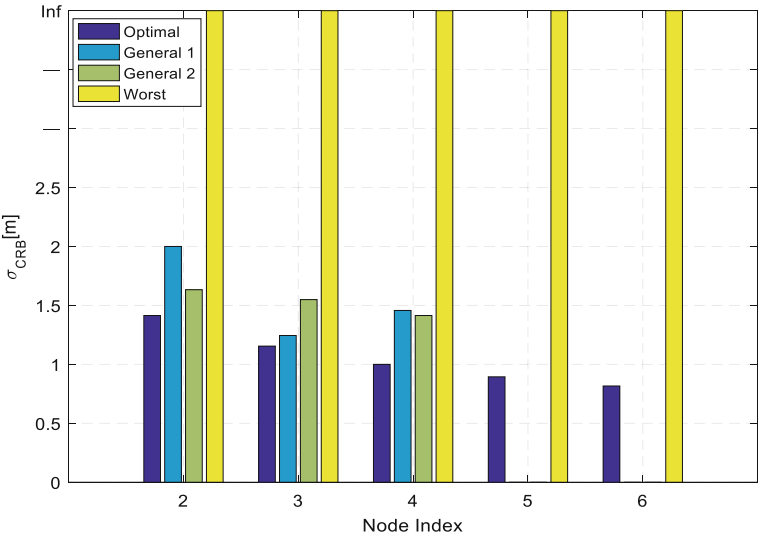


Fig. 2 σ_{CRB} with different topology structure

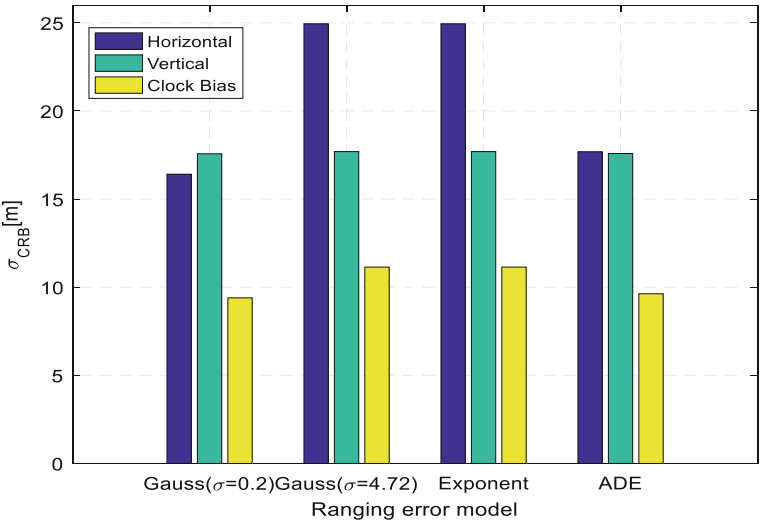


Fig. 3 σ_{CRB} with different ranging error model

Gaussian model. Firstly, we assume that a hybrid GNSS-terrestrial cooperative 3D positioning scenario[1] and the unknown position is the third agent, where cooperative ranging error is ideal Gauss with $\sigma = 0.2$. Therefore, CRB under $\sigma = 4.72$ [15] is a contrast in order to show practical positioning results. Secondly, the Gauss, Exponent and ADE case are compared in the same mean value and variance. It presents that ADE not only adapts to the real indoor environments but also gets the better position accuracy.

5 Conclusion

We have proposed a close expression of lower limit for CRB, which can be used as the optimal position accuracy and to guide the localization algorithm design. Based on analysis of position accuracy, several nodes distribution rules are derived to give insights into 2D localization awareness networks. Also, for agents indoors, Asymmetric double exponent distribution has a better performance than Gauss in terms of positioning accuracy.

Acknowledgements This work is supported by National Natural Science Foundation of China (61601511, 61571464, 61401507).

[1]The scenario is Fig. 3 in [6], and the related coordinates information also are given.

References

1. Shen Y, Wymeersch H, Win MZ (2010) Fundamental limits of wideband localization C part II: cooperative networks. IEEE Trans Inf Theory 56(10):4981C5000
2. Win MZ, Conti A, Mazuelas S, Shen Y, Gifford WM, Dardari D, Chiani M (2011) Network localization and navigation via cooperation. IEEE Commun Mag 49:56C62
3. Shen Y, Win MZ (2010) Fundamental limits of wideband localization part I: a general framework. IEEE Trans Inf Theory 56(10):4956–4980
4. Wang T (2012) Cramer-Rao bound for localization with a priori knowledge on biased range measurements. IEEE Trans Aerosp Electron Syst 48(1):468–476
5. Larsson EG (2004) Cramer-Rao bound analysis of distributed positioning in sensor networks. IEEE Sig Process Lett 11(3):334–337
6. Penna F, Caceres MA, Wymeersch H (2010) Cramer-Rao bound for hybrid GNSS-terrestrial cooperative positioning. IEEE Commun Lett 14(11):1005–1007
7. Qi Y, Kobayashi H (2002) Cramer-Rao lower bound for geolocation in non-line-of-sight environment. In: Proceedings of ICASSP, pp 2473–2476
8. Li S, Hedley M, Collings IB (2015) An empirical ranging error model and efficient cooperative positioning for indoor applications. In: 2015 IEEE international conference on communication workshop (ICCW), pp 773–778
9. Alsindi NA, Alavi B, Pahlavan K (2009) Measurement and modeling of ultrawideband TOA-based ranging in indoor multipath environments. IEEE Trans Veh Technol 58(3):1046–1058
10. Shi X, Mao G, Yang Z, Chen J (2016) Localization algorithm design and performance analysis in probabilistic LOS/NLOS environment. In: 2016 IEEE international conference on communications (ICC), pp 1–6
11. van Trees HL (1968) Detection, estimation, and modulation theory, part I. Wiley, Hoboken, NJ
12. Wymeersch H, Lien J, Win MZ (2009) Cooperative localization in wireless networks. Proc IEEE 97(2):427–450
13. Dardari D, Conti A, Ferner U, Giorgetti A, Win MZ (2009) Ranging with ultrawide bandwidth signals in multipath environments. Proc IEEE 97(2):404–426
14. Nguyen TV, Jeong Y, Shin H, Win MZ (2015) Least square cooperative localization. IEEE Trans Veh Technol 64(4):1318–1330
15. Gentile C, Lopez SM, Kik A (2009) A comprehensive spatial-temporal channel propagation model for the ultra-wideband spectrum 2–8 GHz. In: GLOBECOM 2009–2009 IEEE global telecommunications conference, Honolulu, HI, pp 1–6

Research on Simultaneous Localization and Mapping Algorithm Based on Monocular Vision

Han Zhang, Baowang Lian and Hu Lu

Abstract Simultaneous Localization and Mapping (SLAM) technology can make the robot in the unknown area positioning and building the map. Aiming at the problem of the indoor positioning in a small area, SLAM algorithm based on monocular camera was used. Feature based method was introduced to estimate the position of the robot and build the local map, and then BA (Bundle Adjustment) and graph optimization was used to optimize in the back-end. In order to solve the problem of scale drift in positioning, the bag-of-words (BoW) model and loop closing were used to reduce the system error, and an overall map was constructed at last. The experiment results show that the proposed algorithm can construct a sparse map in a small indoor scene, and the positioning error in the x direction reaches 0.0528 m, which meets the accuracy requirements and effectively reduces the scale drift.

Keywords SLAM · Feature points · Monocular · Pose graph optimization Loop closing

1 Introduction

With the development of science and technology, indoor positioning technology is more and more mature. The positioning technology of mobile robot is the key to indoor positioning, and gradually realizes the positioning, navigation and barrier functions [1]. Simultaneous localization and mapping (SLAM), one of the most advanced research methods of mobile robot localization technology, was first proposed by Randall Smith and Peter Cheseseman in 1988 [2]. SLAM problem

H. Zhang · B. Lian
School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China

H. Lu (✉)
School of Information and Navigation, Air Force Engineering University, Xi'an 710129, China
e-mail: 10956087@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_70

refers to a sensor-equipped mobile device entering an unfamiliar environment, building a map by its own movement, and implementing its ownpositioning [3].

When the sensor is a camera, it becomes a visual navigation. Visual navigation has the advantages of large amount of information, high flexibility, and has an unparalleled status in indoor positioning technology [1]. Monocular cameras are widely used because of their low cost.

There are many monocular SLAM solutions, mainly divided into two parts which are based on the filter and key-frame BA. Early programs are basically using Kalman filter technology to achieve optimization [4]. This algorithm only considers the state of the current moment, and the states of other moments do not participate in the optimization calculation. The error of the current frame will be passed to the next frame, which will result in the accumulation of errors. Therefore, the EKF algorithm is only suitable for small-scale scenes. However, the graph optimization theory based on nonlinear optimization can be applied to a wide range of scenes, so it has become a hot research topic nowadays. In the graph optimization, robot poses and road signs are taken as the nodes and the observation equation is the edge. The optimization problem is expressed as a graph, and the objective error is taken as the objective function, which is minimized to optimize the pose and road sign nodes [5]. During this process, Bundle Adjustment (BA) [6] plays a central role. Due to the large number of landmarks and pose points in the global optimization, the early SLAM was not suitable for real-time calculation, and it was not realized until people recognized the sparsity of Bundle Adjustment (BA).

MonoSLAM is the first vision SLAM system based on a monocular camera [1]. It uses Kalman filter for back-end optimization, but its systematic error is relatively large and computationally intensive. PTAM [7] (Parallel Tracking and Mapping) is the first keyframe-based solution that tracks and constructs maps simultaneously. However, PTAM only applies to small-scale scenes, and the ability to relocate is poor, making it easy to track loss in some complicated scenes. Therefore, following PTAM, Mur-Artal et al. proposed ORB-SLAM (Oriented FAST and Rotated BRIEF-Simultaneous Localization And Mapping) algorithm [8]. This algorithm has three threads: tracking, mapping and loop closing. The feature points are extracted and matched by ORB (Oriented FAST and Rotated BRIEF) [9] feature points method in the algorithm, and three threads are carried out in parallel. The algorithm has the ability of relocalization, and it can detect a large loop, with improving the positioning accuracy.

For monocular system, although it has the advantages of high flexibility and low cost, lack of depth information will result in low positioning accuracy and deviation of trajectory. In order to solve this problem, this paper uses ORB-SLAM algorithm to test in a small-scale scene. The ORB feature points method is used to extract and match the feature points, and initialize them at the front of the visual odometer. At the same time, the results are optimized by graph optimization at the back end, and the sparse map is obtained by loop detection and Sim3 optimization, which can improve the accuracy and reduce the scale drift.

2 Monocular ORB-SLAM Algorithm

2.1 Initialization

Because the monocular SLAM system does not have the depth information, it is necessary to estimate the relative pose of the camera through initialization, and calculate the depth of the map 3D points by triangulation.

In order to estimate the relative pose of the camera, feature points need to be extracted and matched between the current frame and the reference frame, and then extract key frames. In the experiment, feature points were extracted and matched by ORB feature points method, and two threads begin at the same time to calculate homography matrix H and fundamental matrix F [8]. The matrix H is suitable for a planar scene, while the matrix F is suitable for a non-planar scene. H and F matrices are respectively calculated by using the Direct Linear Transformation (DLT) method and the 8-point algorithm [10]. RANSAC (Random Sample Consensus) iteration is used to calculate the symmetrical transmission error between the frames where each pair of matching points is located, and then it is easy to determine whether the match point is the interior point by comparing with the threshold. Finally, accumulate all interior points to get the total score. The process is shown in Eqs. (1) and (2):

$$S_M = \sum_i \left(\rho_M\left(d_{cr}^2\left(x_c^i, x_r^i, M\right)\right) + \rho_M\left(d_{rc}^2\left(x_c^i, x_r^i, M\right)\right)\right) \tag{1}$$

$$\rho_M\left(d^2\right) = \begin{cases} \Gamma - d^2, & if \quad d^2 < T_M \\ 0, & if \quad d^2 \geq T_M \end{cases} \tag{2}$$

M is the H and F matrix, x_c and x_r are pairs of matching points, d_{cr}^2 and d_{rc}^2 are the symmetric transfer errors between frames where the matching points are located, and T_M is an exclusion threshold of invalid data. After multiple selection, compare each score and take the highest one.

According to $R_H = S_H / S_H + S_F$, if the value is greater than 0.45, choose H matrix model, or choose F matrix model. Then through the analytic method, the rotation matrix R and the translation vector t can be decomposed from the H matrix. For the F matrix model, we can transform it into a essential matrix E, then restore the pose vector R, t of the camera through the singular value decomposition (SVD). The essential matrix E is as follows:

$$E = K^{\mathrm{T}} F K \tag{3}$$

After recovering the camera movement, the depth information can be estimated through triangulation, so that the 3D point of the map will be obtained.

2.2 Local Mapping

After obtaining the camera pose and the initialized map 3D points of key frames, then optimize them by using BA (Bundle Adjustment) and graph optimization to minimize the error. We represent the optimization problem in the form of a graph with the camera poses and map points as nodes and observation equation $z = h(x, p)$ as the edge for optimization. Where x is the camera pose, p is the three-dimensional point in the map.

The error function of the observation equation is shown in Eq. (4):

$$e = z - h(x, p) \tag{4}$$

Overall error is obtained as follows:

$$\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\left\|e_{ij}\right\|^{2} = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\left\|z_{ij} - h(x_i, p_j)\right\|^{2} \tag{5}$$

The Levenberg-Marquardt method (LM) is used to iterate and solve the sparse incremental equations $H_i \Delta x_i = b$ for minimizing the system error and obtaining the optimal solution.

A local map can be constructed by optimizing the three-dimensional points of the map and the camera poses in the key frames. Before the complete map is constructed, the local map needs to be updated all the time. Therefore, it is necessary to insert the new key frames and establish the connecting relationship between the current key frame and other key frames, so that the matching between the feature points can be updated. Finally, new three-dimensional points can be generated by triangulation and then optimized by making a new local BA optimization.

2.3 Loop Closing

When feature points are extracted, they are placed in a bag of words vector [11] which then are constructed in a bag of words model. After that, loop detection and relocalization are performed which can reduce the system error and solve the scale drift problem. And then the minimum similarity is obtained by comparing the bag of words vector between the current key frame and the previous key frame. The key frame with the similarity which is not less than the minimum similarity is searched as a candidate key frame in the key frame database. Finally, Closed-loop frames are determined by RANSAC to detect the consistency of each candidate key frame with the previous loop.

Because of the lack of depth information, monocular SLAM systems require seven degrees of freedom (DoFs), with the exception of six rotation and translation

vectors, as well as a scaling factor [12]. The transformation matrix of monocular system is shown in Eq. (6):

$$T = \begin{bmatrix} sR & t \\ 0 & 1 \end{bmatrix} \tag{6}$$

In order to solve the problem of scale drift, Sim3 optimization which is shown in Eq. (7) is performed on the pose during the loop detection:

$$r_{i,j} = \log_{sim(3)}(x_{i,j}x_j x_i) \tag{7}$$

$r_{i,j}$ is the residual between the pose of x_i and the pose of x_j, $x_{i,j}$ is the similar transformation between the pose of x_j and x_i. Graph optimization with similarity constraints can be obtained by minimizing Eq. (8) which is achieved by LM algorithm:

$$R = \sum_{i,j} r_{i,j}^T \Delta_{i,j} r_{i,j} \tag{8}$$

More feature points will be found by using the optimization results, then the scale and pose calibration will be achieved by combining the loop-back frame information with the current frame.

3 Experimental Results and Analysis

The experiment is carried on the laptop of i5-8250U, 2.4 GHz CPU, 8 GB RAM, ubuntu 14.04 operating system.

3.1 ORB-SLAM and PTAM Comparison and Dataset Testing

References [13, 14] provide 22 dataset sequences captured in a vehicle environment and compare various algorithms. The car used for data collection is equipped with two grayscale cameras and two color cameras. The exact coordinates of the path are provided by the laser scanner and the GPS positioning system. The 13th sequence of dataset and the evaluation results [14] are used to compare ORB-SLAM and PTAM algorithm in positioning accuracy. Figure 1 is path comparison charts by respectively using ORB-SLAM algorithm and PTAM algorithm.

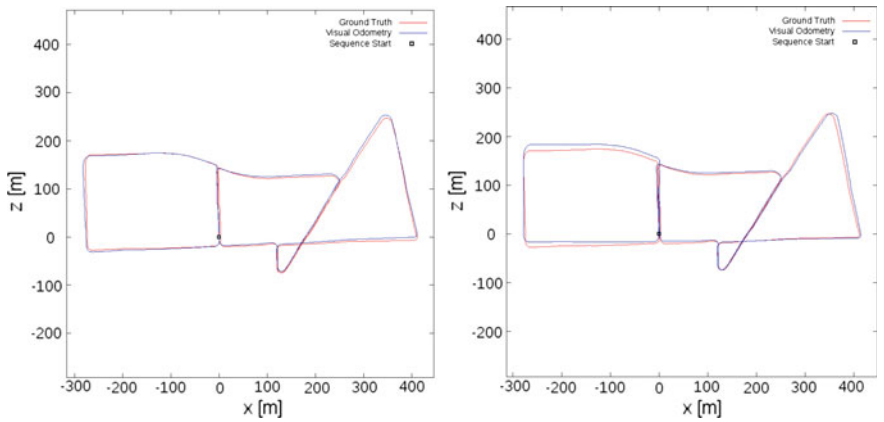


Fig. 1 Comparison of the ground truth (*red*) and the estimated path (*blue*) in ORB-SLAM (Left) and PTAM (Right) algorithm

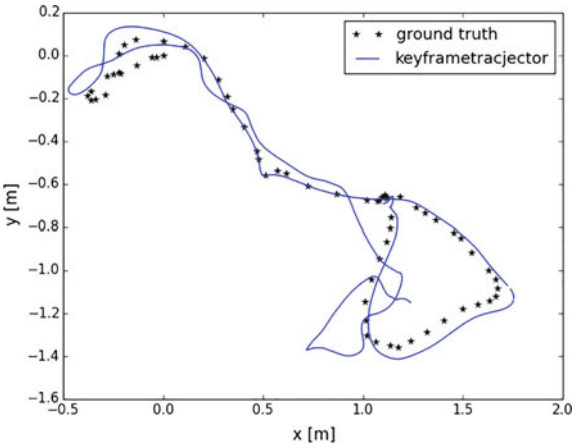


Fig. 2 Comparison between the ground truth (*black*) and the estimated path (*blue*) of FR1 desk sequence

The experiment results show that the ORB-SLAM algorithm has a better match between the estimated trajectory and the real trajectory. With the expansion of the range, the error of the ORB-SLAM algorithm is less than that of the PTAM, especially at the rotation. This shows that ORB-SLAM algorithm has higher accuracy than PTAM algorithm. This is because ORB-SLAM uses the ORB feature with higher matching accuracy, which also has directionality and rotation invariance. However, PTAM extracts FAST corners as feature points which do not have directivity, so that contributes to poor relocalization capability. On the other hand, ORB-SLAM uses graph optimization instead of global BA, and eliminates error accumulation through loop detection, so it is more suitable for larger-scale scenes.

Figure 2 shows the simulation of the TUM RGB-D dataset [15] by using the ORB-SLAM algorithm. The dataset consists of depth pictures and color pictures

captured by Kinect. We chose the sequence FR1 desk that is more suitable for monocular systems and only use the color pictures.

According to the Root Mean Square Error (RMSE) tool which is used to estimate the true pose and estimated pose of the camera provided in reference [15], the RMSE of translation component of FR1 desk is 0.101 m.

3.2 Indoor Environment Experiment

In order to verify the application of monocular ORB-SLAM algorithm in a small indoor scene, the laptop is used as a mobile robot to collect data in a dormitory by using its own camera. The indoor environment is shown in Fig. 3. Because there are a large number of objects in the dormitory corner, it is conducive to the extraction of feature points, so that it is not easy to loss of tracking. The entire experimental environment is a rectangular area of 3 m in length and 1.2 m in width. During the experiment, the laptop is taken in hand and moves at a speed of 0.5 m per second, and two loops were taken at last. In the end, the sparse map and path trajectory of the indoor environment are obtained, which realizes the simultaneous localization and mapping. A total of 4295 pictures were processed throughout the experiment at a frame rate of one frame per second. Figure 4 is the map without loop detection, and Fig. 5 is the map after loop detection.

From Figs. 4 and 5, it can be found that the loop before doing loop detection is not closed. After loop detection and Sim3 optimization, the loop is closed and the trajectory is corrected, which contributes to reduce the scale drift.

Figure 6 is the comparison between the true trajectory and the estimated trajectory. In the figure, the place where the second loop starts is the origin, and make it coincide with the place where the actual path begins, finally the path is drawn. The dashed blue line is the estimated path, and the solid black line is the true path. A total of 368 key frames were inserted during the entire experiment. It can be seen from Fig. 6 that there are differences between the estimated path and the true path. In order to analyze the result, the root mean square error of the true position and the estimated position is thus calculated in the x and y directions respectively. Due to

Fig. 3 Indoor environment

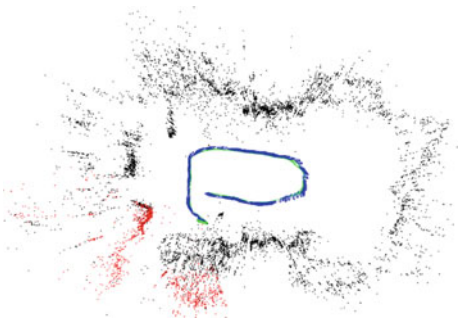


Fig. 4 Sparse map before loop closing. The key frame is draw in blue, the trajectory of camera is in green, and map points are in black and red

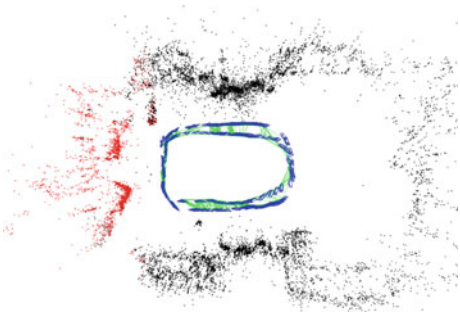


Fig. 5 Sparse map after loop closing

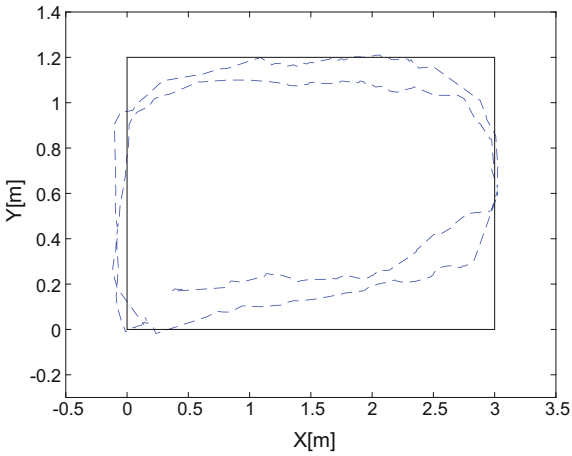


Fig. 6 Comparison of the ground truth (*black*) and the estimated path (*blue*)

the lack of motion capture system, the real coordinates of the actual path can not be known. To make it easy to calculate the measurement error, four points at the corner of the actual path and the estimated path are selected as experimental points. In the process of extracting the coordinates of the point in the estimated path, it is necessary to combine the map point coordinate file generated in the experiment and the

Table 1 Results on position of the measured points

Test point	True values (m)	Estimated values (m)
1	(0, 0)	(−0.0801 ,0.1822)
2	(3, 0)	(2.9324, 0.3049)
3	(3, 1.2)	(3.0088, 0.8547)
4	(0, 1.2)	(0.0085, 0.9620)

pictures corresponding to the key frame to ensure accurate coordinate values. Table 1 shows the 4-point coordinates of the real path and the estimated path.

