Lecture Notes in Electrical Engineering 438

Jiadong Sun
Jingnan Liu
Yuanxi Yang
Shiwei Fan
Wenxian Yu
Editors

China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II

Springer

Lecture Notes in Electrical Engineering

Volume 438

About this Series

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

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

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

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

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

Jiadong Sun · Jingnan Liu · Yuanxi Yang
Shiwei