The root mean square error is calculated as shown in Eq. (9):

$$RMES = [(\sum_{i=1}^{n} (x_i - x_j)^2/n)]^{1/2} \tag{9}$$

x_i is the x-coordinate of the estimated value, x_j is the real-world coordinate and n is 4. As a result, the RMES in the x direction is 0.0528 m, and the RMSE in the y direction is 0.2748 m. It can be seen that through loop detection and Sim3 optimization, the system can achieve high accuracy in the x direction. However, the displacement of the camera will become smaller at the corner which results in a larger drift. So the error in the y direction is still relatively large.

4 Conclusion

In order to reduce the impact of scale drift on the monocular SLAM system, ORB-SLAM algorithm is used to track, build map and detect loop at the same time for estimating the camera pose and map points in the small area scene with monocular camera. After that Sim3 optimization is used to correct the scale, finally, we can get sparse maps and trajectory. The experiment results show that in the range of 3 m in length and 1.2 m in width, the loop detection is performed when the camera reaches the same place, and the scale is calibrated. The root mean square error in the x direction is 0.0528 m, which meets the accuracy requirements. In addition, the algorithm is also applicable to outdoor scenes, so indoor and outdoor positioning can be achieved at the same time. However, there is still a certain degree of scale drift in the monocular system. Therefore, in the future work, further research on the monocular SLAM system is needed to do for improving the positioning accuracy of the system.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Grant Nos. 61473308).

References

1. He J, Li Z (2010) A survey of vision based simultaneous localization and mapping. Appl Res Comput 27(8):2839–2844
2. Li J (2013) Research on SLAM problem of mobile robot based on monocular vision. Harbin Institute of Technology Press, Harbin
3. Durrant-Whyte H, Bailey T (2006) Simultaneous localization and mapping: part I. IEEE Robot Autom Mag 13(2):99–110
4. Liu H, Zhang G, Bao H (2016) A survey of simultaneous localization and mapping based on monocular vision. J Comput Aided Des Comput Graph 28(6):855–867
5. Kummerle R, Grisetti G, Strasdat H et al (2011) G2O: a general framework for graph optimization. In: IEEE international conference on robotics and automation, Shanghai, pp 3607–3613
6. Triggs B, McLauchlan PF, Hartley RI et al (2000) Bundle adjustment a modern synthesis. In: Vision algorithms: theory and practice, Springer, New York, pp 298–322
7. Klein G, Murray D (2009) Parallel tracking and mapping on a camera phone. In: IEEE international symposium on mixed and augmented reality, Orlando, FL, pp 83–86
8. Mur-Artal R, Montiel JMM, Tardos JD (2015) ORB-SLAM: a versatile and accurate monocular SLAM system. In: IEEE transactions on robotic, pp 1147–1163
9. Rublee E, Rabaud V, Konolige K et al (2011) ORB: an efficient alternative to SIFT or SURF. In: International conference on computer vision, Barcelona, pp 2564–2571
10. Hartley RI (1997) In defense of eight-point algorithm. IEEE Trans Pattern Anal Mach Intell 19:580–593
11. Mur-Artal R, Tardos JD (2014) Fast relocalization and loop closing in keyframe-based SLAM. In: IEEE international conference on robotics and automation (ICRA), Hong Kong, pp 846–853
12. Strasdat H, Montiel JMM, Davision AJ (2010) Scale drift-aware large scale monocular SLAM. In: Robotics: science and systems VI, Spain, pp 27–30
13. Geiger A, Lenz P, Stiller C, Urtasun R (2012) Are we ready for autonomous driving? The KITTI vision benchmark suite. In: IEEE conference on computer vision and pattern recognition, Providence, RI, USA, pp 3354–3361
14. http://www.cvlibs.net/datasets/kitti/eval_odometry.php, Cited 28 Nov 2017
15. Sturm J, Engelhard N, Endres F et al (2012) A benchmark for the evaluation of RGB-D SLAM system. In: IEEE/RSJ international conference on intelligent robots and systems, pp 573–580

Geomagnetism Sensing and Calibration Systems of LING QIAO Satellite: Results Over Three Years of In-Orbit Operation

Xi Chen, Chenggong Zheng and Jin Jin

Abstract Geomagnetism-based attitude determination is a popular method for modern small/macro satellites for availability of small form factor, low power magnetometers and their accuracy in meeting the attitude precision requirements of many space missions. However, it is not an easy work to achieve high attitude precision for small mobile communication satellites by geomagnetism-based attitude sensing due to body magnetic interferences, either generated by high power current loop or complex magnetized field of ferromagnetic materials such as semi-rigid cable. In this work, we present the design and in orbit experiment results in the geomagnetism sensing and calibration systems of LING QIAO experimental mobile communication satellite. The LING QIAO satellite has worked in orbit steadily for more than three years, the in orbit results show an long term attitude control precision of 1° was achieved and short term attitude variation given by the MEMS gyroscope is less than 0.3°.

Keywords Satellite · Geomagnetism sensing · Magnetometer · LING QIAO

1 Introduction

Geomagnetism-based attitude determination is fundamental to modern small/macro satellites. Before the stable attitude is established, magnetometers is the only attitude sensor in work. Only after attitude stabilization, star sensors and earth sensors begin to work. In many cases, magnetometers are the only sensors of a satellite. In many cases, magnetometers are used for their small form factor, low power, low

X. Chen
WBC Lab, Shenzhen Tsinghua Research Institute, Shenzhen, China

C. Zheng (✉)
School of Aerospace, Tsinghua University, Beijing, China
e-mail: cgzheng@126.com

J. Jin
Space Center, Tsinghua University, Beijing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_71

cost and their accuracy in meeting the attitude precision requirements of many space missions.

The accuracy of geomagnetism-based attitude sensing is determined by the sensitivity of magnetometers, quality of geomagnetism model and interferences of satellite current loop and body magnetic magnetization.

During the past decades, designing better magnetometers have always been an active research topic in both the academic and industries. Considering that the band pass filter for filtering out the 2nd order harmonics is a performance limiting factor, an orthogonal fluxgate employing digital selective band pass sampling is developed in [1, 2], which allows the fluxgate output to be sampled only once at a single time instance during a number N of excitation cycles. This provides reconstruction of a measured magnetic field with a 40% reduction of noise. In [3, 4], a three-axial search coil magnetometer had been designed and built, which is optimized for 20 MHz to 7 Hz frequencies and for ultra-low (252 μW) power consumption. In [5–7], digital fluxgate magnetometer was developed which substituted part of the analog circuit by digital processors.

The output of magnetometers are compared to the value in magnetic field model of the Earth to estimate the satellite's attitude angles in [8–10]. International Geomagnetic Reference Field (IGRF) model finds out the magnetic fields in the near-Earth space environment, which has a history that goes back 1900s. IGRF model describes the internal part of the geomagnetic field [11, 12] and [13, 14], while CHAOS-6 and T89 (Tsyganenko 1989 model), include simple parameterizations of the external magnetic fields from different magnetospheric sources superimposed on the internal geomagnetic field. Until now, the most commonly used geomagnetic field model is IGRF model.

For a LEO satellite, output of magnetometers are true geomagnetic value superimposed on by body magnetic interferences, which may originates from momentum wheels, body current loop and body magnetic material magnetization. In most of the cases, body magnetic interferences are main error sources of geomagnetism-based attitude determination, which is especially true for a small communication satellite which has high power current loop and ferromagnetic semi-rigid cables.

In this work, we report the design and in orbit experiment results of the geomagnetism sensing and calibration systems of LING QIAO experimental mobile communication satellite. The LING QIAO satellite has worked in orbit steadily for more than three years, the in orbit results show an long term attitude control precision of 1° has been achieved and short term attitude variation given by the MEMS gyroscope is less than 0.3°.

2 The Geomagnetic Sensors of LING QIAO

The Geomagnetism Sensing System of LING QIAO logically includes two separate fluxgate magnetometers. The probe and analog part of the main magnetometer is a COTS product, which is produced by a space company in China that honors more

than ten-times space flying experiences. The probe and analog part of the secondary magnetometer was a self-developed product, which had no space flying experiences. The digital circuitries of the two magnetometers are designed on a single PCB board. The digital circuity of each magnetometer consist of a MCU (Micro-Controller Unit), an ADC (Analog to digital converter) and 1/4 multiplexer. The MCU controls the multiplexer to select the outstanding channel to convert from analog to digital. The MCU reads out the ADC results for processing, which conducts all the work that acts as the so called mathematical platform.

Structure of the self-developed fluxgate magnetometer is given by Fig. 2. The analog part of the self-developed magnetometer has four channels: X, Y and Z channel are 3-axial fluxgate channel probing the environmental magnetic field for DS to hundreds of Hz; an additional channel, i.e., the temperature sensor channel is added to sense the temperature of the 3 fluxgate probes. The X, Y and Z-channel are the same with X-channel given in detail in upper part of Fig. 1.

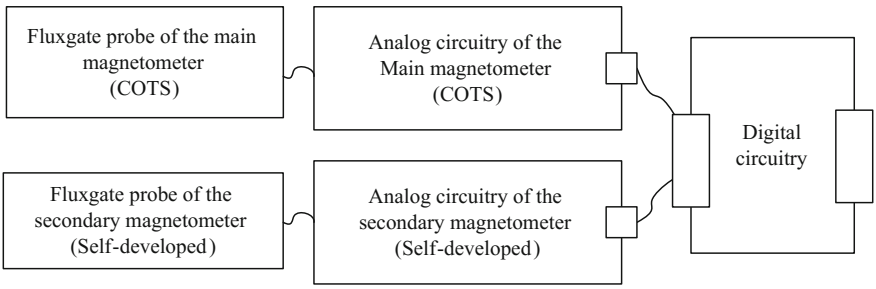


Fig. 1 Block diagram of the geomagnetism sensing system

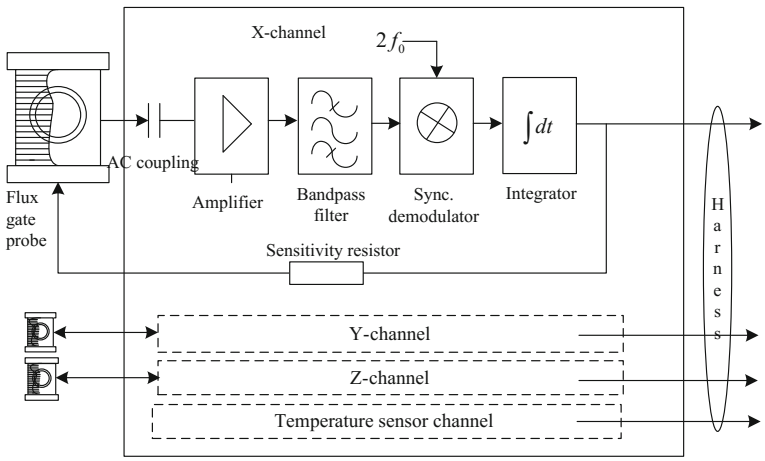


Fig. 2 Analog circuitry of the self-developed fluxgate magnetometer

The output of a fluxgate probe is AC coupled to an analog circuit. The first stage is the low noise amplifier which amplifies all input signals, the second state is a band-pass filter for filtering out signals besides the 2nd harmonic. A synchronization demodulator acts as the third stage to detect the envelope of the 2nd harmonic. The output of the synchronization demodulator is fed to an integrator, which acts as the final stage of the analog circuits. The output of the integrator is proportional to the amplitude of the environmental magnetic field whose direction is parallel to the normal of the fluxgate coil. The feedback path makes the fluxgate coil work around zero regions. The sensitivity resistor in the feedback path controls how much of the output is feedback to the feedback coil.

3 Magnetometer Calibration

3.1 The Calibration Flow

There exist nonlinearity, temperature drift and orthogonal error due to the magnetometer's manufacturing process. Digital calibration method is employed to cancel these errors.

The calibration is done in the processor of the digital circuitry, as given by Fig. 3. During data collection process, original DA output, probe temperature and circuit temperature are collected for calculating the calibration parameters. The finalized calibration parameters are then written to the non-volatile memory of the digital circuitry to run calibration. Calibrated output is the final output of a magnetometer.

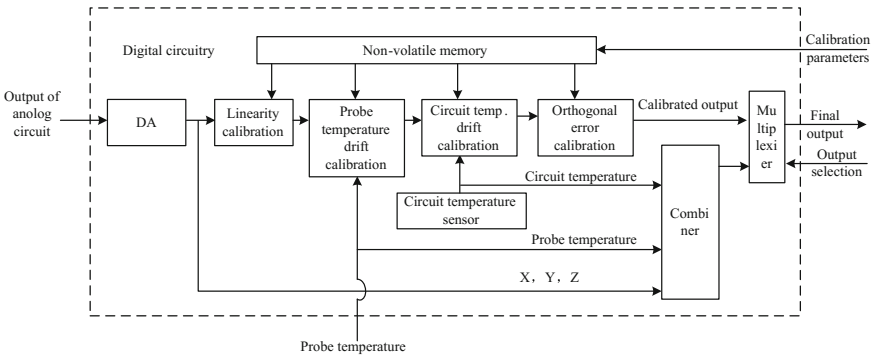


Fig. 3 The calibration flow

3.2 The Linearity and Temperature Drift Calibration

The linearity and temperature drift calibration can be expressed by a single formula as given below with x-axis as an example:

$$\begin{aligned} x_c &= r'_x r_x [\, k_{x,N-1} \quad k_{x,N-2} \quad \cdots \quad 1\,] K_{T_0,x} \\ &= r'_x r_x (k_{x,N-1} x^{N-1} + k_{x,N-2} x^{N-2} + \cdots + k_{x,1} x + k_{x,0}) \end{aligned} \tag{1}$$

where x is the x-axis DA output, x_c is the linearity and temperature drift calibrated x-axis output. $k_{x,N-1}, \ldots, k_{x,0}$ is x-axis linearity coefficient, r_x is x-axis temperature drift coefficient and r'_x is x-axis circuit drift coefficient. All the above coefficient can be obtained in during ground experiment or by standard magnetic lab.

3.3 The Orthogonal Error Calibration

Given a reference coordinate $O{-}x'y'z'$, and $O{-}xyz$ is the actual polar coordinate. Z-axis and Z′-axis coincide, Y-axis lies in the plane $z'Oy'$, the projection line of X on plane $z'Ox'$ is x''.

$$\begin{cases} B_x = B_{x'} \cos q \cos p + B_{y'} \cos \alpha' + B_{z'} \cos \gamma' \\ B_y = B_{y'} \sin \beta + B_{z'} \cos \beta \\ B_z = B_{z'} \end{cases} \tag{2}$$

The relationship of $O{-}x'y'z'$ and $O{-}xyz$ can be formulated as (2).

Where $B_{x'}$, $B_{y'}$, $B_{z'}$ are the projection of magnetic field intensity B in $O{-}x'y'z'$ (see Fig. 5).

$$\begin{cases} B_{x1} = B \sin\theta \cos\phi \\ B_{y1} = B \sin\theta \sin\phi \\ B_{z1} = B \cos\theta \end{cases} \tag{3}$$

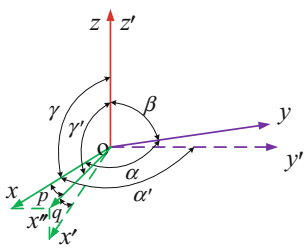


Fig. 4 The reference coordinate and the actual coordinate

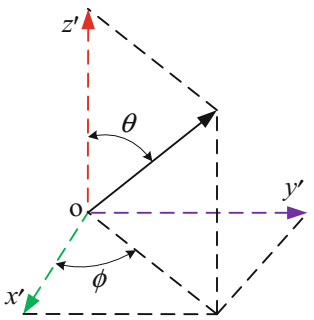


Fig. 5 The projection of magnetic field in $O-x'y'z'$

The orthogonal degree of three axis is quite small, i.e. $\left|\frac{\pi}{2}-\alpha\right|$, $\left|\frac{\pi}{2}-\beta\right|$, $\left|\frac{\pi}{2}-\gamma\right|$ are not more than 1°. Hence we have $\gamma-\gamma'\leq 0.00006°$, $\alpha-\alpha'\leq 0.02°$, $\left|\frac{\pi}{2}-\alpha\right|-p\leq 0.02°$, $q=\left|\frac{\pi}{2}-\gamma'\right|$ (Fig. 4).

so

$$\cos q=\cos\left|\frac{\pi}{2}-\gamma'\right|=\sin\gamma'\approx\sin\gamma$$
$$\cos p\approx\cos\left|\frac{\pi}{2}-\alpha\right|=\sin\alpha$$
$$\cos\alpha'\approx\cos\alpha$$
$$\cos\gamma'\approx\cos\gamma$$

Hence Eq. (2) can also be formulated as

$$\begin{cases}B_x=B_{x'}\sin\alpha\sin\gamma+B_{y'}\cos\alpha+B_{z'}\cos\gamma\\ B_y=B_{y'}\sin\beta+B_{z'}\cos\beta\\ B_z=B_{z'}\end{cases}\tag{4}$$

So we have orthogonal error calibration formula as

$$\begin{cases}B_x=(x_c-(y_c-z_c\cos\beta)\cos\alpha/\sin\beta-z_c\cos\gamma)/\sin\alpha\sin\gamma\\ B_y=(y_c-z_c\cos\beta)/\sin\beta\\ B_z=z_c\end{cases}\tag{5}$$

4 Interference Cancellation

As a small communication satellite, LING QIAO has very high power output when its mobile communication payload is working. Besides, all the other subsystems all generate magnetic interferences to its magnetometer. Such interferences are time-varied and very strong so cannot be ignored for a communication satellite.

Subsystem interconnection of LING QIAO is given by Fig. 6. To cancel interference, the measured real-time current value of all subsystems are periodically sent to the attitude determination and control subsystem (ADCS) via its CAN bus. During the ground manufacture process, the magnetic interference of each subsystem to the magnetometer is measured and modeled for in-orbit use.

Such a cancellation system reduces the impact of body inference to less than 10% or its original value, as confirmed by repeated test results.

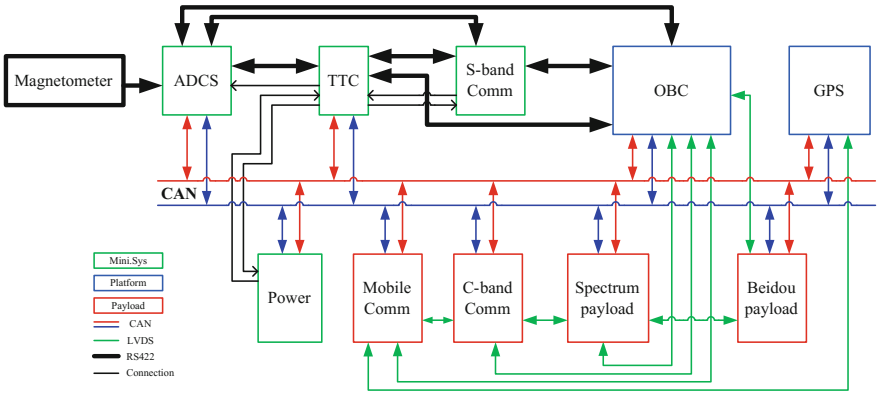


Fig. 6 Subsystem interconnection of LING QIAO

5 In-Orbit Results

A 24-h in-orbit trace of the secondary magnetometer on July 17–18, 2017 is shown in Fig. 7. The trace shows that the magnetometer had been working well on the satellite and had been proving a reliable geomagnetic measurement values for the attitude determination and control system. The maximum geomagnetism measured in LING QIAO orbit is no more than 45,000 nT.

In Fig. 8, we show the relationship between the output of the secondary magnetometers and latitude.

In Figs. 9 and 10, we give the 24-h trace of the MEMS gyroscope of ADCS system of LING QIAO satellite and the attitude of the LING QIAO satellite. According to this result, short term attitude variation of LING QIAO is less than 0.3°, which is good enough for the requirement of mobile communications.

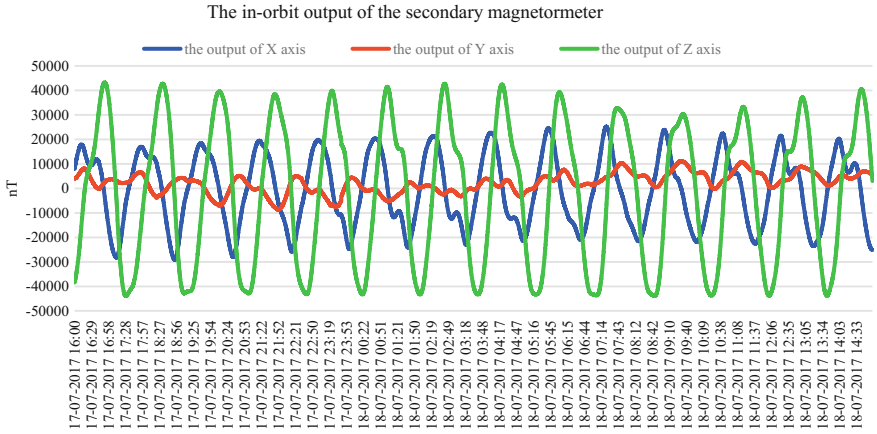


Fig. 7 24-h in-orbit log July 17–18, 2017

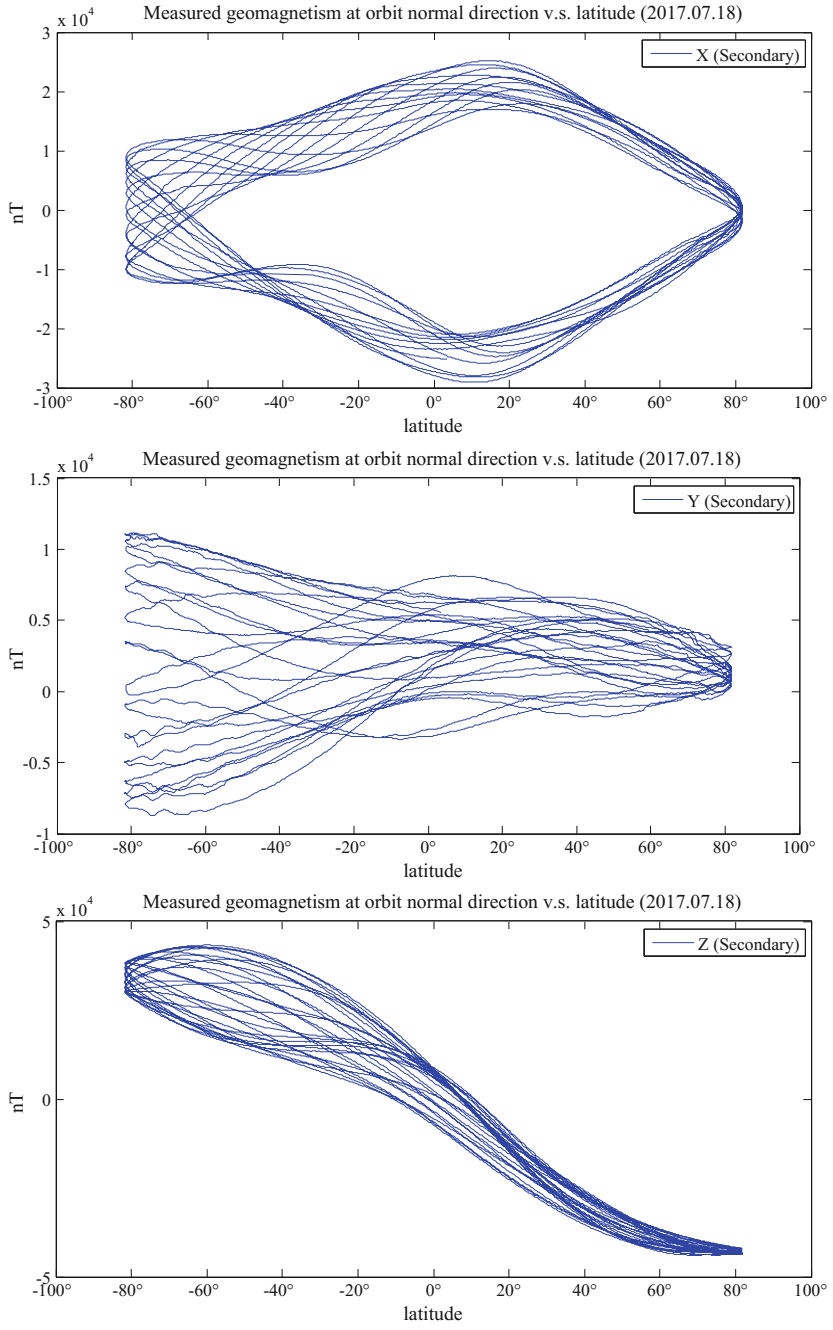


Fig. 8 Measured geomagnetism of the secondary magnetometers versus latitude

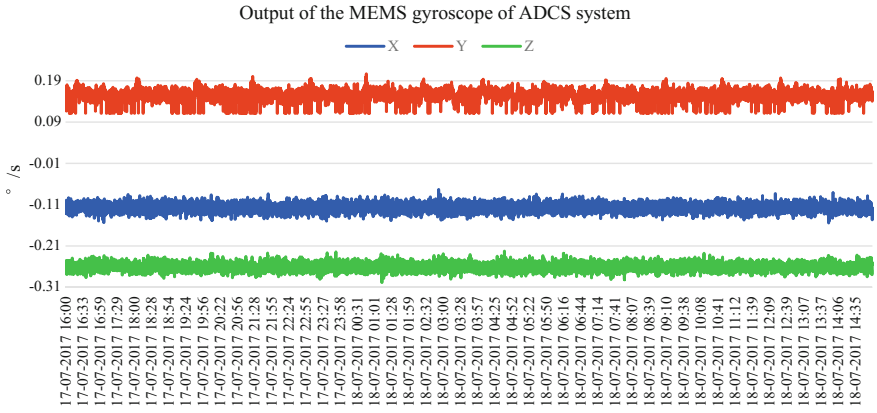


Fig. 9 The 24-h trace of the MEMS gyroscope of ADCS system of LING QIAO satellite July 17–18, 2017

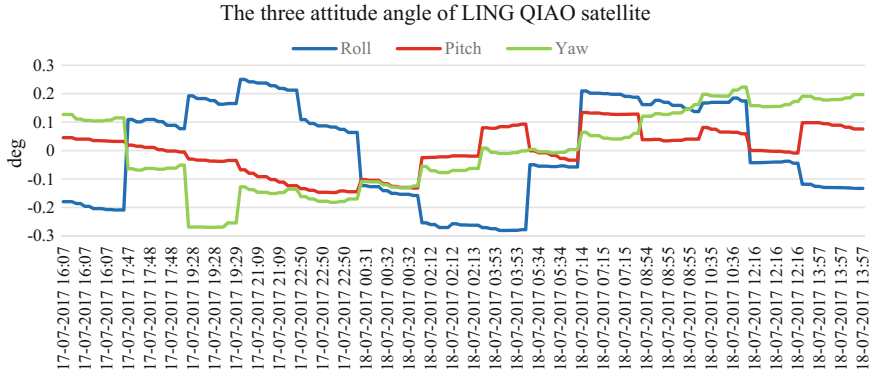


Fig. 10 The three attitude angle of LING QIAO satellite July 17–18, 2017

6 Conclusions

Geomagnetism-based attitude determinations have been widely used in modern small/macro satellites, which is attribute to the availability of small form factor, low power magnetometers and their accuracy meeting the attitude precision requirements of the space mission. In this work, we report the design of the geomagnetism sensing and calibration systems and in orbit results of LING QIAO experimental mobile communication satellite. The LING QIAO satellite has worked in orbit steadily for more than three years, the in orbit results show that an long term attitude control precision of 1° has been achieved and the short term attitude variation given by the MEMS gyroscope is less than 0.3°. In future space missions, the

self-developed magnetometer will replace the COTS magnetometer. We are designing a new generation of magnetometers that is expected to be more compact, low power and light weight.

Acknowledgements This work was supported in part by the Science, Technology and Innovation Commission of Shenzhen Municipality (No. JSGG20160429165838848).

References

1. Grosz A, Paperno E, Amrusi S, Zadov B (2011) A three-axial search coil magnetometer optimized for small size, low power, and low frequencies. IEEE Sens J 11:1088–1094
2. Auster H et al (1995) Concept and first results of a digital fluxgate magnetometer. Meas Sci Technol 6:477–481
3. Heinrich G et al (2003) Array of miniaturized fluxgate sensors for nondestructive testing applications. Sens Actuators A 106:326–328
4. Iguchi K, Matsuoka A (2014) A digital-type fluxgate magnetometer using a sigma-delta digital-to-analog converter for a sounding rocket experiment. Meas Sci Technol 828–838
5. Rani BS, Gomathy C, Sowmya B, Narmadha R (2010) Attitude determination and control system of Sathyabamasat. In: Recent advances in space technology services and climate change 2010, pp 326–329
6. Li C, Wang L, Li X (2013) Method of attitude-aided magnetometers/SINS/GNSS integration. In: 2013 6th international conference on information management, innovation management and industrial engineering, pp 304–308
7. Chen X, Wang J, Gao W (2015) A digital fluxgate magnetometer and geomagnetic measurement system. 201510098779.X
8. Guler DC, Kaymaz Z, Hajiyev C (2017) Geomagnetic models at low earth orbit and their use in attitude determination. In: 2017 8th international conference on recent advances in space technologies, pp 301–305
9. Weiss E, Grosz A, Amrusi S et al (2012) Orthogonal fluxgate employing digital selective bandpass sampling. IEEE Trans Magn 48:4089–4091
10. Zhou X (2007) Design and achievement of a fully digital fluxgate magnetometer. Master degree, Zhejiang University, Jiangsu
11. Drljaca PM, Kejik P, Vincent F et al (2004) Single core fully integrated CMOS miro-fluxgate magnetometer. Sens Actuators A 110:236–241
12. Lee S-H, Rhee S-W, Shin I (2002) On-orbit magnetic field correction of magnetometers and geometry-based attitude determination for small satellites. In: Proceedings of IEEE aerospace conference, pp 5-2309–5-2317
13. Chen YT, Lu CC, Sung GM et al (2011) Design and implementation of driving circuits for micro fluxgate sensors employing multiple-harmonic characteristics. In: Proceedings of IEEE international conference on nano/micro engineered and molecular systems, pp 1071–1074, 20–23
14. Somov Y, Butyrin S, Hajiyev C (2015) In-flight calibration of magnetometer and attitude determination system for land-survey mini-satellite. In: IEEE metrology for aerospace, pp 277–282

Research on Acoustic Velocity Correction Algorithm in Underwater Acoustic Positioning

Bingsheng Zhang, Tianhe Xu and Ruru Gao

Abstract With the development of national marine strategic plan, the demand of underwater high-precision positioning is increasing greatly. Acoustic wave, as an information carrier, plays an important role in underwater high-precision positioning. In this paper, the method of determining sound velocity in seawater and several classical models for calculating sound velocity are introduced. The indirect method is used to calculate the sound velocity of sea water, that is, the sound speed is calculated by the sound velocity empirical model according to the depth, temperature and salinity. Then, several classical sound speed models are used to select the optimal model under different computational conditions, and the most suitable model for the study sea area is obtained. Calculation of sound velocity profile based on the real time observation data of Argo measured by the official website combined with the optimal classical sound velocity model. The sound velocity is corrected by the weighted average sound velocity method and the layered ray tracing method, and the depth of the sea is recovered. Results show that the accuracy of tracking method is higher than that of the weighted average sound velocity method, which can achieve decimeter level precision in depth. The error between the calculated depth and the actual value by using the weighted average sound velocity method is large.

Keywords Sound velocity correction · Underwater acoustic positioning
Ray tracing · Weighted average

B. Zhang · R. Gao
School of Geology Engineering and Surveying, Chang'an University,
No.126, Yanta Road, Xi'an, Shanxi, China
e-mail: 1716892458@qq.com

T. Xu (✉)
Institute of Space Science, Shandong University, Weihai, Shandong, China
e-mail: thxugfz@163.com

T. Xu
State Key Laboratory of Geo-information Engineering, Xi'an, Shanxi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_72

1 Introduction

With the rapid development of the social economy and the rapid growth of the population, the land resources are scarce, and the exploration and exploitation of resources are extend into the ocean. Therefore, the demand for high precision underwater positioning is increasing [1]. The most important location parameter of underwater acoustic positioning technology is the sound speed. The sound velocity in sea water is influenced by many factors, such as temperature, salinity, pressure and so on. The change of temperature in sea water has the greatest influence on the sound velocity [2]. There are two ways to obtain the velocity profile of the sea water: one can be measured directly by the sound velocity profiler, and the other is to calculate the sound velocity according to the empirical model of the sound velocity [3]. The direct measurement method is of high precision, but the general price is expensive [4, 5]. In marine applications, we often use indirect methods. Based on the measured data of buoys in China's Argo real-time data center, we analyze and calculate the sound velocity profile. We use the weighted average sound velocity method and the stratified ray tracing method to correct the sound velocity in seawater [6], and calculate the depth of seawater and analyze the accuracy.

2 The Empirical Model of the Sound Velocity

Sound velocity is an important parameter in ocean positioning, which is a function of temperature, salinity, depth or pressure. Since 1950s, scholars have put forward different models of sound speed experience. The formula is as follows [1]:

1. Dell Grosso formula

$$\begin{aligned} C &= C' + C_p \\ C' &= 1448.6 + 4.618T - 0.0523T^2 + 0.00023T^3 \\ &\quad + 1.25(S-35) - 0.11(S-35)T + 2.7\times 10^{-8}(S-35)T^4 \\ &\quad - 2\times 10^{-7}(S-34)^4(1+0.577T-0.0072T^2) \\ C_p &= 0.160518D + 1.0279\times 10^{-7}D^2 \\ &\quad + 3.451\times 10^{-12}D^3 - 3.503\times 10^{-16}D^4 \end{aligned} \tag{1}$$

Although the model is simple and convenient, the application range is too small.

2. W.D. Wilson formula

$$C = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.01T)(S - 35) + 0.016D \quad (2)$$

The formula is suitable for temperature between −4 and 30 °C, and the suitable salinity is between 0 and 37. This formula is suitable for the China's sea area.

3. Leroy formula [7]

$$C = 1492.9 + 3(T - 10) - 0.006(T - 10)^2 - 0.04(T - 18)^2 + 1.2(S - 35) - 0.01(S - 35)(T - 18) + D/61 \quad (3)$$

The formula is suitable for temperature between −2 and 34 °C, and the suitable salinity is between 20 and 42, the suitable depth is between 0 and 8000 m.

4. Medwin formula

$$C = 1449.2 + 4.6T - 5.5 \times 10^{-2}T^2 + 2.9 \times 10^{-4}T^3 + (1.34 - 0.017)(S - 35) + 0.016D \quad (4)$$

The formula is suitable for temperature between 0 and 35 °C, and the suitable salinity is between 0 and 45, the suitable depth is between 0 and 1000 m.

5. Mackenzie formula [8]

$$C = 1448.96 + 4.591T - 5.304 \times 10^{-2}T^2 + 2.374 \times 10^{-4}T^3 + 1.34 \times (S - 35) + 1.63 \times 10^{-2}D + 1.675 \times 10^{-7}D^2 - 1.025 \times 10^{-2}T(S - 35) - 7.139 \times 10^{-13}TD^3 \quad (5)$$

The formula is suitable for temperature between −2 and 30 °C, and the suitable salinity is between 25 and 40, the suitable depth is between 0 and 8000 m.

6. EM formula

The surface sound velocity model:

$$C^1 = 1449.05 + T(4.57 - T(0.0521 - 0.00023T)) + (1.333 - T(0.0126 - 0.00009T))(S - 35) \quad (6)$$

In the deep water:

$$C = C^1 + 16.5D \tag{7}$$

where T is the temperature with unit of centigrade, the S is the salinity with unit of ppt, and the D is the depth with unit of meter.

3 Analysis and Calculation of Sound Velocity Profile

3.1 The Brief Introduction of Argo

Argo global ocean observation network is a large-scale ocean observation program put forward by scientists in the field of oceanography, atmosphere and other related fields in the early twentieth Century. To be more accurate, more rapid and large scale global upper ocean water temperature and salinity profile measurement should be collected, in order to improve the precision of weather forecast, effectively prevent the worsening global climate disasters (such as hurricanes, floods and hail, drought and other weather disasters) the threat of human [9].

Since China began to implement the Argo program in early 2002, 390 Argo profile buoys have been put in the Pacific and India oceans. 101 buoys are still working at the sea. These buoys can provide millions of data of sea water temperature and salinity profiles each year [10, 11]. The Argo buoy collects vertical profiles of temperature, salinity, and water depth in the range of 0–2000 m [12–14]. The following is the type and regional distribution of Chinese buoy (Fig. 1).

3.2 Calculation of Sound Velocity Profile

In this paper, the data of No. 2902712 buoy in October 31, 2017 is taken as an example. The Argo data has a fixed data format, and the velocity profile of the Argo buoy measurement value is calculated. The calculation results are shown in Fig. 2. Every sound velocity empirical formula has its certain range of use [15]. When the depth is greater than 1000 m, there are more data beyond the range of salinity or the range of temperature. Therefore, this paper only draws a velocity profile of 0–1000 m in depth.

From Fig. 2, it can be seen that the depth is within the range of 0–200 m from the measured buoy data, In addition to the Dell Grosso model, the sound velocity of the other models remains almost the same. As the depth increases, the water temperature decreases and the sound velocity decreases. The sound velocity of the EM stratified simplified model is slightly lower than those of other 4 models. The

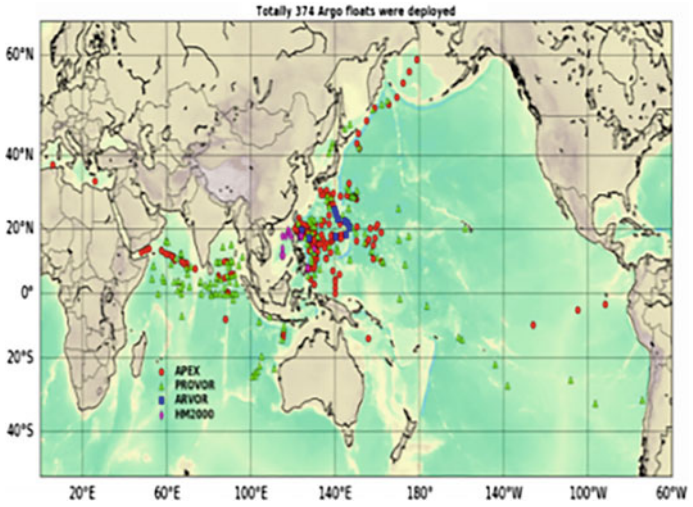


Fig. 1 Types and distribution of buoys in China

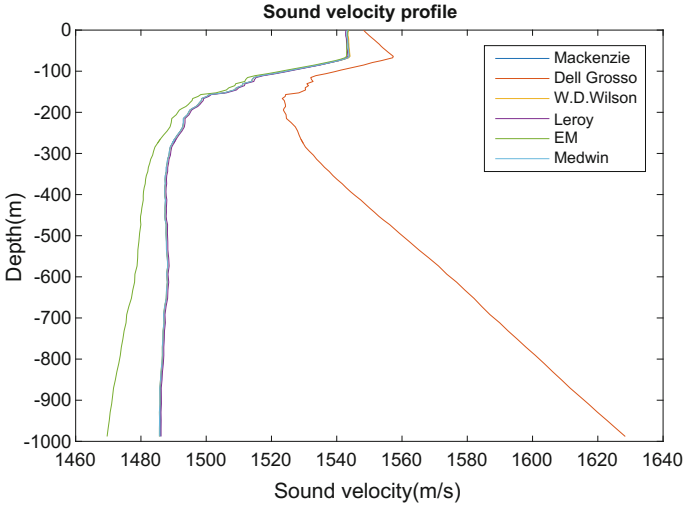


Fig. 2 Calculated values of different sound velocity models in October 31, 2017

Dell Grosso model has a relatively large deviation value and may not be applicable to the current sea area because of its smaller scope of application. Other models have high computational similarity, which can be used for the calculation of sound velocity profiles in the current sea area.

To further verify the correctness of the above, a set of simulated data is used to test the results compared with the above conclusions. In the applicability of the 6 models, when the depth is 0–1000 m, the calculation temperature is 10, 20 and 30 °C, respectively. During the calculation, the salinity is 30, 32, 34 and 36 ppt respectively. Tables 1 and 2 shows the results of the sound velocity calculation when the depth is 50 m and the depth is 800 m.

The unit of calculation results in the table: m/s, the value of s in the table is the value of salinity, and the unit is ppt (Figs. 3 and 4).

The sound velocity calculation results of the two table and the above pictures show that the sound velocity of the Dell Grosso model is larger than that of the other models when the depth of water is larger. The correctness of the Argo buoy data is verified. The results show that the Dell Grosso model is not suitable for the current research area, and the other models are more suitable.

Table 1 Calculation results of 6 kinds of sound velocity when the water depth value is 50 m

Temperature (°C)	Model type	Salinity			
		30s	32s	34s	36s
10	Dell Grosso	1484.9	1487.2	1489.4	1491.7
	W.D. Wilson	1484.6	1487.1	1489.6	1492.0
	Leroy	1483.9	1486.5	1489.1	1491.6
	Mackenzie	1484.4	1486.9	1489.4	1491.9
	EM	1483.7	1486.1	1488.6	1491.0
	Medwin	1484.2	1486.8	1489.5	1492.1
20	Dell Grosso	1517.2	1520.3	1521.8	1524.1
	W.D. Wilson	1516.6	1518.9	1521.2	1523.5
	Leroy	1517.1	1519.4	1521.8	1524.1
	Mackenzie	1515.8	1518.1	1520.3	1522.6
	EM	1516.3	1518.3	1521.1	1523.4
	Medwin	1516.6	1518.9	1521.2	1523.5
30	Dell Grosso	1564.4	1560.4	1556.3	1552.3
	W.D. Wilson	1541.1	1543.2	1545.2	1547.4
	Leroy	1540.2	1542.3	1544.5	1546.6
	Mackenzie	1541.0	1543.1	1545.1	1547.2
	EM	1541.1	1543.2	1545.3	1547.3
	Medwin	1539.7	1542.3	1545.0	1547.7

Table 2 Calculation results of 6 kinds of sound velocity when the water depth value is 800 m

Temperature (°C)	Model type	Salinity			
		30s	32s	34s	36s
10	Dell Grosso	1617.5	1617.8	1618.1	1618.4
	W.D. Wilson	1496.6	1499.0	1501.8	1504.1
	Leroy	1497.1	1499.6	1502.1	1504.7
	Mackenzie	1496.8	1499.2	1501.7	1504.2
	EM	1496.9	1499.3	1501.7	1504.2
	Medwin	1496.2	1498.8	1501.5	1504.1
20	Dell Grosso	1655.0	1653.2	1651.3	1649.4
	W.D. Wilson	1528.6	1530.9	1533.2	1535.5
	Leroy	1529.4	1531.7	1534.1	1536.4
	Mackenzie	1528.4	1531.2	1534.1	1536.1
	EM	1528.3	1531.3	1533.1	1535.9
	Medwin	1527.7	1530.3	1533.2	1535.5
30	Dell Grosso	1684.9	1680.8	1676.8	1672.7
	W.D. Wilson	1553.1	1555.2	1557.3	1559.3
	Leroy	1552.5	1554.6	1556.8	1558.9
	Mackenzie	1553.3	1555.4	1557.5	1559.5
	EM	1553.5	1555.6	1557.6	1559.7
	Medwin	1551.7	1554.3	1557.0	1559.6

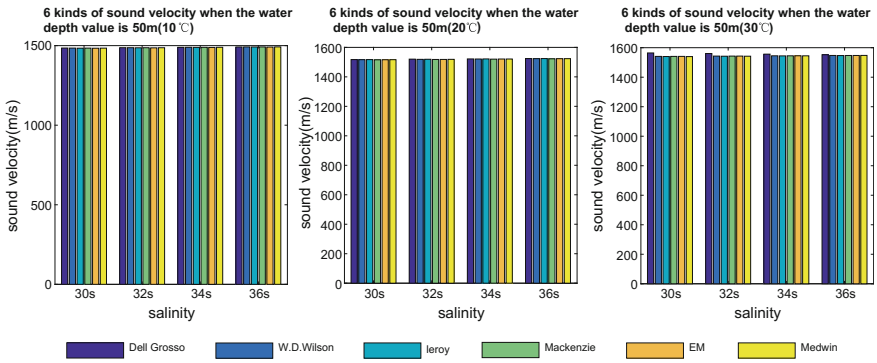


Fig. 3 6 kinds of sound velocity when the water depth value is 50 m (10, 20, 30 °C)

3.3 The Sound Velocity Tracking Method

Because sea water is not a homogeneous medium, sound waves do not propagate along the straight line in seawater. If we use fixed sound velocity to calculate, there will be large errors. Therefore, it is necessary to study high accuracy acoustic line tracking. Acoustic line tracking is usually implemented by layer by layer method.

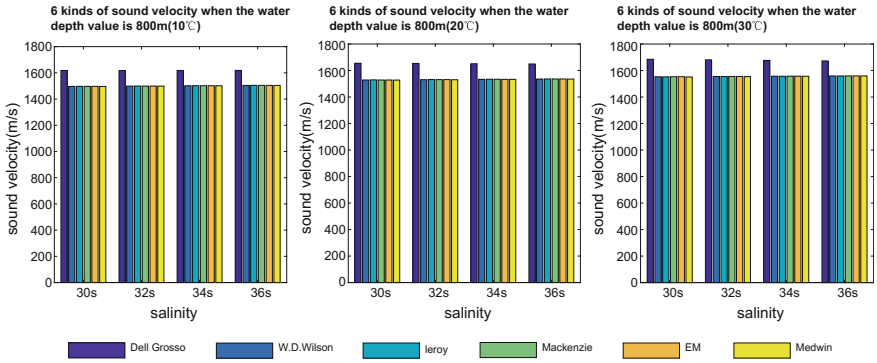


Fig. 4 6 kinds of sound velocity when the water depth value is 800 m (10, 20, 30 °C)

The two adjacent sound velocity points in the sound velocity profile are regarded as a layer, and the change of the class velocity can be assumed as constant or constant gradient. The calculation process of the former is simple and the latter has high accuracy [16].

3.3.1 The Constant Sound Speed Ray-Tracing Method

Assuming that the beam passes through N water layers during the propagation process, the sound velocity in each water layer is constant, and the schematic diagram is shown in Fig. 5.

From the Fig. 5, according to the Snell rule, we can get

$$\sin \theta_i = pC_i \tag{8}$$

The formula for calculating the horizontal displacement y_i of the beam in the layer i and the propagation time of this layer.

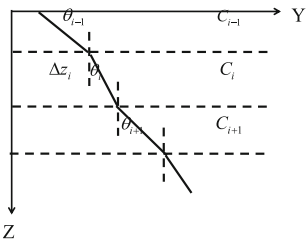


Fig. 5 The principle of constant sound speed ray-tracing method

$$y_i = \Delta z_i \tan\theta_i = \frac{\sin\theta_i \Delta z_i}{\cos\theta_i} = \frac{pC_i\Delta z_i}{(1-(pC_i)^2)^{1/2}}$$
$$t_i = \frac{y_i/\sin\theta_i}{C_i} = \frac{\Delta z_i}{C_i(1-(pC_i)^2)^{1/2}} \tag{9}$$

If the water has many layers, when the water layer is thinner, each layer of the acoustic velocity will not change much, assuming that each layer velocity gradient or set velocity constant, then the final calculation results and actual results are similar, will not lead to large deviation.

3.3.2 The Constant Gradient Ray-Tracing Method

It is assumed that sound waves undergo N equal gradient water, that is, the propagation velocity of sound waves in each layer is equal to the gradient. The sound velocity gradient g_i in layer i can be expressed in the following formula:

$$g_i = (C_{i+1} - C_i)/\Delta z_i \tag{10}$$

where Δz_i is the depth of the layer i, and C_{i+1} is the sound velocity of the i + 1 layer. Because the propagation of underwater sound satisfies the Snell rule, under the constant gradient of sound speed, the actual propagation path of the beam in the i level is a R_i with a certain radius of curvature and a continuous arc [17]. The formula is as follows:

$$R_i = -1/pg_i \tag{11}$$

The horizontal displacement of acoustic line in layer i can be expressed as [1, 18]:

$$y_i = R_i(\cos\theta_{i+1} - \cos\theta_i) = \frac{\cos\theta_i - \cos\theta_{i+1}}{pg_i} \tag{12}$$

where θ_i is the angle of incidence of the sound line of layer i, and θ_{i+1} is the angle of incidence of the i + 1 layer.

The length of the arc of the beam propagating on this layer is:

$$S_i = R_i(\theta_i - \theta_{i+1}) \tag{13}$$

The time of this section is:

$$t_i = m\ln[1 + \frac{g_i \Delta z_i}{C_i}]$$
$$m = \frac{\arcsin[p(C_i + g_i \Delta z_i)] - \arcsin(pC_i)}{pg_i^2 \Delta z_i} \tag{14}$$

The path of this method is in good agreement with the real acoustic line, and the accuracy of the calculation error is high, but when the data layer is more, the computation is large.

3.4 The Weighted Average Sound Velocity Method

The data of the sound velocity profile is stratified in the vertical direction, and a weighted mean value is required to be calculated by the sound velocity data of the discrete layers. A sound velocity is used in the whole water area, and the specific steps are as follows [19]:

Step 1: Fit the sound velocity distribution curve according to the measured sound velocity at different depths;

Step 2: The velocity curves are stratified along the depth so that the sound velocity varies linearly with depth in each layer;

Step 3: The mean sound velocity is calculated as [20]:

$$S = \frac{1}{H}\sum_{i=0}^{N-1} \frac{(C_i + C_{i+1})(Z_i - Z_{i+1})}{2} \tag{15}$$

where H is the water depth, C_i and C_{i+1} are the sound velocity of the layer i and $i + 1$ respectively, Z_i and Z_{i+1} are the depths of layers i and $i + 1$ respectively.

This method is simple, convenient in calculation and high in efficiency, but there is still a large residual error (Fig. 6).

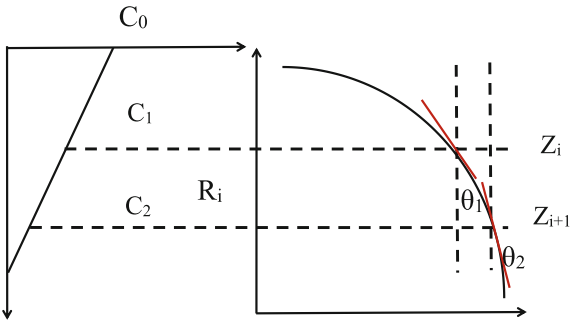


Fig. 6 The principle of constant gradient ray-tracing method

4 Example Analysis

According to the No. 2902712 buoy (130.02° longitude, latitude 7.48°) data (October 31, 2017) and the No. 2902744 buoy (111.81° longitude, latitude 12.25°) data (November 22, 2017), the water depth of about 1000 m can be calculated (sound velocity profile in the case of W.D. Wilson model), the results are shown in Figs. 7 and 8.

From Fig. 7, we can see that the sound velocity presents a positive gradient in the depth 0–70 m range, while the depth is in the range of 70–1000 m, showing a negative gradient, the maximum value of sound speed is about 1545 m/s, and the minimum value is about 1485 m/s.

From Fig. 8, we can see that the velocity of sound is a positive gradient in the range of 0–50 m, while the depth is in the range of 50–1000 m, showing a negative gradient, the maximum value of sound speed is about 1541 m/s, and the minimum value is about 1483 m/s.

Sound line will appear bending phenomenon in seawater, ray bending ultimate result is the measured value of water depth, which is very different from the actual value of depth, this phenomenon must be corrected. The sound ray bending is due to refraction of sound wave propagation in water is caused, if the ray tracking can determine the exact depth of the position, all this must be based on the correct calculation of sound velocity, the correct calculation depends on the sound velocity profile.

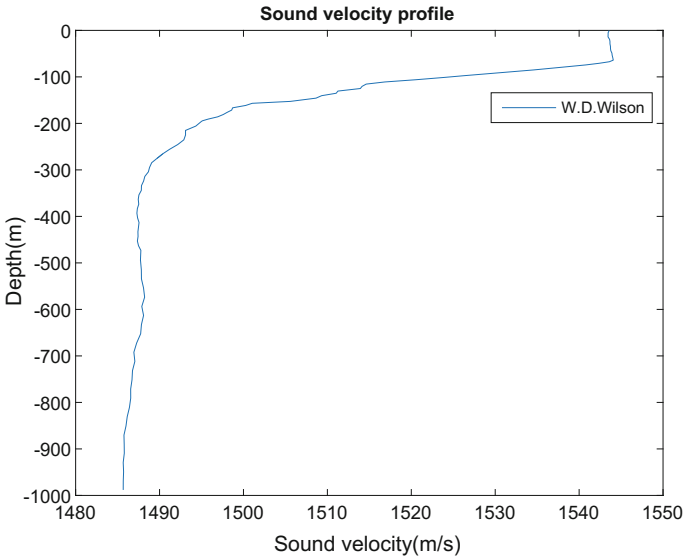


Fig. 7 Sound velocity profile data of No. 2902712 buoy

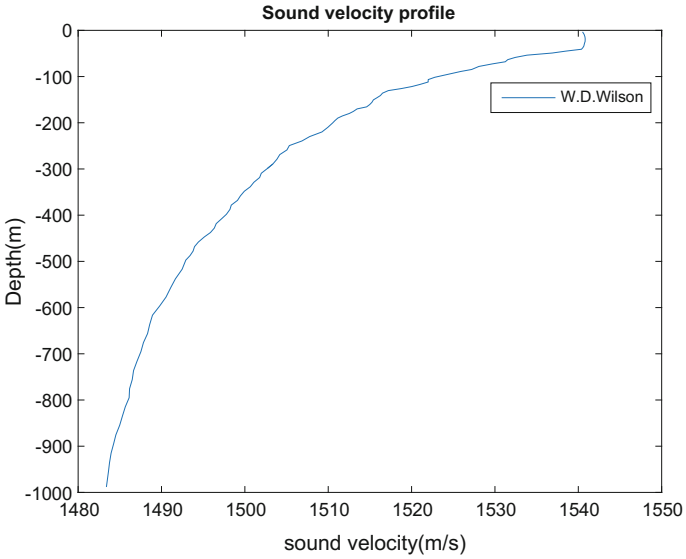


Fig. 8 Sound velocity profile data of No. 2902744 buoy

The sound velocity profile is stratified to get the measurement time t (round-trip) at a certain incidence angle and depth (the actual measurement depth is less than 1000 m), and then the sound velocity distribution is changed to 50 m level. Assuming that a certain initial incidence angle is incident, the beam propagation time of every 50 m water layer is subtracted from t/2 until the time difference between the remaining time is less than that of the beam propagating in the layer. In practical engineering, we take the average gradient of the first point and the last point of each layer as the sound velocity gradient of the layer, and calculate the propagation time of the sound wave in the time interval according to the constant sound velocity gradient. In the end, because the rest time is not enough, the sound line passes through a whole layer. Therefore, when dealing with this layer, the average sound velocity of the layer is taken as constant sound velocity to get the vertical distance that the sound line passes through in this time. The exact depth of the sea water is obtained by adding the vertical distance of the sound line at each layer.

The accuracy of the water depth calculated by the sound velocity tracking method and the weighted mean sound velocity method under different sound velocity profiles is compared and analyzed.

① Set angle theta to 1.5°, under the sound velocity profile of Fig. 7, a number of water depth values are calculated by using the sound velocity tracking method and the weighted average method. The results are shown in Tables 3 and 4.

Table 3 The value of water depth calculated by the sound velocity tracking method

Points	Actual depth (m)	Sound velocity tracking method (m)	Deviation (m)
1	85.1	84.95	−0.15
2	185.9	185.57	−0.33
3	275.2	275.11	−0.09
4	393.1	393.06	−0.04
5	473.1	473.07	−0.03
6	573.7	573.68	−0.02
7	692.1	692.12	0.02
8	771.2	771.21	0.01
9	889.9	889.98	0.08
10	988.2	988.31	0.11

Table 4 The value of water depth calculated by the weighted average sound velocity method

Points	Actual depth (m)	Weighted average sound velocity method (m)	Deviation (m)
1	85.1	82.15	−2.95
2	185.9	181.84	−4.06
3	275.2	271.23	−3.97
4	393.1	389.65	−3.45
5	473.1	470.03	−3.07
6	573.7	571.08	−2.62
7	692.1	690.02	−2.08
8	771.2	769.53	−1.67
9	889.9	888.87	−1.03
10	988.2	986.79	−1.41

It can be seen from Tables 3 and 4 that the difference of the depth calculated by the sound velocity tracking method and the actual value is small, with decimeter or even centimeter level precision. The depth of the weighted mean sound velocity method is far from the actual value, and the precision is not so high.

② Set angle theta to 1.5°, under the sound velocity profile of Fig. 8, The results are shown in Tables 5 and 6.

From Tables 5 and 6, the same conclusion as ① can also be drawn.

It can be concluded that the value of water depth calculated by the sound velocity tracking method is smaller than the actual value, and the accuracy is higher than that of the weighted mean sound speed method from ① and ②.

Table 5 The value of water depth calculated by the sound velocity tracking method

Points	Actual depth (m)	Sound velocity tracking method (m)	Deviation (m)
1	84.7	84.65	−0.05
2	179.9	179.92	0.02
3	289.2	289.24	0.04
4	387.5	387.62	0.12
5	497.2	497.37	0.17
6	596.6	596.67	0.07
7	695.3	695.41	0.11
8	775.5	775.74	0.24
9	894.2	894.42	0.22
10	988.3	988.45	0.15

Table 6 The value of water depth calculated by the weighted average sound velocity method

Points	Actual depth (m)	Weighted average sound velocity method (m)	Deviation (m)
1	84.7	79.17	−5.53
2	179.9	172.82	−7.08
3	289.2	281.11	−8.09
4	387.5	378.95	−8.55
5	497.2	488.59	−8.61
6	596.6	588.07	−8.53
7	695.3	687.11	−8.19
8	775.5	767.76	−7.74
9	894.2	887.11	−7.09
10	988.3	981.96	−6.34

5 Conclusions

In this paper, several classic sound speed empirical models are selected for the optimal model under different conditions, and an optimal model suitable for the area is obtained. On the basis of the algorithm of sound speed error correction, combined with the most suitable sound velocity model, the data of buoy 2902712 in October 31, 2017 and the 2902744 buoy data in November 22, 2017 were used, and the sound velocity profiles were obtained respectively. The water depth is calculated using weighted average velocity method and ray tracing method, respectively. The results show that the method of tracking velocity in the depth direction can reach decimeter accuracy. The accuracy of the weighted mean sound velocity method is not so high, and it is far from the actual value.

Acknowledgements The research was supported by National Key Research and Development Program of China (2016YFB0501701) and National Natural Science Foundation of China (Grant No. 41574013, 41731069).

References

1. Zhao J, Liu J (2008) Multi-beam sounding and image data processing. Wuhan University Press, Wuhan
2. Zhang H (2010) Underwater navigation and positioning technology. Wuhan University Press, Wuhan
3. Sun G (2007) Research and application of sound speed correction method for sun leather multi-beam sounding system. Ocean University of China
4. Geng X (1997) Precise acoustic bathymetry. Master's thesis, Department of Electrical and Computer Engineering, University of Victoria, Victoria, B.C
5. Vincent HT, Hu SLJ (2000) Method and system for determining underwater effective sound velocity. United States patent application, U.S. Navy Case Number: 78160
6. Guan Z (1990) Accuracy analysis of sound velocity correction formula in China. Oceanography 3:32–37
7. Zhu X (2008) Research on three-dimensional location simulation of deep sea. Chinese Academy of Surveying and Mapping Science, Beijing
8. Zhou F, Zhao J, Zhou C (2001) Determination of classic experiential sound speed formulae in multibeam echo sounding system. J Oceanogr Taiwan Strait 20(4):411–419
9. Xu J, Zhu B (2000) Argo global ocean observation network and the development of China's marine monitoring technology. Mar Technol 20(2):15–17
10. Vaas AE (1964) Refraction of the direct monotonic sound ray. Technical memorandum 322, Naval Underwater Ordnance Station, Newport
11. Zhang S (2010) Spatial distribution and change analysis of the number of argo observation points. Mar Technol 20:34–40
12. Wu B, Wang S (2014) Application range analysis of sound velocity empirical formula. Acoust Technol 12:504–507
13. Wen H, Li H, Cai Y, Cheng P, Zhu G (2012) Study on sea level change based on argo buoy, satellite altimetry and GRACE data. J Surveying Mapp 05:696–702
14. Yang F, Lu X, Li J et al (2011) Precise positioning of underwater static Mobjects without sound speed profile. Mar Geodesy 34(2):138–151
15. Zheng T, Zhou Y, Bian S (2009) Data processing and mapping of multi-beam bathymetry. Acta Marinea Sinica 06:112–117
16. Zhao J (2017) Modern marine surveying and mapping (1). Wuhan University Press, Wuhan
17. Lu X, Huang M (2012) Algorithm for improving average sound speed in constant gradient acoustic line tracking [J]. Journal of Wuhan University: InfSci Ed 37(5):590–593
18. Lan H (2007) Research on deep-sea underwater acoustic transponder positioning and navigation technology. Harbin Engineering University, Harbin
19. Yi C (2009) Error analysis of two water depth positioning system. Pet Geophys Explor 44 (2):136–139
20. Nie Z, Wang Z, Li S (2015) Comparison of methods for acoustic velocity correction in underwater acoustic positioning. Ocean Bull 04:423–427

Digital Beamforming Correction Algorithm Under the Matching Error of Space-Frequency Steering Vector

Chao Ren and Yongxiang Zheng

Abstract In real application, errors exist in environment, signal source and array sensors. The mismatch between the assuming steering vector of the desired signal and the real one will make the array beamformer to suppress the desired signal as interference. For the problem of beamforming performance degradation due to the signal matching errors, a new parallel digital multi-beam forming algorithm under the Space-frequency wide-band constraint is proposed. Taking the uniform circular array (UCA) as an example, the solution to the multi-beam forming algorithm is deduced by paralleling the single-star constraint of uncertainty set of the space-frequency vector. At the same time, solving process of the optimal weight is given. The algorithm effectively improves the output signal-to-noise ratio (SNR) of the digital multi-beam former. The algorithm is verified by Computer simulations, demonstrating its robustness.

Keywords Space-frequency constraint · Digital multi-beam forming
Steering vector

1 Introduction

As we known, digital beamforming technique has broad prospects in radar, digital communication, sonar and many other fields. However, in real application, errors exist in environment, signal source and array sensors [1, 2]. The mismatch between the assuming steering vector of the desired signal and the real one will make the array beamformer to suppress the desired signal as interference, therefore the phenomenon is called the self-cancellation of signal. The traditional Capon beamforming algorithm is hard to obtain the ideal effect in real application due to the sensitivity to the errors of the steering vector. In order to solve the problem of

C. Ren (✉) · Y. Zheng
Unicore Communications, Inc., Fengxiandonglu Road, Haidian District,
Beijing 100081, China
e-mail: chaoren@navchina.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_73

the mismatch between the assuming steering vector and the real one, in recent years, a class of robust algorithms based on the uncertainty set of the steering vector have been proposed one after another.

The robust MVDR beamforming algorithm under the constraint of spherical uncertainty set guarantee the minimum expecting response of the array is no less than 1, at the same time constraining the signal steering vector being located in the spherical uncertainty set [3]. Jian Li and his team had proposed a robust Capon beamforming algorithm based on the constraint of uncertain set of steering vectors [4]. In this paper, taking the UCA as an example, a new parallel digital multi-beam forming algorithm under the Space-frequency wide-band constraint is proposed, based on robust Capon beamforming algorithm. By paralleling multiple single-star constraint of uncertainty set of the space-frequency vector, desired signal in each frequency sub-band which is constrained is almost not lost. Then the solution to the multi-beam forming algorithm is given, as well as the solving process of the optimal weight. The proposed algorithm robustly processes space-frequency steering vector with random error, effectively improving the output SNR of the digital multi-beam former. The algorithm is verified by computer simulations, demonstrating its robustness.

2 Space-Frequency Beamformer

Consider a uniform circular array with M elements, locating at the origin of the coordinates, as shown in Fig. 1. To make it simple, electromagnetic signals, propagation media, array states and environments can be idealized.

Supposing a plane wave $s(t)$ is coming from the direction of (θ, φ) towards the UCA with the radius $R = \lambda/2$, where λ is the wavelength of the carrier signal. The spherical coordinate system is used to represent the arriving direction of the plane

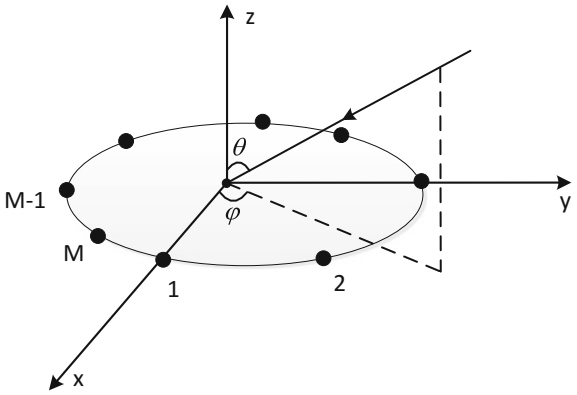


Fig. 1 The uniform circular array

wave. According to the phase relationship of each element relative to the coordinate origin, the space steering vector of the array can be obtained as

$$\boldsymbol{S}_a = \left[e^{j\xi\cos(\varphi - r_1)},\ e^{j\xi\cos(\varphi - r_2)},\ \ldots,\ e^{j\xi\cos(\varphi - r_m)} \right]^{\mathrm{T}} \tag{1}$$

where $\xi = \frac{2\pi R}{\lambda}\sin\theta$, $r_m = \frac{2\pi}{M}m$, $(m = 1, 2, \ldots, M)$. The array response of the signal can be expressed as $S_a s(t)$.

Supposing there exist N wide-band plane wave $s_i(t)$, $i = 1, 2, \ldots N$, then the response of the array can be expressed as

$$\boldsymbol{X}(t) = \sum_{i=1}^{N} \boldsymbol{S}_{ai} s_i(t) + \boldsymbol{N}(t) \tag{2}$$

$$\boldsymbol{N}(t) = [n_1(t), n_2(t), \ldots\ n_M(t)]^{\mathrm{T}} \quad m = 1, 2, \ldots\ M \tag{3}$$

where $\boldsymbol{s_{ai}}$ represents the space steering vector of signal $s_i(t)$ coming from (θ_i, φ_i), and $\boldsymbol{N}(t)$ represents the Gaussian white noise, $n_m(t)$ represents the thermal noise of the mth receiving channel.

For wide-band signal, the structure of a space-frequency beamformer is shown in Fig. 2. Space-frequency beamformer is a suboptimal choice relative to the space-time beamformer, but it has advantage in reducing the dimensionality of matrix. Discrete Fourier transform (DFT) is performed on the delayed signal of each element, and then beamforming is performed separately in each frequency bins, finally the time-domain signal is restored by an Inverse Discrete Fourier transform (IDFT). Performing DFT of K points to the received signals, we can get the following frequency-domain model as [5, 6]

$$\boldsymbol{X}(f_k) = \boldsymbol{A}(f_k)\boldsymbol{S}(f_k) + \boldsymbol{N}(f_k) \quad k = 1, 2, \ldots, K \tag{4}$$

$$\boldsymbol{A}(f_k) = [\boldsymbol{S}_{a1}(\theta_1, \varphi_1, f_k) \quad \boldsymbol{S}_{a2}(\theta_2, \varphi_2, f_k) \quad \ldots \quad \boldsymbol{S}_{aN}(\theta_N, \varphi_N, f_k)]^{\mathrm{T}} \tag{5}$$

$$\boldsymbol{S}(f_k) = [\boldsymbol{S}_1(f_k) \quad \boldsymbol{S}_2(f_k) \quad \ldots \quad \boldsymbol{S}_N(f_k)]^{\mathrm{T}} \tag{6}$$

$$\boldsymbol{N}(f_k) = [\boldsymbol{N}_1(f_k) \quad \boldsymbol{N}_2(f_k) \quad \ldots \quad \boldsymbol{N}_M(f_k)]^{\mathrm{T}} \tag{7}$$

where $\boldsymbol{X}(f_k)$, $\boldsymbol{S}(f_k)$, and $\boldsymbol{N}(f_k)$ are the DFT of the received signals, the desired signal and the noise of the kth bin respectively, as shown in Fig. 2.

The received signals can also be expressed as

$$\boldsymbol{X}(f_k) = [x_{1k}, x_{2k}, \ldots, x_{Mk}]^{\mathrm{T}} \quad k = 1, 2, \ldots, K \tag{8}$$

Defining vector $\boldsymbol{w}(f_k)$ of dimension $M \times 1$ as the adaptive weight vectors for the kth bin

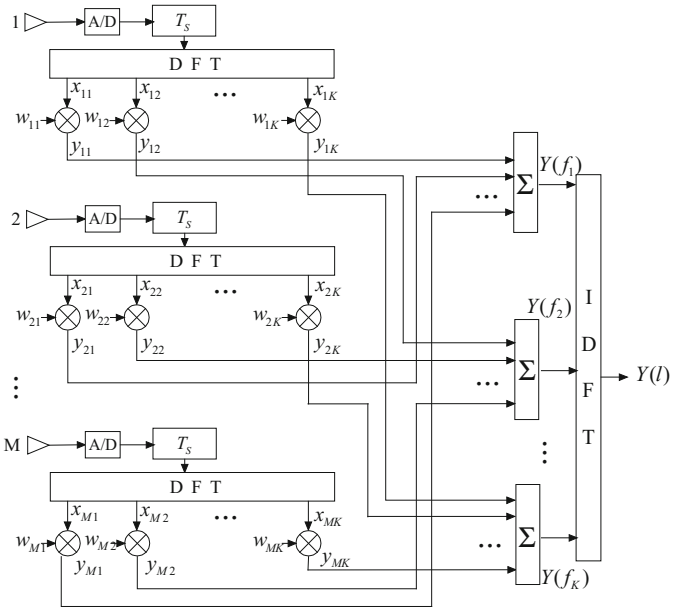


Fig. 2 Structure of space-frequency beamformer

$$\boldsymbol{w}(f_k) = [w_{1k}, w_{2k}, \ldots, w_{Mk}]^{\mathrm{T}} \quad k = 1, 2, \ldots, K \tag{9}$$

Then, the output signal of the space-frequency beamformer is

$$\boldsymbol{Y}(f_k) = \boldsymbol{w}^H(f_k)\boldsymbol{X}(f_k) = \sum_{m=1}^{M} y_{mk} \quad k = 1, 2, \ldots, K \tag{10}$$

3 Beamformer with Parallel Constraint of Space-Frequency Steering Vector

The standard Capon [4, 7] beamforming algorithm can be expressed as a linear constrained quadratic optimization problem

$$\begin{cases} \min\limits_{w} & \boldsymbol{w}^H\boldsymbol{R}\boldsymbol{w} \\ s.t. & \boldsymbol{w}^H\boldsymbol{a}_0 = 1 \end{cases} \tag{11}$$

where $\boldsymbol{a}_0$ is the steering vector of the desired signal, $\boldsymbol{R}$ is the covariance matrix,The optimal solution is

$$\boldsymbol{w}_{\text{Capon}} = \frac{\boldsymbol{R}^{-1}\boldsymbol{a}_0}{\boldsymbol{a}_0^H \boldsymbol{R}^{-1}\boldsymbol{a}_0} \tag{12}$$

The steering vector $\boldsymbol{a}_0$ usually cannot be accurately obtained in real application, therefore the assumed steering vector $\bar{\boldsymbol{a}}_0$ is used, restricting $\boldsymbol{a}_0$ within the uncertainty spherical as follow

$$\|\boldsymbol{a}_0 - \bar{\boldsymbol{a}}_0\|^2 \leq \varepsilon \tag{13}$$

where ε is the parameter of the spherical uncertainty set. Formula (11) can be transformed as a quadratic problem [8–10]

$$\begin{cases} \min\limits_{\hat{\boldsymbol{a}}_0} \quad \boldsymbol{a}_0^H \boldsymbol{R}^{-1}\boldsymbol{a}_0 \\ s.t. \quad \|\boldsymbol{a}_0 - \bar{\boldsymbol{a}}_0\|^2 \leq \varepsilon \end{cases} \tag{14}$$

Taking the space-frequency beamformer in Fig. 1 as a model, DFT is performed on the signals received from the M array elements. The received signal is divided into K bins in the frequency domain. Based on the position information of the signal source and the array, the direction of arrival(DOA) of the desired signals are estimated as $\{\theta_1, \varphi_1, \theta_2, \varphi_2, \ldots, \theta_N, \varphi_N,\}$. The space-frequency steering vector of the desired signal in the kth frequency bin can be obtained as $\boldsymbol{A}(f_k) = [\boldsymbol{S}_{a1}(\theta_1, \varphi_1, f_k) \quad \boldsymbol{S}_{a2}(\theta_2, \varphi_2, f_k) \quad \ldots \quad \boldsymbol{S}_{aN}(\theta_N, \varphi_N, f_k)]$.

In real application, the covariance matrix of the sampled data $\hat{\boldsymbol{R}}$ is usually used instead of $\boldsymbol{R}$. The sample covariance matrix of the kth frequency bin can be expressed as

$$\begin{aligned} \hat{\boldsymbol{R}}(f_k) &= \frac{1}{L} X(f_k) X^H(f_k) \\ &= \frac{1}{L} \begin{bmatrix} \sum\limits_{l=1}^{L} x_{1k}(l) x_{1k}^*(l) & \cdots & \sum\limits_{l=1}^{L} x_{1k}(l) x_{Mk}^*(l) \\ \cdots & \cdots & \cdots \\ \sum\limits_{l=1}^{L} x_{Mk}(l) x_{1k}^*(l) & \cdots & \sum\limits_{l=1}^{L} x_{Mk}(l) x_{Mk}^*(l) \end{bmatrix} \quad k = 1, 2, \ldots, K \end{aligned} \tag{15}$$

where L is the number of snapshots on frequency domain and $x_{mk}(l)$ represents the DFTed signal received by the mth array element, on the kth frequency bin of the lth snapshot.

The space-frequency constraint equation of the nth desired signal on the kth frequency bin is

$$\begin{cases} \min \quad E\{\left|Y^{(n)}(f_k)\right|^2\} = \boldsymbol{w}^{(n)}(f_k)^H \boldsymbol{R}(f_k) \boldsymbol{w}^{(n)}(f_k) \\ s.t. \quad \boldsymbol{S}_{an}^H(\theta_n, \varphi_n, f_k) \boldsymbol{w}^{(n)}(f_k) = 1 \end{cases} \tag{16}$$

Constraining the space-frequency steering vector of the nth desired signal within the spherical uncertainty set, the desired signal output power is maximized

$$\begin{cases} \max\limits_{S_a} & \frac{1}{\boldsymbol{S}_{an}^H(\theta_n,\varphi_n f_k)\hat{\boldsymbol{R}}^{-1}(f_k)\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)} \\ s.t. & (\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)-\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k))^H(\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)-\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k))\le\varepsilon_n \end{cases} \tag{17}$$

where $\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)$ Is the real space-frequency steering vector of the nth desired signal, $\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k)$ is an estimate of $\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)$, and ε_n is a parameter in constraint equation of the nth desired signal.

Equation (17) can also be rewritten as

$$\begin{cases} \min\limits_{S_a} & \boldsymbol{S}_{an}^H(\theta_n,\varphi_n,f_k)\hat{\boldsymbol{R}}^{-1}(f_k)\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k) \\ s.t. & \left\|\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)-\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k)\right\|^2\le\varepsilon_n \end{cases} \tag{18}$$

Equation (18) constrains the output SNR of one desired signal to be optimal. However, the estimated matching error of different signal in the same system is similar, we could set $\varepsilon=\max\{\varepsilon_1,\varepsilon_2,\ldots\varepsilon_N\}$, then the beamformer with parallel constraint of space-frequency steering vector is given as

$$\boldsymbol{Y}(f_k)=[\,y^{(1)}(f_k)\quad y^{(2)}(f_k)\quad\ldots\quad y^{(N)}(f_k)\,]^{\mathrm{T}}=\boldsymbol{W}^{\mathrm{T}}(f_k)\boldsymbol{X}(f_k) \tag{19}$$

$$\boldsymbol{W}(f_k)=[\,\boldsymbol{w}^{(1)}(f_k)\quad\boldsymbol{w}^{(2)}(f_k)\quad\ldots\quad\boldsymbol{w}^{(N)}(f_k)\,]^{\mathrm{T}} \tag{20}$$

$\boldsymbol{W}(f_k)$ can be solved by using the Lagrange multiplier methodology in the following way

$$\begin{aligned} L(\boldsymbol{S})=&\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)\hat{\boldsymbol{R}}^{-1}(f_k)\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)\\ &+\lambda\left(\left\|\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)-\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k)\right\|^2-\varepsilon\right) \end{aligned} \tag{21}$$

Then we get the optimal steering vector of the desired signals,

$$\begin{aligned} \boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)&=[\frac{\hat{\boldsymbol{R}}^{-1}(f_k)}{\lambda}+\boldsymbol{I}]^{-1}\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k)\\ &=\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k)-[\boldsymbol{I}+\lambda\hat{\boldsymbol{R}}(f_k)]^{-1}\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k) \end{aligned} \tag{22}$$

By replacing the estimated steering vector $\bar{\boldsymbol{S}}_{an}(\theta_n,\varphi_n,f_k)$ with the optimal one $\boldsymbol{S}_{an}(\theta_n,\varphi_n,f_k)$, we get this

$$\boldsymbol{w}^{(n)}(f_k) = \frac{[\hat{\boldsymbol{R}}(f_k) + \frac{1}{\lambda}\boldsymbol{I}]^{-1}\bar{\boldsymbol{S}}_{an}(\theta_n, \varphi_n, f_k)}{\bar{\boldsymbol{S}}_{an}^{H}(\theta_n, \varphi_n, f_k)[\hat{\boldsymbol{R}}(f_k) + \frac{1}{\lambda}\boldsymbol{I}]^{-1}\hat{\boldsymbol{R}}(f_k)[\hat{\boldsymbol{R}}(f_k) + \frac{1}{\lambda}\boldsymbol{I}]^{-1}\bar{\boldsymbol{S}}_{an}(\theta_n, \varphi_n, f_k)} \tag{23}$$

There are N constraint equations to the multi-beam former for N desired signals as follows

$$g(\lambda) = \left\|[\boldsymbol{I} + \lambda\hat{\boldsymbol{R}}(f_k)]^{-1}\bar{\boldsymbol{S}}_{an}(\theta_n, \varphi_n, f_k)\right\|^2 = \varepsilon \tag{24}$$

It's the same way to solve the loading value for N constraint equations, except that the covariance matrix is different in each frequency bin. First of all, decomposing the covariance matrix $\hat{\boldsymbol{R}}(f_k) = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H$, where $\boldsymbol{V}$ is the matrix of feature vector and $\boldsymbol{\Lambda}$ is a diagonal matrix. The eigen values are, $\gamma_{k1} \geq \gamma_{k2} \geq \ldots \gamma_{kM}$, respectively. Let $\boldsymbol{z}(f_k) = [z_{k1}, z_{k2}, \ldots, z_{kM}]^{\mathrm{T}} = \boldsymbol{V}^H\bar{\boldsymbol{S}}_{an}(\theta_n, \varphi_n, f_k)$, K constraint equations are given in K bins

$$g(\lambda) = \sum_{i=1}^{M}\frac{|z_{ki}|^2}{(1+\lambda\gamma_{ki})^2} = \varepsilon_k \quad k \in [1, K] \tag{25}$$

Set ε_{nk} be the matching error of space-frequency vector in the kth frequency bin of the nth desired signal. It is determined by the allowable angle offset $\Delta\theta$, which is generally within 10° in real application. ε_{nk} can be calculated by the given $\Delta\theta$ as $\varepsilon_{nk} = \left\|\boldsymbol{S}_{an}(\theta_n, \varphi_n, f_k) - \bar{\boldsymbol{S}}_{an}(\theta_n, \varphi_n, f_k)\right\|^2 \quad n \in [1, N]$, $k \in [1, K]$, set $\varepsilon = \max\{\varepsilon_{nk}\}$, we can get

$$G(\lambda) = \sum_{i=1}^{M}\frac{|z_{1i}|^2}{(1+\lambda\gamma_{1i})^2} + \ldots + \sum_{i=1}^{M}\frac{|z_{ki}|^2}{(1+\lambda\gamma_{ki})^2} + \ldots + \sum_{i=1}^{M}\frac{|z_{Ki}|^2}{(1+\lambda\gamma_{Ki})^2} = K\varepsilon \tag{26}$$

Set $\eta = \sum_{i=1}^{M}|z_{1i}|^2 \approx \sum_{i=1}^{M}|z_{Ki}|^2$, $\gamma_{\min} = \min\{\gamma_{kM}\}$, $\gamma_{\max} = \max\{\gamma_{k1}\}$. When $\lambda \geq 0$, we have $\eta/(1+\lambda\gamma_{\max})^2 \leq \varepsilon \leq \eta/(1+\lambda\gamma_{\min})^2$. When $\varepsilon \leq \eta$, there must be a unique solution λ to the constraint equation, where $\lambda \in \left((\sqrt{\eta/\varepsilon} - 1)/\gamma_{\max}, (\sqrt{\eta/\varepsilon} - 1)/\gamma_{\min}\right)$. When $\varepsilon = \eta$, we have $\lambda = 0$. When $\varepsilon > \eta$, we have $\lambda < 0$ satisfying the constraint equation. $G(\lambda)$ is a monotonically decreasing function, there must be a unique solution $\lambda \in (-1/\gamma_{\max}, 0)$. Therefore, the solution $\boldsymbol{W}(f_k)$ to the beamformer with parallel constraint of space-frequency steering vector can be determined. Using the formula (10) to process the space-frequency filtering, the frequency-domain matrix of each desired signal is obtained as $[\,y^{(n)}(f_1) \quad y^{(n)}(f_2) \quad \ldots \quad y^{(n)}(f_K)\,] \quad n = 1, 2, \ldots, N$. We can use the IDFT to get the time-domain signal. Although the algorithm adds K − 1 covariance matrix calculation, but it ensures that each satellite output SNR optimal.

4 Numerical Examples

A UCA of seven elements receiving eight satellite navigation signals is used to verify the algorithm. The DOA of the signals which are marked as '*' is shown in Fig. 3, and the parameters of the satellite signals are shown in Table 1.

Supposing that the matching error of the space-frequency steering vector is mainly caused by the estimated deviation of the DOA. The estimated deviation sets 5°. DFT of 512 points is performed on the received signals, and the parameter in Eq. (26) is chosen as $\varepsilon = 0.4695$, $\eta = 7$. The algorithm constrains the power of the satellite signals in frequency bins within signal bandwidth from loss. Figure 4 shows the output SNR of the proposed beamformer, compared with beamformer constraining only the fourth satellites and beamformer constraining multiple satellites.

Next, we are analyzing the beamforming performance. Considering a UCA of seven elements, with the expecting DOA of the satellite signal is (30°, 150°). The bandwidth of signal is 20 MHz, and the sampling rate is 62 MHz, the input SNR is −30 dB. Supposing the matching error of the space-frequency steering vector is 10°, DFT of 512 points is performed on the received signals. The duration of received signal is approximate 132 μs, with snapshot of 8192. The parameter ε is then calculated as 0.4695, Figs. 5 and 6 show the array pattern and contour map of the proposed beamformer respectively. It can be seen that the algorithm can effectively correct the matching error of the space-frequency steering vector.

Figure 7 shows the output SNR when the matching error changing from 0° to 10° using the standard Capon beamformer (SCB), the robust Capon beamformer (RCB) and the proposed one. Obviously, the proposed beamformer with parallel constraint of space-frequency steering vector improves the output SNR of the system. As the matching error increases, the beamformer can still maintain a certain robustness and the performance is obviously superior to the other two beamformer.

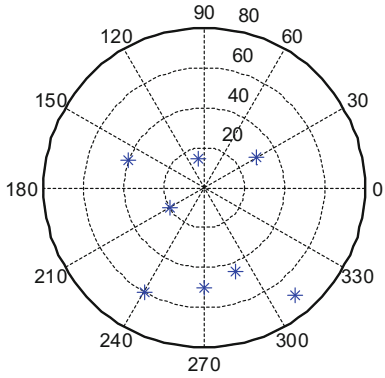


Fig. 3 The DOA of satellites

Table 1 The parameters of satellite signals

Satellite no.	DOA (θ, φ)	SNR (dB)	Doppler (KHz)
1	(30°, 30°)	−30	2.5
2	(15°, 100°)	−32	3.5
3	(40°, 160°)	−26	−4.5
4	(20°, 210°)	−28	4.5
5	(60°, 240°)	−30	−2.5
6	(50°, 270°)	−31	−0.5
7	(45°, 290°)	−29	1.5
8	(70°, 310°)	−29	0.5

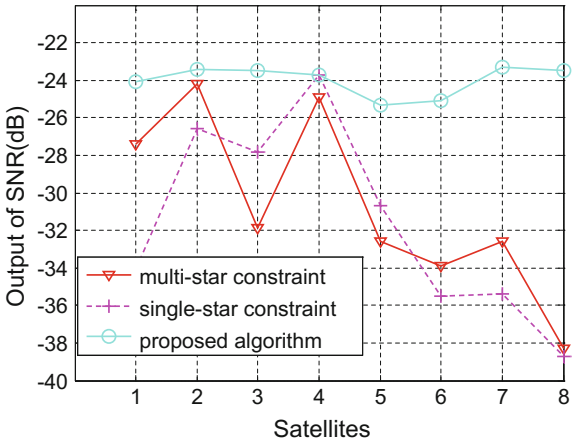


Fig. 4 Output SNR of different constraints

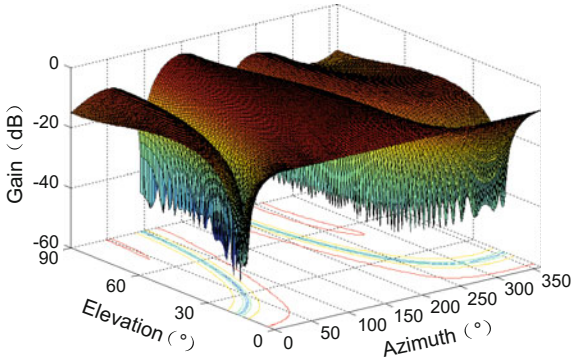


Fig. 5 Array pattern

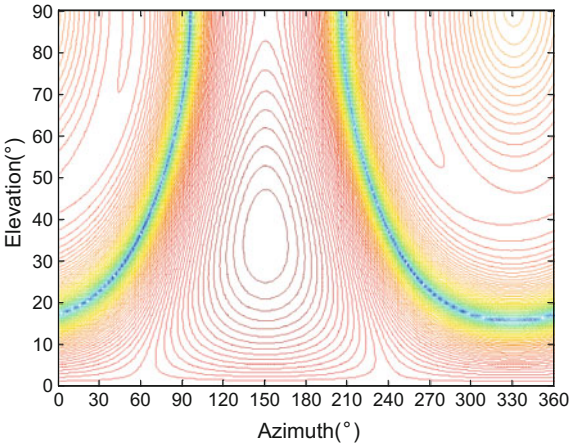


Fig. 6 Gain contour map

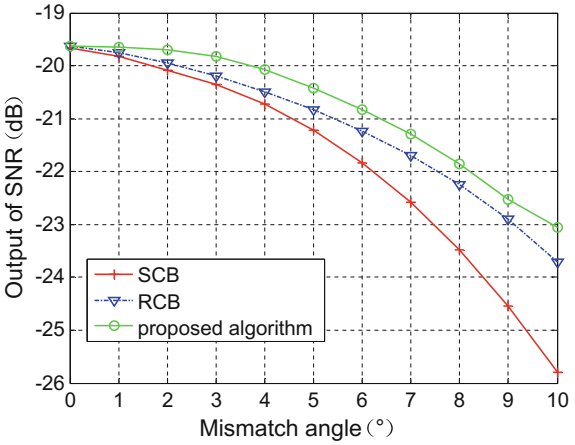


Fig. 7 Output SNR varies with matching error

5 Conclusions

In applications of adaptive antenna, there usually exists matching error between the expecting steering vector of the desired signal and the real one, which degrades the performance of the beamformer. Based on the wide-band signal model, a new parallel digital multi-beam forming algorithm under the Space-frequency wide-band constraint is proposed. By paralleling multiple single-star constraint of uncertainty set of the space-frequency vector, the solution to the constraint equation of is derived, meanwhile, the solving process of the optimal weight is given. The method improves the anti-jamming capability of the satellite navigation receiver and effectively improves the distortion of the satellite signal in the actual application environment.

References

1. Lee JH, Cheng KP, Wang CC (2006) Robust adaptive array beamforming under steering angle mismatch. Sig Process 86:296–309
2. Bao ZQ, Wu SJ, Zhang LR (2006) Fast and robust GSC beamformer based on variable diagonal loading. In: 2006 CIE international conference on radar, Oct 2006, pp 934–937
3. Vorobyov SA, Gershman AB, Luo ZQ (2003) Robust adaptive beamforming using worst case performance optimization: a solution to the signal mismatch problem. IEEE Trans Sig Process 51(2):313–323
4. Li J, Stotica P, Wang ZS (2003) On robust capon beamformer and diagonal loading. IEEE Trans Sig Process 51(7):1702–1715
5. Duan, ZL, Li YL, Xu SB (2015) Multiple signal detection based on spatial-frequency adaptive processing using fast subspace decomposition method. In: IEEE international conference on signal processing, 2015, pp 1–4
6. Gupta IJ, Moore TD (2004) Space-frequency adaptive processing (SFAP) for radio frequency interference mitigation in spread-spectrum receivers. IEEE Trans Antennas Propag 52 (6):1611–1615
7. Li J, Stoica P, Wang Z (2004) Doubly constrained robust Capon beamformer. IEEE Trans Sig Process 52:2407–2423
8. Elnashar A, Elnoubi S, Elmikati H (2005) Robust adaptive beamforming with variable diagonal loading. In: 2005 6th IEE international conference on 3G and beyond, Nov 2005, pp 1–5
9. Chen H, Zhao Y, Liu C (2014) A novel robust capon algorithm for wideband beamforming. In: IEEE radar conference, pp 950–953
10. Somasundaram SD, Parsons NH, Li P (2015) Reduced-dimension robust capon beamforming using Krylov-subspace techniques. IEEE Trans Aerosp Electron Syst 51(1):270–289

Analysis and Selection of Global Marine Gravity/Gravity Gradient Aided Navigation Matching Areas

Lin Wu, Hubiao Wang, Hua Chai, Lu Zhang, Houtse Hsu and Yong Wang

Abstract With the development of underwater passive navigation in recent years, gravity or gravity gradient matching navigation are utilized to correct the accumulated navigation errors of inertial navigation system. The accuracy and success rate of gravity/gravity gradient matching navigation are seriously affected by the features of gravity changes in the sailing area. In this paper, a parameter named average gravity difference (AGD) was introduced here to express the level of gravity changes in a certain area. With the overall survey of global marine gravity, the AGD values of local areas were calculated and analyzed. After a series of experiments, the relationship of AGD values and navigation accuracy in local areas can be generally determined. As the criterion of selection of matching areas was induced, marine gravity matching areas for navigation can be selected. Simulation results showed that, take gravity as an example, the AGD values can range from about 0.6–2.7 mGal/nautical mile (1 nautical mile = 1 n mile = 1852 m) in different areas of the oceans around the world. With present technologies, gravity matching navigation could achieve an accuracy of 0.5 n miles in matching areas which have AGD values larger than 2.0 mGal/n mile. By this criterion, South China Sea, North Indian Ocean and some other areas can be selected. In these areas, better navigation results can be obtained since they have rougher gravity changes. These results would be significant for the development of gravity aided navigation technologies and the construction of underwater PNT system.

Keywords Gravity aided navigation · Matching areas selection
Underwater navigation · Gravity matching · Gravity gradient

L. Wu (✉) · H. Wang · H. Chai · L. Zhang · H. Hsu · Y. Wang
State Key Laboratory of Geodesy and Earth's Dynamics,
Institute of Geodesy and Geophysics, Chinese Academy of Sciences,
Wuhan 430077, China
e-mail: linwu@whigg.ac.cn

L. Zhang · Y. Wang
University of Chinese Academy of Sciences, Beijing 100049, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_74

1 Introduction

In recent years, exploration and development of ocean are playing a more and more important role in several fields like economy, diplomacy, national strategy, and so on. Therefore the development and construction of underwater position, navigation and timing (PNT) system become a significant component in comprehensive PNT system of our country. Currently, during long distance and endurance underwater sailing, additional aided navigation methods are needed to correct the accumulated navigation errors of inertial navigation systems. Gravity/gravity gradient matching navigation are passive and their navigation errors are not increase with time, so they are appropriate for aided underwater navigation [1–9].

The accuracy and success rate of gravity/gravity gradient matching navigation are seriously affected by the features of gravity changes in the sailing area. Better navigation results may be gained while gravity changes are much rougher in the matching area, and vice versa [10]. In this paper, a parameter named average gravity difference (AGD) was introduced here to express the level of gravity changes in a certain area. With the overall survey of global marine gravity/gravity gradient, the AGD values of local areas were calculated and analyzed. After a series of experiments, the relationship of AGD values and navigation accuracy in local areas can be generally determined. As the criterion of selection of matching areas was induced, marine gravity matching areas for navigation can be selected.

2 Average Gravity Differences in Different Regions of the World

2.1 Average Gravity Difference

Matching algorithm is one of the most important parts in gravity/gravity gradient aided navigation systems, while it is mostly influenced by the features of gravity changes. In this paper, an effective and efficient parameter was introduced here to present the degree of gravity changes in regions or areas, quantificationally. It is average gravity difference (AGD) which was defined as average gravity difference between neighbouring grids in gravity database or maps:

$$AGD = \frac{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n-1}\sigma_{ij}}{m*(n-1)} + \frac{\sum_{i=1}^{n}\sum_{j=1}^{m-1}\omega_{ij}}{n*(m-1)}}{2} \tag{1}$$

where σ_{ij} ($i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n-1$) and ω_{ij} ($i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m-1$) are the gravity differences between neighbouring grids in x and y directions, respectively.

2.2 Gravity/Gravity Gradient AGD Values in Regions of the World

The 1′ × 1′ resolution global marine gravity model calculated and released by D. T. Sandwell et al. in Scripps Institution of Oceanography, Univ. of California, San Diego. was adopted in this paper for analysis and simulation [11]. The earth (longitude 0°–360°, latitude −80° to 80°) was divided into 28 regions as showed in Fig. 1. The sizes of these regions are three types: 50° × 35°, 50° × 50° and 0° × 25°. The 28 regions were numbered from 1 to 28 as Fig. 1. Equation (1) was taken to calculate the gravity anomaly/vertical gravity gradient AGD values of these regions, the results were presented in Table 1.

In Table 1, the unit of gravity anomaly AGD values is mGal/n mile while the unit of gravity gradient AGD values is E/n mile (1 mGal = 10^{-5} m/s^2, 1 E = 10^{-9}/s^2, 1 n mile = 1 nautical mile = 1852 m). As the 28 regions divided here, the gravity anomaly AGD values of regions range from 0.5641 to 2.6736 mGal/n mile, the vertical gravity gradient AGD values of regions range from 1.5062 to 6.3568 E/n mile. Based on the results, curves of AGD values from region 1 to 28 can be drawn as in Fig. 2. From this figure it can be found that curves of gravity anomaly and gravity gradient AGD values have similar tendency.

Additionally, the gravity anomaly AGD value of the whole earth can be calculated as 1.1895 mGal/n mile, the gravity gradient AGD value was 2.9401 E/n mile.

From Fig. 1 and Table 1 and it can be seen that, AGD values were much different from one region to another. In some regions, the gravity/gravity gradient changes smoothly while in some other regions they change much roughly. So the position accuracy of gravity aided navigation would be much different from one region to another. Again, it underlines the importance of matching areas selection.

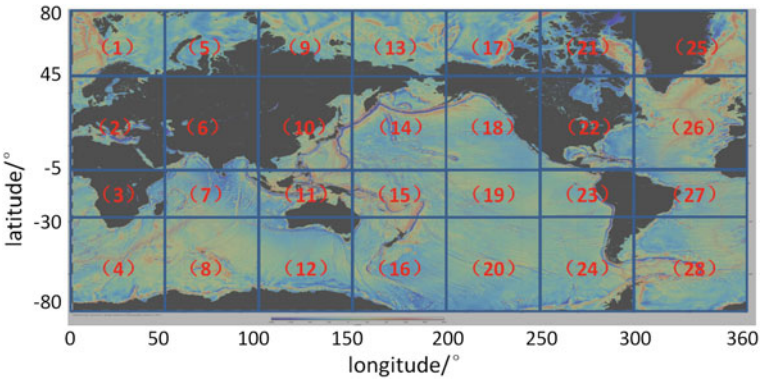


Fig. 1 Numbered regions of the world

Table 1 Gravity/gravity gradient AGD values in different regions of the world

Region no.	Gravity anomaly AGD (mGal/n mile)	Gravity gradient AGD (E/n mile)
1	0.6465	1.5773
2	1.5619	3.3601
3	1.3899	3.1489
4	0.8132	1.9305
5	0.6801	1.5062
6	2.6025	6.2903
7	1.4669	3.5989
8	0.6395	1.7671
9	0.8281	1.9470
10	2.6736	6.3568
11	1.4561	3.5106
12	0.5952	1.5806
13	0.8133	2.0740
14	1.3439	2.9043
15	1.9111	4.2506
16	0.8594	2.2379
17	1.1667	3.0820
18	1.0044	2.9424
19	1.0304	3.0601
20	0.5641	1.7110
21	0.6910	2.0081
22	1.7733	4.2405
23	2.2177	5.4480
24	0.7788	2.1058
25	0.7477	1.9290
26	1.1649	2.9129
27	1.1526	2.7645
28	0.7333	2.0782

3 Selection of Gravity Aided Navigation Matching Areas

3.1 Position Accuracies in Regions with Different AGD Values

Eight gravity maps located in different regions were chosen to the simulation tests. The AGD values of these maps are 0.846, 1.009, 1.218, 1.425, 1.865, 2.591, 3.486 and 4.791 mGal/n mile. These values had almost covered the range of AGD values in regions of the world. Two of the maps with AGD values 1.218 and 4.791 mGal/n mile were chosen to demonstrate the 3D gravity distribution as in Fig. 3. In this

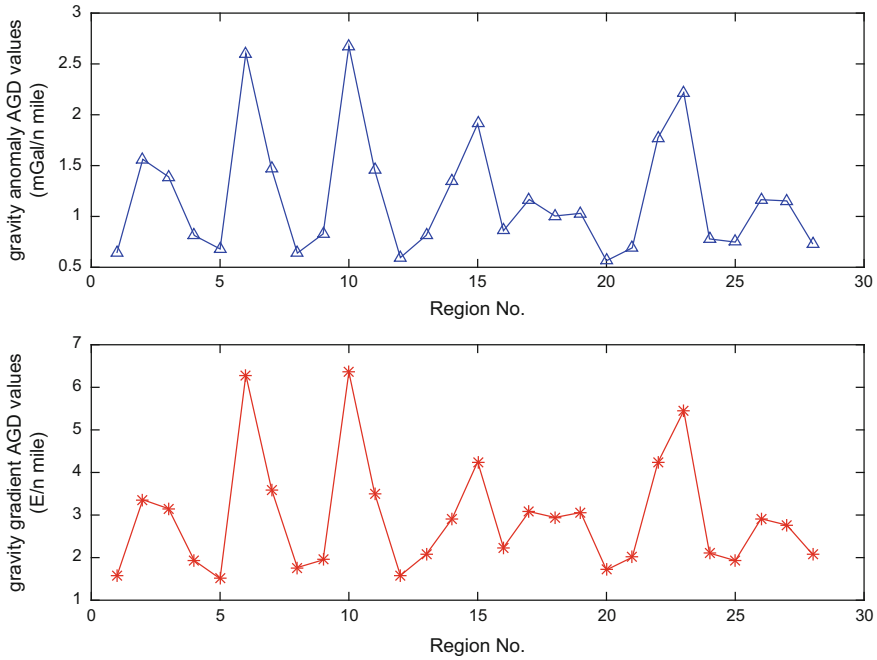


Fig. 2 Curves of AGD values in different regions of the world

figure it can be seen intuitively that, the gravity anomaly changes much rougher in the area with larger AGD value, smoother in the area with smaller AGD.

Actual dynamic gravimetry accuracy and operating conditions were referenced to design the simulation parameters. After series of tests in eight gravity maps, average position errors can be gathered and compared. The average position errors in regions with different AGD values were displayed in Fig. 4.

3.2 Analysis, Selection and Discussion

In Fig. 4 it appears that matching areas with different AGD values played a much more serious role. In most cases, the average position errors got smaller when AGD values increased.

In particular while AGD values <1.865 mGal/n mile, average position errors became much larger, as well as the matching length need to be at least 120 n miles to obtain a 0.5 n mile positioning accuracy. Especially for AGD = 0.846 mGal/n mile, the average position errors remained larger than 1.3 n mile even the matching length were enlarged to 150 n mile. By contrast, while AGD >2.591 mGal/n mile average position errors got much smaller, also the matching length need to be just no more than 50 n miles to obtain a 0.5 n mile positioning accuracy [12].

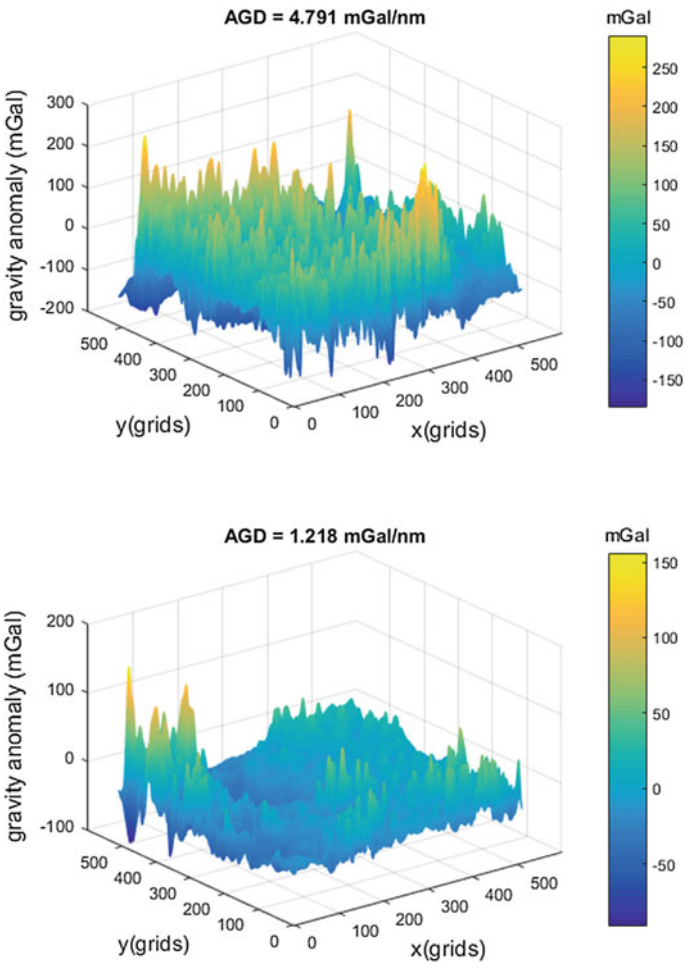


Fig. 3 Distribution of gravity anomalies in different regions

Based on these results, a general criterion of gravity aided navigation matching areas selection can be constructed: if the gravity anomaly AGD values larger than 2.0 mGal/n mile, gravity matching navigation could achieve an accuracy of 0.5 n miles in these matching areas with present technologies. Besides, it can be found in Table 1 that in regions with gravity anomaly AGD values close to or larger than 2.0 mGal/n mile, their gravity gradient AGD values were larger than 4 E/n mile.

With this criterion, the 28 regions presented in Fig. 1 were surveyed. Regions 6, 10, 15, 22, 23 were selected to measure up to the criterion, as indicated by orange color in Fig. 5 which corresponding to areas of the Northern Indian Ocean, South China Sea, Coral Sea and Eastern Pacific Ocean. In these areas, better navigation results can be obtained since they have rougher gravity changes.

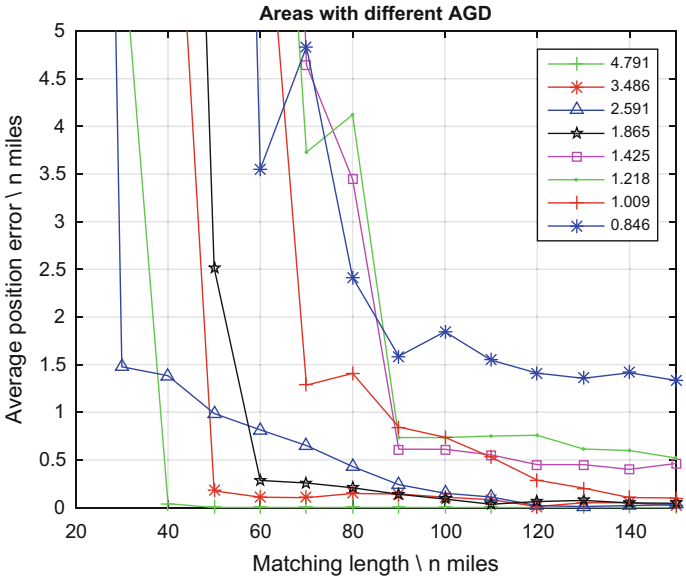


Fig. 4 Average position errors in different regions

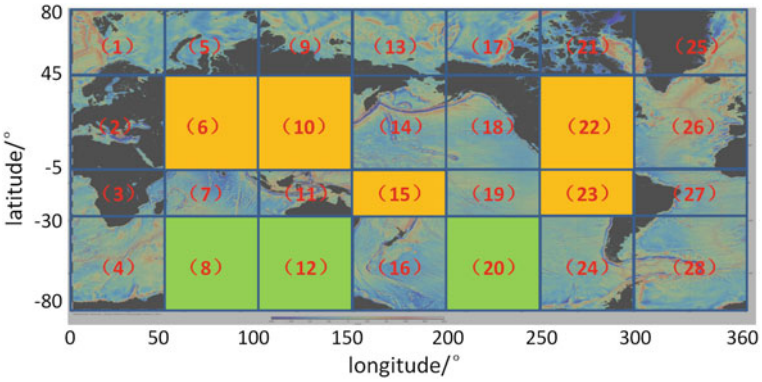


Fig. 5 Selection of gravity/gravity gradient aided navigation matching areas

Conversely, regions 8, 12, 20 with smallest AGD values were also selected, as indicated by green color in Fig. 5 which corresponding to areas of the Southern Indian Ocean, Southwest Pacific Ocean. In these areas, bad navigation results may be obtained since the gravity changes smoothly.

4 Conclusion

Selection of matching areas is significant for the performance of gravity matching aided navigation system. In this paper, a parameter named average gravity difference (AGD) was introduced for the expression of the level of gravity changes in a certain area. With the overall survey of global marine gravity, the AGD values of local areas were calculated and analyzed. After a series of experiments, the relationship of AGD values and navigation accuracy in local areas was generally determined. As the criterion of matching areas selection was induced, marine gravity matching areas for navigation were selected.

Simulation results showed that, the AGD values can range from about 0.6 mGal/n mile to greater than 2.7 mGal/n mile in different regions of the world. With present technologies, gravity matching navigation could achieve an accuracy of 0.5 nautical miles in matching areas which have AGD values larger than 2.0 mGal/n mile. With this criterion, Northern Indian Ocean, South China Sea and some other areas can be selected. In these areas, better navigation results can be obtained since they have rougher gravity changes. Further selection and study should be implemented in these areas.

These results would be significant for the development of gravity aided navigation technologies and the construction of underwater PNT system. However, more detailed analysis of marine gravity matching areas with smaller scale is still needed. Also, further simulations and practical experiments should be implemented to verify the matching areas selection criterions.

References

1. Moryl J, Rice H, Shinners S (1998) The universal gravity module for enhanced submarine navigation. In: Proceedings of IEEE/ION position location and navigation symposium 1998, Palm Springs, CA, USA, 20–23 April 1998, pp 324–331
2. Rice H, Kelmenson S, Mendelsohn L (2004) Geophysical navigation technologies and applications. In: Proceedings of IEEE/ION position location and navigation symposium 2004, Monterey, CA, USA, 26–29 April 2004, pp 618–624
3. Wang Z, Bian S (2008) A local geopotential model for implementation of underwater passive navigation. Prog Nat Sci 18(9):1139–1145
4. Wu L, Ke X, Hsu H, Xiong C, Wang Y (2013) Joint gravity and gravity gradient inversion for subsurface object detection. IEEE Geosci Remote Sens Lett 10(4):865–869
5. Deng Z, Ge Y, Guan W, Han K (2010) Underwater map-matching aided inertial navigation system based on multi-geophysical information. Front Electr Electron Eng China 5(4): 496–500
6. Wu L, Wang H, Hsu H, Chai H, Wang Y (2015) Research on the relative positions-constrained pattern matching method for underwater gravity-aided inertial navigation. J Navig 68(5):937–950
7. Wang H, Wu L, Hsu H, Chai H, Zheng H, Wang Y (2016) Technology of gravity aided inertial navigation system and its trial in South China Sea. IET Radar Sonar Navig 10(5): 862–869

8. DeGregoria A (2009) Gravity gradiometry and map matching: an aid to aircraft inertial navigation systems. Master's Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH
9. Richeson J (2008) Gravity gradiometry aided inertial navigation within non-GNSS environments. Ph.D. Dissertation, University of Maryland, College Park, MD
10. Wang H, Wu L, Chai H, Xiao Y, Hsu H, Wang Y (2017) Characteristics of marine gravity anomaly reference maps and accuracy analysis of gravity matching-aided navigation. Sensors 17(8):1851
11. Sandwell DT, Müller RD, Smith WHF, Garcia E, Francis R (2014) New global marine gravity model from CryoSat-2 and Jason-1 reveals buried tectonic structure. Science 346(6205): 65–67
12. Wu L, Wang H, Chai H, Zhang L, Hsu H, Wang Y (2017) Performance evaluation and analysis for gravity matching aided navigation. Sensors 17(4):769

Building the Crab Pulsar Timing Model with XPNAV-1 Observations

Hai-feng Sun, Jian-yu Su, Liang Zhao, Zhi-wei Liu and Hai-yan Fang

Abstract This paper addresses the problem of building the X-ray pulsar timing model that is an essential work for X-ray pulsar navigation (XNAV). We propose a theoretical discussion of the differences between the X-ray and radio pulsar timing and develop a complete processing framework. In this framework, the Z_m^2-test method is deduced by the theory of the Fourier Series Estimator (FSE) in the complex form. A fast maximum likelihood (FAML) estimator is applied for the estimation of pulse time of arrival (TOA). Then the model parameters are deduced based on a least-squares strategy. To sustain this framework, about 173 ks observations with X-ray pulsar navigation-I (XPNAV-1) satellite over 30 days are utilized for experimental studies. The results show that the proposed framework is valid, and the phase residuals with the X-ray model parameters is 2.57 milliperiods.

Keywords X-ray pulsar navigation · X-ray timing model · Crab pulsar XPNAV-1 satellite

1 Introduction

X-ray pulsar navigation (XNAV) is a developing autonomous navigation technique, which offers a potentially revolutionary navigational solution throughout our Solar System and beyond. Since the concept of the XNAV was first presented in 1981,

H. Sun · J. Su · Z. Liu · H. Fang (✉)
School of Aerospace Science and Technology, Xidian University,
Xi'an 710071, China
e-mail: hyfang@xidian.edu.cn

L. Zhao
Science and Technology on Space Physics Laboratory,
Beijing 100076, China

Z. Liu
Beijing Institute of Astronautical Systems Engineering,
Beijing 100076, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_75

this technique has attracted widespread attention in these days [1–3], and several important projects have been launched for demonstration and verification. About 1 month observations from the Gamma-ray Burst Polarimeter (POLAR) on the Chinese space laboratory Tiangong-2 (TG-2) were analyzed, and preliminary estimation accuracy of the TG-2 orbit is about 19.7 km [4]. The recently launched X-ray pulsar navigation-I (XPNAV-1) is a China's experimental satellite designed for XNAV [5], and it is reported that more than 10 pulsars will be characterized for their pulse frequency and intensity to create a navigation database in the future observation plan. The Station Explorer for X-ray Timing and Navigation Technology (SEXTANT) is a NASA funded technology demonstration enhancement to the Neutron-star Interior Composition Explorer (NICER) an International Space Station (ISS) payload, whose primary objective is to demonstrate real-time orbit determination with uncertainty better than 10 km with 2 weeks of measurements in the highly dynamic ISS orbit [6]. China's X-ray Modulation Telescope (HXMT) was launched in June 2017, which carries three main payloads onboard: the High Energy X-ray telescope, the Medium Energy X-ray telescope and the Low Energy X-ray telescope. Some in-orbit experiments are being performed to investigate the feasibility of XNAV [7].

For an XNAV system, a timing model building is a fundamental work. It is necessary to monitor these navigation candidates and to update new models with X-ray observations especially for the deep-space situation where real-time communication is impossible. In X-ray astronomy, the phase misalignment between the radio and X-ray pulses has important repercussions for the understanding of the spatial origin of the emission processes [8], and one can get the phase misalignment by comparing the radio and X-ray timing models. The X-ray timing model might help confirm the variations in radio dispersion measure (DM) through comparing timing residuals ($\sim 10^{-3}$ ns).

In contrast to the radio timing process, the X-ray timing process may encounter two issues: (1) the orbital period of an X-ray observation platform, is generally far less than that of the ground-based radio telescopes; (2) the detected X-ray signals are actually random and discrete photon TOAs with low radiation intensity. Aiming at these two issues, this paper presents a processing framework and accurate experimental analysis for X-ray timing. The rest of this paper is organized as follows. Section 2 presents the details of the X-ray timing processing framework. Section 3 deals with the experimental phase of our work. Section 4 summarizes the concluding remarks.

2 X-Ray Timing Processing Framework

It is essential to provide precise procedures to implement time transformations, frequency search, pulse TOA estimation and model parameters determination.

2.1 Forming Barycentric TOAs

Geometric and relativistic effects are included in the observed accumulated phase, and observed TOAs should be deduced to barycentric TOAs [9–11]

$$t_{\mathrm{SSB}} - \tau_{\mathrm{obs}} = \Delta_{\mathrm{C}} + \Delta_{\mathrm{P}} + \Delta_{\mathrm{R}} + \Delta_{\mathrm{E}} + \Delta_{\mathrm{S}} + \Delta_{\mathrm{D}} + \Delta_{\mathrm{A}} \tag{1}$$

where Δ_{C} is clock corrections, Δ_{P} parallax delay, Δ_{R} Roemer delay, Δ_{E} Einstein delay, Δ_{S} Shapiro delay, Δ_{D} Solar system dispersion delay, and Δ_{A} atmospheric delay. This equation holds for isolated pulsars, and a binary system transformation involves the propagation delay due to binary orbital motion. Full details about the transformations are not focused here, but the differences of the radio and X-ray timing procedures.

(1) Significant dispersion exists for the radio waves in the interplanetary medium because of the electron content of the solar wind, whereas the Solar system dispersion delay is considered to be zero for X-ray radiation [9]. Besides, the atmospheric delay is negligible in the X-ray TOAs transformation as the X-ray instruments are always set above the Earth's atmosphere.
(2) For radio observations, synchronously 1000 or more pulses are averaged and orthogonal polarizations are de-dispersed and summed to get a mean total intensity pulse profile, and then the measured topocentric pulse TOAs are corrected to the barycentric ones at SSB because the velocity of the radio observatory relative to the pulsar in the SSB reference frame changes slowly. However, the X-ray TOAs transformation is of a different situation. The barycentric frequency differs from the observed frequency for the motion of the observation platform relative to the SSB [12], and the Doppler frequency shift $v_{\mathrm{d}}(t)$ can be expressed by

$$v_{\mathrm{d}}(t) = -v_{\mathrm{s}}(t)\left(\frac{\mathrm{d}\Delta_{\mathrm{R}}}{\mathrm{d}t} + \frac{\mathrm{d}\Delta_{\mathrm{P}}}{\mathrm{d}t} + \frac{\mathrm{d}\Delta_{\mathrm{E}}}{\mathrm{d}t} + \frac{\mathrm{d}\Delta_{\mathrm{S}}}{\mathrm{d}t}\right) \tag{2}$$

Here the terms in the right-hand side are respectively the time-derivatives of the Roemer delay, parallax delay, Einstein delay and Shapiro delay. A Doppler frequency shift about 10^{-4} Hz in magnitude is introduced by the first term $\mathrm{d}\Delta_{\mathrm{R}}/\mathrm{d}t$, also called the Roemer rate. The effect of the sum of the other three time-derivative terms on the Doppler frequency shift, also called the high-orders of the Doppler frequency shift, is ignorable.

2.2 Pulse Frequency Search

Complete classes of tests for uniformity on the pulse period were derived in [11], among which Pearson's χ^2-test and Z_m^2-test seem to be the most popular in the X-ray astronomy. The latter is an improvement over Pearson's χ^2-test for it is independent on the sizes of bins. However, the Z_m^2-test still suffers with respect to the choice of smoothing parameters. If the smoothing parameter is small, the corresponding test is very powerful against broad peaks, and if large, narrow peaks have priority in detection. Here the Z_m^2-test method is deduced by the theory of the Fourier Series Estimator (FSE) in the complex form.

Consider a series of barycentric TOA t_i $(i = 1, 2, \ldots, N)$ for a pulsar. With respect to an assumed pulse frequency v, the phases of the barycentric TOA can be calculated as $\phi_i = vt_i$. A cost function (CF) is given by the measure of the distance between a periodic density function $f(\phi)$ and uniform density 1, and can be expressed by

$$Z^2 = N \int_0^1 (f(\phi) - 1)^2 d\phi \tag{3}$$

It is clear that if a periodic signal is in presence, the Z^2 statistic is large and the null hypothesis should be rejected. If we specify $\hat{f}_m(\phi)$ as the Fourier Series Estimator (FSE) with m harmonics, then the FSE in the complex form is

$$\hat{f}_m(\phi) = \sum_{k=-m}^{m} \hat{F}_k \exp(j2\pi k\phi), \quad j = \sqrt{-1} \tag{4}$$

where $\hat{F}_k$ is an unbiased estimate of the k-th Fourier coefficient equaling

$$\hat{F}_k = 1/N \sum_{i=1}^{N} \exp(-j2\pi k\phi_i) \tag{5}$$

Submitting Eqs. (4) and (5) into (3), one can obtain the Z_m^2-test cost function

$$Z_m^2 = 2N \sum_{k=1}^{m} \left|\hat{F}_k\right|^2 \tag{6}$$

By processing the observations, we find that the estimation accuracy of the Pearson's χ^2-test is not only related to the number of bins, but also to the selected frequency step. An improper number of bins or a small frequency step may lead to plenty of burrs on the curve of the cost function; however, the Z_m^2-test CF appears smooth and no signs of burrs occur despite a very small frequency step.

2.3 Pulse TOA Estimation

The phase of the pulse peak is used as representative of the X-ray pulse phase or pulse TOA. Two major peak-determination techniques are commonly used to find the pulse phase: (1) fitting a parabolic or Lorentzian functions to the highest bin in the profile [8]; (2) using the phase estimation technique to obtain the phase or time offset between the observed profile and standard profile [13, 14]. The former are available for the observed profile with high signal to noise (SNR), but a large fitting error may occur when the curve is unsmooth in the neighbourhood of the pulse peak. The latter are measuring the displacement of two similar profiles instead of determining the location of the pulse peak, among which the most commonly used methods are the cross correlation (CC) and maximum likelihood (ML) techniques. Here we use a fast maximum likelihood (FAML) estimator as an alternative to the ML estimator since the ML estimator utilizes the X-ray TOAs for the algorithm implementation, resulting in a high computational complexity that linearly grows with the observation time.

The logarithmic form of the FAML estimator with respect to the phase offset ϕ_{D} between the observed and standard pulsar rate functions is [13]

$$\mathrm{J}(\phi_{\mathrm{D}}) = \sum_{i=1}^{N_{\mathrm{b}}} C_i \ln[\lambda(\phi_{\mathrm{D}} + i/N_{\mathrm{b}})] \tag{7}$$

where C_i is the statistical photon count within the i-th bin after the epoch folding procedure, N_{b} the number of bins contained in one phase cycle, and $\lambda(\phi)$ is the standard pulsar rate function which equals

$$\lambda(\phi) = \lambda_{\mathrm{b}} + \lambda_{\mathrm{s}} h(\phi) \tag{8}$$

Here λ_{b} and λ_{s} are the effective background and source arrival rates respectively. Instead of calculating the observed pulsar rate function, this equation uses statistical photon count directly to estimate the phase offset through maximizing Eq. (8)

$$\hat{\phi}_{\mathrm{D}} = \arg \max_{\phi_{\mathrm{D}} \in (0,1)} \mathrm{J}(\phi_{\mathrm{D}}) \tag{9}$$

Note that this form is similar to CC estimator except the employment of the logarithm. This kind of variance stabilizing transformations can mitigate the effects of epoch folding noises on the phase estimation accuracy.

2.4 Model Parameters Determination

Pulse timing is achieved by fitting the barycentric pulse TOAs to a Taylor expansion of pulse phase. The total accumulated phase $\Phi(t)$ is

$$\Phi(t) = \Phi_0 + \sum_{k=1}^{n} \frac{1}{k!} \nu^{(k-1)} (t - t_0)^k \tag{10}$$

Here the subscript 0 in this equation denotes a parameter evaluated at the reference epoch t_0, and thus Φ_0 denotes an accumulated initial phase at. The k-order frequency derivative terms $\nu^{(k)}$ are fittable model parameters. If $k = 0$, $\nu^{(0)} \equiv \nu_0$ denotes the pulse frequency at t_0.

The observation time t can be set as the proper time t^{psr} of a pulsar mass center [9, 12], or the barycentric time t^{SSB} of the SSB in the use of TCB or TDB. Different time system of the observation time t will lead to different timing parameters. One can also choose the Geocentric Coordinate Time (TCG) in the model; however, some periodic terms should be added into the Taylor expansion for the revolution of the Earth. For young pulsars, there exist large stochastic variations in their spin-down rates, so more high-order frequency derivatives are required to remove red timing noises [12]. For a binary system, a further complicated timing model incorporating the pulsar orbital period has to be taken into consideration if the fitted positions are fixed in the geocenter or the SSB. Taking the partial derivative of $\Phi(t)$ with respect to t, one can get the spin frequency expression

$$\nu_{\mathrm{s}}(t) = \sum_{k=1}^{n} \frac{1}{(k-1)!} \nu^{(k-1)} (t - t_0)^{k-1} \tag{11}$$

The pulse peak TOA t_i^{pk} at the SSB and a given timing model are used to form the i-th pre-fit phase residual

$$r_i = \Phi\left(t_i^{\mathrm{pk}}\right) - N_i \tag{12}$$

where N_i is the nearest integer to $\Phi\left(t_i^{\mathrm{pk}}\right)$. Note that if the given timing model is accurate, each phase residual $r_i \to 0$. After eliminating the phase jumps, the model parameters can be fitted based on the least-squares strategy.

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{13}$$

where $\boldsymbol{Y} = \left[\Phi_{\mathrm{w}}\left(t_1^{\mathrm{pk}}\right) - \Phi_0, \Phi_{\mathrm{w}}\left(t_2^{\mathrm{pk}}\right) - \Phi_0, \ldots, \Phi_{\mathrm{w}}\left(t_{\mathrm{m}}^{\mathrm{pk}}\right) - \Phi_0\right]^{\mathrm{T}}$, $\boldsymbol{X}$ is the Vandermonde matrix formed by Δt_i^{pk} with m columns, $\boldsymbol{\beta} = [\nu^{(n-1)}/n!, \nu^{(n-2)}/(n-1)!, \ldots, \nu^{(0)}]^{\mathrm{T}}$, and $\boldsymbol{\varepsilon}$ is constructed by pre-fit phase residuals ε_i s.

The elements of $\boldsymbol{\varepsilon}$ are uncorrected random variables with mean $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and covariance matrix $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{I}$.

The measurements equal the true values plus random, uncorrelated errors with constant variance, and the least squares problems is to find the estimate of $\boldsymbol{\beta}$ to minimize [15]

$$Q = min||\boldsymbol{Y} - \hat{\boldsymbol{Y}}||_2^2 = \min||\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}||_2^2 \tag{14}$$

where $||\cdot||_2^2$ means the Euclidean norm of a vector. Finding the least squares estimate $\hat{\boldsymbol{\beta}}$ is developed by the QR method. By the orthogonal triangular decomposition method, the matrix $\boldsymbol{X}$ is decomposed into an $m \times n$ triangular matrix $\boldsymbol{R}$ and a $n \times n$ unitary matrix $\boldsymbol{Q}$, so that $\boldsymbol{X}$=$\boldsymbol{QR}$. The least squares solution is

$$\hat{\boldsymbol{\beta}} = \boldsymbol{R}^{-1}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Y} \tag{15}$$

3 Observations and Results

Observations of the Crab pulsar were derived from the Time-resolved Soft X-ray Spectrometer (TSXS) on board the XPNAV-1 satellite with an orbital altitude of 500 km [5]. The TSXS consists of Wolter-I type lens of four nested mirror shells that focus X-ray photons onto a silicon drift detector (SDD) operating in the 0.5–10 keV range, with a combined nominal collecting area of 30 cm^2 @1.5 keV and a field of view 15′ [5]. We used 1 month of public observations (MJD 57709-57740) in event mode which records the X-ray TOAs with a resolution of 1.5 μs.

The observations were reduced using the following screening criteria: (1) the effective arrival rate exceeds 10 ph/s; (2) the observation time $T_{\mathrm{obs}} \geq 2300$ s. After this screening, 58 eligible subsets were selected, and about 2450 000 events were contained with a total exposure time 173 ks. In our analysis, the radio monthly timing ephemeris is derived from Jodrell Bank observations daily at 610 MHz and weekly at 1420 MHz, which is published on the web (see http://www.jb.man.ac.uk/pulsar/crab/all.gro). Two radio ephemeris records are used here which are denoted by ‘Jodrell Bank I’and ‘Jodrell Bank II’ in Table 1.

Table 1 Spin parameters for the Crab pulsar

Parameters	Jodrell Bank I	Jodrell Bank II	XPNAV-1
ν (Hz)	29.7478547478211	29.6471215546085	29.6478547377(2)
$\dot{\nu}$ (10^{-10} Hz/s)	−3.6872	−3.68942	−3.689502(8)
$\ddot{\nu}$ (10^{-20} Hz/s^2)	−3.89	1.55	−6.2(1)
t_0 (MJD)	57715.000000295	57738.000000232	57715.00000295

To confirm the result of the pulse frequency searching, we searched the frequency in individual subset via the Z_m^2-test (m = 10), and compared the searched frequencies with the model frequencies from the radio ephemeris. As shown in Fig. 1, this plot clearly shows the slow-down trend of the rotation that is determined by the rotation of the magnetic dipole, through a combination of classical dipolar radiation at the rotation frequency and an outflow of energetic particles. The Z_m^2-test technique is less sensitive to the low average flux, and the rms of frequency residuals between the searched and model frequencies is about 2.6×10^{-6} Hz with a reduced chi-square of 1.04 for 56 degrees of freedom.

In Fig. 2, the pre-fit phase residuals were achieved with initial model parameters. A lineally increasing trend and four phase-jumps are presence in the plot of the pre-fit phase residuals, and the reason is that the Crab pulsar phase cannot be well described by the initial model parameters. Predictably, there will be more phase jump points in the pre-fit phase residuals, if one select lower accurate spin parameters.

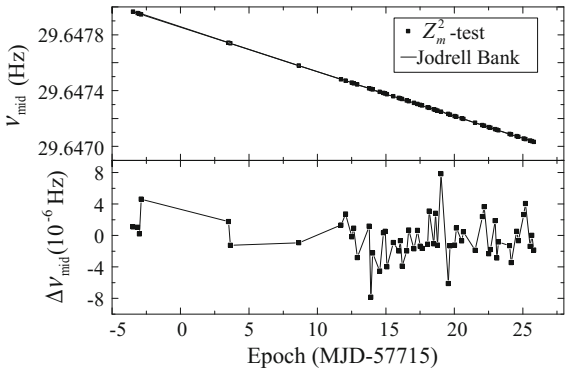


Fig. 1 Top panel has the searched pulse frequency and the radio model frequency. Bottom panel gives the frequency residuals

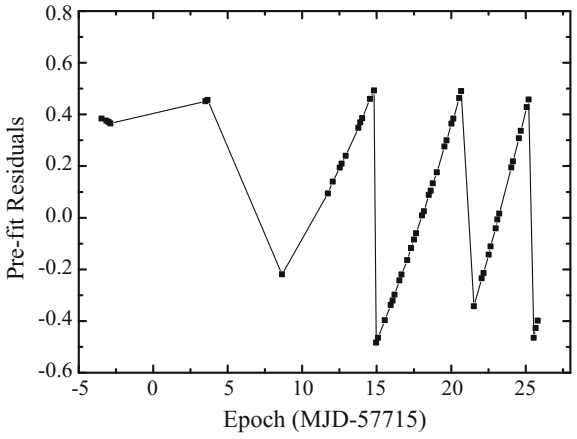


Fig. 2 Pre-fit phase residuals achieved by using initial model parameters with errors

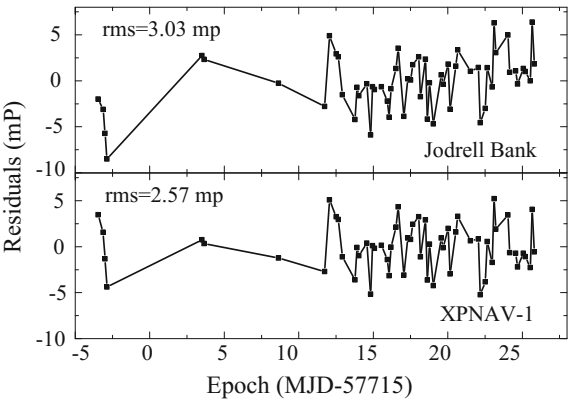


Fig. 3 Phase residuals for the Crab pulsar with the Jodrell Bank radio ephemerides (top panel) and with fitted parameters via the presented processing framework (bottom panel)

The pulsar position is fixed with RA $= 05^{\mathrm{h}}34^{m}31^{s}.972$ and DEC $= 22°00'52''.07$ in the J2000 coordinates. We phase-connected all the sets of the X-ray TOAs, and the spin parameters obtained by minimizing the phase residuals are listed in Table 1, which is denoted as 'XPNAV-1'.

The post-fit phase residuals were calculated for the Crab pulsar with the given radio and X-ray parameters in Fig. 3. The linear trend is not seen in the data, indicating that there are no systemic errors for this fit. Note that the rms of the post-fit phase residuals with the X-ray model parameters is 2.57 milliperiods (mPs) with a reduced chi-square of 1.018 for 56 degrees of freedom, which is slightly smaller than that with the radio one derived from the Jodrell Bank observatory. However, the rms given by the radio ephemeris is about 0.6 mPs which is far less than 3.03 mPs calculated from the XPNAV-1 observations, implying that the obtained rms residuals may overestimate the true uncertainty on the X-ray timing since the errors of the X-ray pulse TOAs are larger than that of the radio pulse TOAs.

All the barycentric TOAs were folded via the X-ray and radio model parameters to achieve the pulse profiles, as shown in Fig. 4. By contrast, both the two pulse profiles have steep and smooth curve-shapes, and exhibit a virtually perfect agreement, meaning that the discrepancies between them are negligible and the model parameters are accurate.

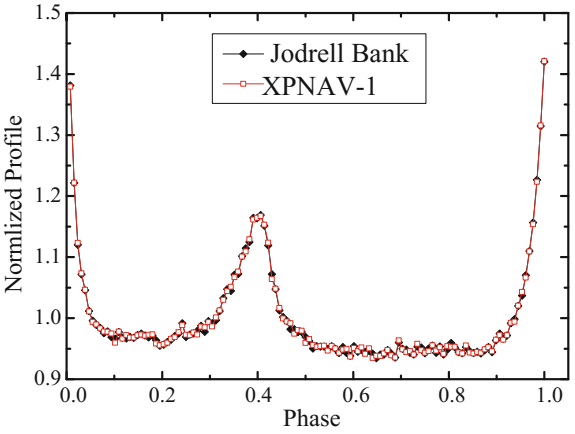


Fig. 4 Crab pulse profiles obtained by folding all the TOAs with X-ray/radio model parameters

4 Conclusions

This paper developed a processing framework for the X-ray timing model to overcome the effect of the dynamics of observation platforms and the low X-ray radiation intensity. The model parameters are obtained by using the X-ray observations of the Crab pulsar with XPNAV-1 satellite for 31 days on the basis of least-squares strategy. Sometimes it seems difficult to get the initial frequency range with a high accuracy, resulting in the total accumulated phase ambiguity, so future work will be dedicated to solving the model parameters with large initial errors.

Acknowledgements Funding support for this work was provided by the National Key Research and Development Program of China [Grant No. 2017YFB0503300]; the National Natural Science Foundation of China [Grant No. 61603287].

References

1. Sheikh SI, Pines DJ, Ray PS, Wood KS, Lovellette MN, Wolff MT (2006) Spacecraft navigation using X–ray pulsars. J Guid Control Dynam 29(1):49–63
2. Liu J, Fang JC, Kang ZW, Wu J, Ning XL (2015) Novel algorithm for X–ray pulsar navigation against doppler effects. IEEE T Aero Elec Syst 51(1):228–241
3. Wang YD, Zheng W, Sun S, Li L (2013) X-ray pulsar-based navigation system with the errors in the planetary ephemerides for Earth-orbiting satellite. Adv Space Res 51(12):2394–2404
4. Zheng SJ, Ge MY, Han DW et al (2017) Test of pulsar navigation with POLAR on TG–2 space station. Sci Chin Ser G 47(9):1–9
5. Shuai P, Zhang XY, Huang LW et al (2017) X–ray pulsar navigation test satellite data analysis. Aerosp Control Appl 43(2):1–6
6. Mitchell JW, Hassouneh MA, Winternitz LMB et al (2015) SEXTANT–station explorer for X–ray timing and navigation technology. In: AIAA guidance, navigation, and control conference, Kissimmee, FL, USA, 2015, p 16

7. Li TP, Xiong SL, Zhang SN et al (2018) Insight-HXMT observations of the first binary neutron star merger GW170817. Sci China Phys Mech 61(3):031011
8. Rots AH, Jahoda K, Lyne AG (2004) Absolute timing of the Crab pulsar with RXTE. Astrophys J 605:129–132
9. Edwards RT, Hobbs GB, Manchester RN (2010) Tempo2, a new pulsar timing package—II. The timing model and precision estimates. Mon Not R Astron Soc 372(4):1549–1574
10. Sun HF, Sun X, Fang HY et al (2018) Building X-ray pulsar timing model without the use of radio parameters. Acta Astron. 143:155–162
11. De Jager OC, Swanepoel JWH, Raubenheimer BC (1989) A powerful test for weak periodic signals with unknown light curve shape in spare data. Astron Astrophys 221:180–190
12. Hobbs GB, Edwards RT, Manchester RN (2010) TEMPO2, a new pulsar–timing package–I: an overview. Mon Not R Astron Soc 369(2):655–672
13. Xue MF, Li XP, Sun HF, Fang HY (2016) A fast pulse phase estimation method for X-ray pulsar signals based on epoch folding. Chinese J Aeronaut 29(3):746–753
14. Sun HF, Bao WM, Fang HY, Shen LR, Xue MF, Li XP (2016) Effect of X-ray energy band on the X–ray pulsar based navigation. Aerosp Sci Technol 58:150–155
15. Rice JA (2007) Mathematical statistics and data analysis, 3rd edn. Thomson Higher Education, Belmont, CA

Spatial Coordinate Autonomous Establishment Technique for High-Precision Local Positioning System

Zhen Wu, Zheng Yao and Mingquan Lu

Abstract To overcome the lack of accuracy, reliability, availability, and anti-jamming in satellite navigation system and meet the needs of military and civilian users for positioning, navigation, and timing (PNT) at anytime, anywhere and in any environment, there is an urgent need to study the local PNT spatiotemporal baseline network based on land-based pseudolites. However, in the existing local positioning system, the requirement for accurate mapping of base stations coordinates in advance limits the mobility, flexibility and rapid deployment capability of the system. In order to solve this problem, this paper studies the technique of establishing local positioning system spatial coordinate independently in high precision. Based on the shortest path algorithm and the nonlinear conjugate gradient method, the proposed technique can independently estimate the relative position of each pseudolite and the network structure of the local positioning system according to the partial ranging information between the pseudolite base stations, and establish the local positioning system spatial coordinate. The results of performance analysis show that compared with other existing space coordinate establishment techniques, the technique proposed not only has the higher precision in the spatial coordinate establishment but also has lower computational complexity. Moreover, it is practical that it can handle the circumstance in which partial distance information in the network is missing. It can greatly simplify the laying costs and shorten the deployment time of local PNT system, reduce the maintenance difficulty and greatly improve the mobility and flexibility of the system in practice.

Keywords Local positioning system · Spatial coordinate · Pseudolite
Euclidean distance matrix · Shortest path · Nonlinear conjugate gradient method

Z. Wu · Z. Yao (✉) · M. Lu
Tsinghua University, Beijing, China
e-mail: yaozheng@tsinghua.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 499,
https://doi.org/10.1007/978-981-13-0029-5_76

1 Introduction

Safe, reliable and diversified positioning and navigation means are required in future military and civilian applications. The satellite navigation system has advantages including a large range of coverage, all-weather and all-day availability, high precision, convenient application and the support for an unlimited number of users, which has become the preferred positioning, navigation and timing (PNT) means worldwide, but it is limited to the physical characteristics of signals, constellation configurations and the environment. In some harsh scenarios, there exist problems including serious satellite obstacles, weak signals leading to poor positioning accuracy, degraded integrity, reliability and anti-jamming ability. In order to overcome these drawbacks of the satellite navigation system and meet the needs of all military and civilian users for positioning, navigation, and timing anytime, anywhere and in any environment, there is an urgent need to study the local PNT spatiotemporal baseline network based on land-based pseudolites.

The signal structure and positioning principle of the land-based pseudolite system are similar to those of the satellite navigation system. Not only can it be tightly integrated with the satellite navigation system to provide more equivalent visible satellites, but also it has the ability to position and navigate independently in high precision in the coverage area, which can be used as a local backup navigation system under the circumstance in which the satellite navigation system is not available because of jamming and blocking.

When the pseudolite navigation system works independently, the phase center coordinates of all the pseudolite base stations need to be accurately mapped in advance, and then the information is broadcast to the users through ephemeris, which is used for solving the positioning equation. The measurement accuracy of phase center coordinates will directly affect the error of positioning solution. However, the requirement for calibration in advance greatly increases the deployment time of the pseudolite system. In the military and civilian areas, there are a lot of applications requiring high mobility, flexibility and rapid deployment for the pseudolite system. In some environments lacking the prior mapping condition, like fire sites, disaster areas, the application of the pseudolite system will be limited.

In recent years, the problem of autonomous positioning of beacons draws more and more attention, especially when the noisy or unavailable distance measurements between some beacons exist and make the problem more difficult. The existing research methods mostly start with the Euclidean distance matrix. Based on the low-rank property of the Euclidean distance matrix, SVD is used to reconstruct Euclidean distance matrix in [1], which is unelaborate. The research in [2] modifies the low-rank matrix complete algorithm Optspace, in order to complete the Euclidean distance matrix, but the restored low-rank matrix is not guaranteed to be a Euclidean distance matrix. The semidefinite programming relaxation (SDR) technique [3] relaxes the problem to a convex optimization problem by relaxing the constraints of the problem, and [4, 5] further summarize and propose various relaxation methods, obtaining different convex optimization problems.

Although the convex optimization problem can find the optimal solution of the relaxing problem, it still costs a great deal of computation burden.

Given the limitations of existing techniques and the urgent need of the spatial coordinate autonomous establishment in the construction of the local positioning system, this paper presents the technique for autonomous establishment of local positioning system spatial coordinate in high precision, which is based on the shortest path algorithm and nonlinear conjugate gradient method. The problem is modeled as the Euclidean distance matrix completion problem on the principle of minimum variance. The gradient expression of the cost function is given through rigorous mathematical derivation, and then the initial value is estimated by the shortest path algorithm, and the convergent solution is found by using the nonlinear conjugate gradient method. The relative position of each pseudolite and the pseudolite network constellation is recovered according to the partial ranging information between the pseudolite base stations.

The results of performance analysis show that compared with other existing techniques, the technique proposed not only has the higher precision in the spatial coordinate establishment, but also has lower computational complexity. It is practical that it can handle the missing distance information in the network, which can significantly simplify the laying costs and shorten the deployment time, reduce maintenance difficulty of the pseudolite system, greatly improving the mobility and flexibility of the local positioning system in practice.

2 Problem Description

In a p-dimensional pseudolite system that contains k pseudolites, the coordinates matrix of pseudolites is represented as $\boldsymbol{X} = [\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{x}_k]^{\mathrm{T}} \in \mathbb{R}^{k \times p}$. The Euclidean distance between each pseudolite pair is $d_{ij} = \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|$, which forms the Euclidean distance matrix $\boldsymbol{D} = \left(d_{ij}^2\right)$.

Because the Euclidean distance matrix only contains the distance information between pseudolites, the corresponding Euclidean distance matrix is the same no matter how the original pseudolite network is translated, rotated and reflected. This means that when obtaining the Euclidean distance matrix from position vectors, the absolute position information is lost, only leaving the relative position between pseudolites. Therefore, only relative position information can be recovered from the Euclidean distance matrix. To obtain the absolute position, additional anchor information is required. The absolute position of all nodes can be obtained by matching the anchor points in the relative position to the known anchor points.

In traditional recovering method Multidimensional Scaling (MDS) [6], the relationship between $\boldsymbol{D}$ and $\boldsymbol{X}$ is $-0.5\boldsymbol{JDJ} = \boldsymbol{JX}(\boldsymbol{JX})^{\mathrm{T}}$, which can be obtained by defining the centering matrix $\boldsymbol{J} = \boldsymbol{I}_k - \boldsymbol{1}_k\boldsymbol{I}_k^{\mathrm{T}}/k$. The spatial dimension of distribution of pseudolite system is 2 or 3 generally, so $\boldsymbol{JDJ}$ is a low-rank matrix. $-0.5\boldsymbol{JDJ}$ can be decomposed as $\boldsymbol{BB}^{\mathrm{T}}$ and then $\boldsymbol{B} = \boldsymbol{JX}$, which means geometrically moving the geometric center of pseudolites to the origin of coordinate.

However, in practical applications, the ranging between pseudolites may have following problems, resulting in the inapplicability of MDS. One is the noise of ranging results and the other is that the distances between some pseudolites cannot be measured due to obstacles and overlarge layout range, as shown in Fig. 1. So the observed Euclidean distance matrix $\boldsymbol{D}_{obs}$ is incomplete in most cases.

Based on the problem description, introducing noise matrix $\boldsymbol{N} = (n_{ij})$ and measuring matrix $\boldsymbol{E} = (e_{ij})$, where $e_{ij} = 1$ means the distance between i and j is measurable and $e_{ij} = 0$ unmeasurable. Then the Euclidean distance matrix measured in practice can be represented as $\boldsymbol{D}_{obs} = (\tilde{d}_{ij}^2), \tilde{d}_{ij} = (d_{ij} + n_{ij}) \times e_{ij}$.

At this point, there are some unknown entries in the Euclidean distance matrix. Therefore, the problem of the autonomous establishment of the pseudolite system spatial coordinate can be summed up to recover the complete Euclidean distance matrix according to partial distances between k pseudolites in the pseudolite system, and then obtain the relative position of each pseudolite, and establish the pseudolite system spatial coordinate.

In order to overcome the shortcomings of traditional algorithms, such as low precision, high complexity and inability to deal with more missing distances, in next section, we propose a new technique based on the shortest path algorithm and the nonlinear conjugate gradient method to establish the local positioning system spatial coordinate.

3 The Proposed Algorithm

We consider the minimum variance of the Euclidean distance matrix $\|\boldsymbol{D} - \boldsymbol{D}_{obs}\|_F^2/2$ as the optimization goal. The relationship between the Euclidean distance matrix $\boldsymbol{D}$ and the coordinate matrix $\boldsymbol{X}$ is $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}})\boldsymbol{I}^{\mathrm{T}} + \boldsymbol{I}\mathrm{diag}(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}})^{\mathrm{T}} - 2\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}$ where $\mathrm{diag}(\cdot)$ represents the vector containing the diagonal entries of the matrix, so the optimization goal is reformulated as

$$f(\boldsymbol{X}) = \left\|\mathcal{P}_E(\sigma(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}) - \boldsymbol{D}_{obs})\right\|_F^2/2 \tag{1}$$

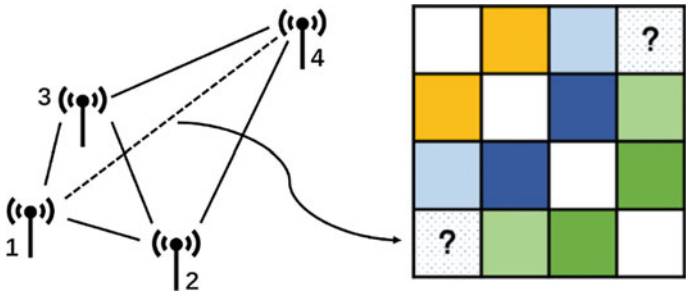


Fig. 1 The diagram of Euclidean distance matrix with unknown entries

where $\mathcal{P}_E(\cdot)$ means assigning unknown matrix elements 0 and $\sigma(\boldsymbol{Y}) = \text{diag}(\boldsymbol{Y})\boldsymbol{I}^{\mathrm{T}} + \boldsymbol{I}\text{diag}(\boldsymbol{Y})^{\mathrm{T}} - 2\boldsymbol{Y}$.

This problem is an unconstrained nonlinear problem, so the nonlinear conjugate gradient method is chosen as an optimal algorithm for this problem. The key of nonlinear conjugate gradient method is to construct a group of mutually conjugate direction, in order to accelerate the efficiency of descent of cost function, and the conjugate direction can be constructed by the combination of the last search direction and the gradient direction at the current point. The nonlinear conjugate gradient method has the characteristics of global convergence, which not only overcomes the shortcoming of slow convergence of steepest descent method, but also avoids the disadvantage that Newton method needs to compute the Hessian matrix and larger demand for memory.

Considering that the cost function is a nonconvex function when $p < k - 1$, which is the general case, it is necessary to estimate the missing entries in the Euclidean distance matrix in advance, which makes the optimization algorithm have a good initial value. By comparing the results of different initialization strategies, this paper selects the value of shortest path between unmeasurable pseudolite pairs as the initial estimate of unknown distances.

3.1 Expression of Gradient Direction

In the nonlinear conjugate gradient method, the conjugate direction is computed from the negative gradient direction at the current point and the last search direction in each iteration. In this section, the expression of the gradient of the cost function is deduced.

The directional derivative of a function $f(\boldsymbol{X})$ along a given vector $\boldsymbol{H}$ at a given point $\boldsymbol{X}$ is defined as $\mathrm{D}f(\boldsymbol{X})[\boldsymbol{H}]$. The relationship between the directional derivative $\mathrm{D}f(\boldsymbol{X})[\boldsymbol{H}]$ and Euclidean gradient $\nabla f(\boldsymbol{X})$ is $\mathrm{D}f(\boldsymbol{X})[\boldsymbol{H}] = \langle \nabla f(\boldsymbol{X}), \boldsymbol{H}\rangle$, so we deduce the expression of gradient by find a $\boldsymbol{V}$ that satisfies $\mathrm{D}f(\boldsymbol{X})[\boldsymbol{H}] = \langle \boldsymbol{V}, \boldsymbol{H}\rangle$.

Note that the cost function $f(\boldsymbol{X})$ can be expressed as $f(\boldsymbol{X}) = \|h(\boldsymbol{X})\|_F^2/2$ where $h(\boldsymbol{X}) = \mathcal{P}_E(\sigma(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}) - \boldsymbol{D}_{obs})$, and $(\boldsymbol{X}+t\boldsymbol{H})(\boldsymbol{X}+t\boldsymbol{H})^{\mathrm{T}}$ can be expressed as $\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} + \boldsymbol{\Delta}$ where $\boldsymbol{\Delta} = t\boldsymbol{X}\boldsymbol{H}^{\mathrm{T}} + t\boldsymbol{H}\boldsymbol{X}^{\mathrm{T}} + t^2\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$, then we have

$$\begin{aligned} f(\boldsymbol{X}+t\boldsymbol{H}) &= \left\langle \mathcal{P}_E(\sigma(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} + \boldsymbol{\Delta}) - \boldsymbol{D}_{obs}), \sigma(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} + \boldsymbol{\Delta}) - \boldsymbol{D}_{obs}\right\rangle/2, \\ f(\boldsymbol{X}) &= \left\langle \mathcal{P}_E(\sigma(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}) - \boldsymbol{D}_{obs}), \sigma(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}) - \boldsymbol{D}_{obs}\right\rangle/2. \end{aligned} \tag{2}$$

So when $t \to 0$, the increment of the cost function along vector $\boldsymbol{H}$ at point $\boldsymbol{X}$ is

$$\begin{aligned} & f(\boldsymbol{X}+t\boldsymbol{H}) - f(\boldsymbol{X}) \\ &= \left\langle \mathcal{P}_E(\sigma(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}) - \boldsymbol{D}_{obs}), \sigma(\boldsymbol{\Delta})\right\rangle + \langle \mathcal{P}_E(\sigma(\boldsymbol{\Delta})), \sigma(\boldsymbol{\Delta})\rangle/2 \\ &= \left\langle h(\boldsymbol{X}), \sigma(t\boldsymbol{X}\boldsymbol{H}^{\mathrm{T}} + t\boldsymbol{H}\boldsymbol{X}^{\mathrm{T}})\right\rangle + o(t) \\ &= 2t\left\langle h(\boldsymbol{X}), \sigma(\mathrm{Sym}(\boldsymbol{H}\boldsymbol{X}^{T}))\right\rangle + o(t), \end{aligned} \tag{3}$$

where Sym$(\cdot)$ represents the symmetric part of a matrix.

According to the definition of directional derivative, we have

$$\begin{aligned} &\mathrm{D}\, f(\boldsymbol{X})[\boldsymbol{H}] \\ &= \lim_{t\to 0} \frac{f(\boldsymbol{X}+t\boldsymbol{H}) - f(\boldsymbol{X})}{t} \\ &= 4\left\langle h(\boldsymbol{X}), \mathrm{Sym}(\mathrm{diag}(\mathrm{Sym}(\boldsymbol{H}\boldsymbol{X}^{\mathrm{T}}))\boldsymbol{1}^{\mathrm{T}})\right\rangle - 4\left\langle h(\boldsymbol{X}), \mathrm{Sym}(\boldsymbol{H}\boldsymbol{X}^{\mathrm{T}})\right\rangle \\ &= \langle 4(\mathrm{eye}(\mathrm{Sym}(h(\boldsymbol{X}))1) - \mathrm{Sym}(h(\boldsymbol{X})))\boldsymbol{X}, \boldsymbol{H}\rangle \end{aligned} \tag{4}$$

where eye$(\cdot)$ represents the squared matrix whose diagonal is the vector.

In general $\boldsymbol{D}_{obs}$ is symmetric, then $h(\boldsymbol{X})$ is symmetric too and the expression of Euclidean gradient is $\nabla f(\boldsymbol{X}) = 4(\mathrm{eye}(h(\boldsymbol{X})\boldsymbol{1}) - h(\boldsymbol{X}))\boldsymbol{X}$.

3.2 Algorithm Steps

To sum up, the overall algorithm proposed in this paper is as follows.

Algorithm Spatial Coordinate Autonomous Establishment Technique for High-precision Local Positioning System

Input: noisy and partially missing Euclidean distance matrix $\boldsymbol{D}_{obs}$, cost function $f(\boldsymbol{X})$, threshold for step size $\in_1$, threshold for the norm of the search direction $\in_2$

Output: absolute positions of pseudolites $\boldsymbol{X}$

1. Employ Floyd-Warshall algorithm to get the shortest paths to estimate the missing elements in $\boldsymbol{D}_{obs}$ and get $\hat{\boldsymbol{D}}$;
2. Employ MDS algorithm on $\hat{\boldsymbol{D}}$ to get initial position estimate $\boldsymbol{X}_0$;
3. Employ nonlinear conjugate gradient method to get the convergent solution;
   repeat
   - 3.1 Compute negative gradient direction $\Delta\boldsymbol{X}_n = -\nabla f(\boldsymbol{X}_n)$;
   - 3.2 Compute conjugate direction $\boldsymbol{P}_n = \begin{cases} \Delta\boldsymbol{X}_n & n = 0 \\ \Delta\boldsymbol{X}_n + \beta_n \boldsymbol{P}_{n-1} & \text{otherwise} \end{cases}$;
   - 3.3 Perform Armijo line search and get step size α_n;
   - 3.4 Update $\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \alpha_n \boldsymbol{P}_n$;

   until $\alpha_n < \in_1$ or $\|\boldsymbol{P}_n\|_F < \in_2$

4 Simulation

In this section, we test the proposed technique, called FWCG for short, and other existing algorithms on algorithm performance and running time, including Optspace [2], SDR [5], conjugate gradient method on the manifold, called CGRM for short [7] and gradient descent algorithm.

Simulations are performed in a unit square area, with k pseudolites laid randomly, assuming the detection region of every pseudolite is a circular area whose centralized radius is R. The distances between pseudolites are unmeasurable and then unknown if they are more than R, and others are accompanied with the gaussian noise with a variance of λ^2. Then the proposed algorithm and other algorithms are tested, matching the relative position with the actual position, and calculating the centralized average distance error of all pseudolites. The Optspace and SDR don't need the initial value. In addition, it is ensured that there is no pseudolite whose number of connections is less than $p+1$ in all simulated pseudolite networks to avoid the possibility of multiple structures of space coordinate.

Figure 2 shows the results of the proposed algorithm when $\lambda = 0.01$. The proposed FWCG can ensure that the deviation of the estimated location of each pseudolite is acceptable. More pseudolites can support the smaller detection range of each pseudolite, which means the stronger ability to resist missing distances.

Then we test the result of FWCG and other four algorithms in the noise-free environment when R is set to different values and $k = 20$. As Fig. 3 shows, the centralized average error of Optspace is always around 0.1, which is obviously poor than other algorithms, and whose convergent speed is slow meanwhile. The centralized average error of the gradient descent method decreases slowly with the

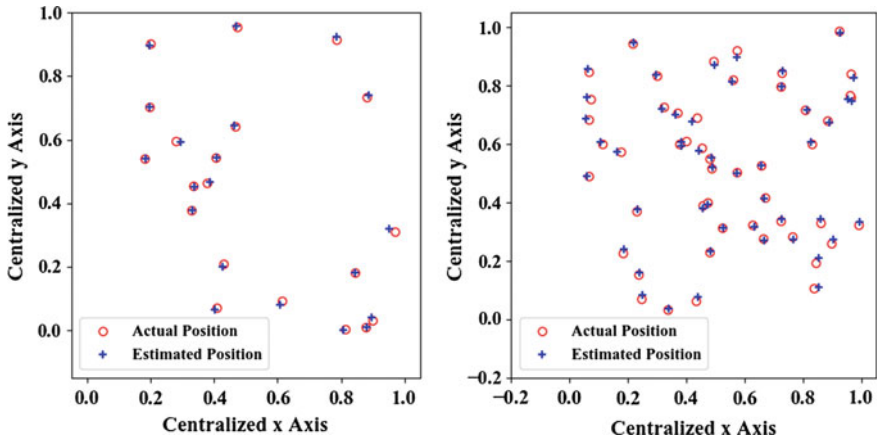


Fig. 2 The recovery result of FWCG (Left: the centralized average error is 7.61×10^{-3} when $\lambda = 0.01$, R = 0.5, k = 20. Right: the centralized average error is 9.76×10^{-3} when $\lambda = 0.01$, R = 0.3, k = 50)

increase of R and remains at the magnitude of 0.01. FWCG and CGRM perform closely, and when the detection distance is large, FWCG, CGRM and SDR can accurately establish the local positioning system space coordinate. In comparison, the proposed FWCG can achieve higher precision.

Because the performances of FWCG, SDR and CGRM are more outstanding than Optspace and gradient descent, following simulation mainly focuses on these three algorithms.

We test the positioning accuracy of three algorithms versus different noise standard deviation when $R = 0.7, k = 20$. As shown in Fig. 4, the positioning accuracy of three algorithms is proportional to the noise standard deviation. The error difference between the proposed algorithm FWCG and CGRM is very small. The performances of both are slightly better than that of SDR. Figure 5 shows the running time of three algorithms in $R = 0.5, k = 20$ and $R = 0.5, k = 100$, it is obvious that FWCG is much faster than the other two algorithms. The computation burden of FWCG is about 1/7 of that of CGRM. When $k = 20$, the computation burden of FWCG is about 1/18 of that of SDR, and is about 1/170 of that of SDR when $k = 100$. On the one hand, because of the large increase of the computation and storage of SDR in the case of larger number of pseudolites, the testing computer may not even meet the demand for memory. On the other hand, in the process of optimization, CGRM requires a large number of SVD operations. In contrast, the proposed FWCR only need matrix addition and multiplication, the computational burden is far less than the other two algorithms, so FWCR is also suitable as a real-time spatial coordinate autonomous establishment algorithm.

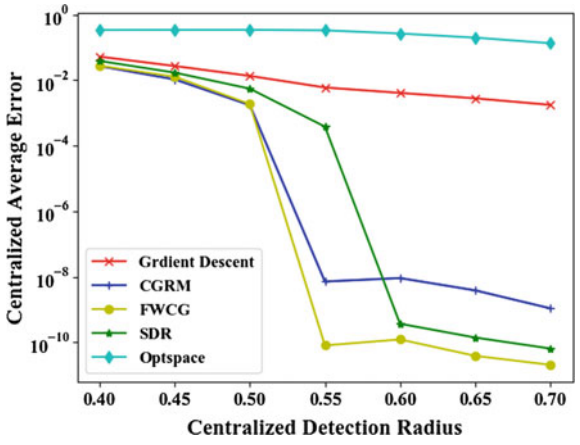


Fig. 3 Curves of centralized average distance error of algorithms versus centralized detection range when λ = 0, k = 20

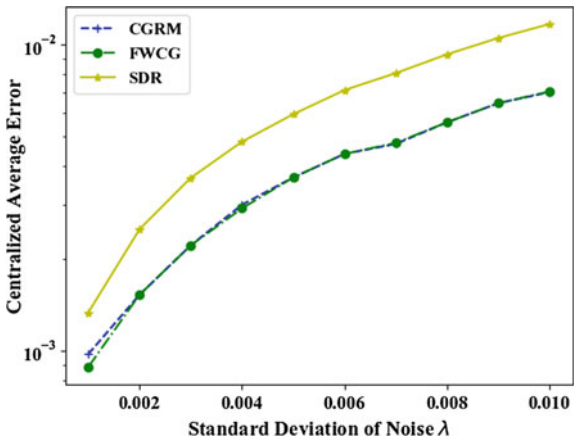


Fig. 4 Curves of centralized average distance error of algorithms versus standard deviation of noise when R = 0.7, k = 20

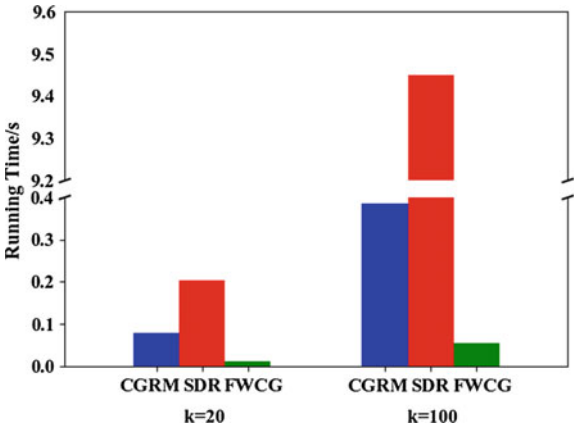


Fig. 5 Comparison of running time of algorithms when k = 20 and k = 100

5 Conclusion

In this paper, we design the technique for the autonomous establishment of the local positioning system spatial coordinate in high precision to handle the missing ranging information in the network. The technique models the problem as the Euclidean distance matrix completion problem on the principle of minimum variance, strictly deducing the expression of the gradient of the cost function, taking the value of the shortest path as the initial value for unknown distances and using the nonlinear conjugate gradient method to get the convergent solution. The relative position of each pseudolite and the pseudolite network constellation is recovered according to the partial ranging information between the pseudolite base stations.

Furthermore, analyzing in aspects of coordinate establishment accuracy and algorithm complexity, compared with other existing algorithms, the proposed algorithm not only has higher positioning accuracy but also can handle the problem

that there are missing distances in the network, and the performance is elevated as the number of pseudolites increases. Moreover, the algorithm proposed has a less computational complexity and is suitable for a real-time algorithm to be used in practice. It can significantly simplify the laying costs, shorten the deployment time, reduce maintenance difficulty of the pseudolite system and greatly improve the mobility and flexibility of the system in applications.

References

1. Drineas P, Javed A, Magdon-Ismail M, Pandurangant G, Virrankoski R, Savvides A (2006) Distance matrix reconstruction from incomplete distance information for sensor network localization. In: 2006 3rd annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON'06), vol 2. IEEE, pp 536–544, 28 Sept 2006
2. Taghizadeh MJ, Parhizkar R, Garner PN, Bourlard H, Asaei A (2015) Ad hoc microphone array calibration: Euclidean distance matrix completion algorithm and theoretical guarantees. Sig Process 28(107):123–140
3. Biswas P, Ye Y (2004) Semidefinite programming for ad hoc wireless sensor network localization. In: Proceedings of the 3rd international symposium on information processing in sensor networks. ACM, pp 46–54, 26 Apr 2004
4. Ding Y, Krislock N, Qian J, Wolkowicz H (2010) Sensor network localization, Euclidean distance matrix completions, and graph realization. Optim Eng 11(1):45–66
5. Dokmanic I, Parhizkar R, Ranieri J, Vetterli M (2015) Euclidean distance matrices: essential theory, algorithms, and applications. IEEE Sig Process Mag 32(6):12–30
6. Shang Y, Ruml W, Zhang Y, Fromherz MP (2003) Localization from mere connectivity. In: Proceedings of the 4th ACM international symposium on mobile ad hoc networking & computing. ACM, pp 201–212, 1 June 2003
7. Nguyen L, Kim S, Shim B (2016) Localization in internet of things network: matrix completion approach. In: Information theory and applications workshop (ITA). IEEE, pp 1–4, 31 Jan 2016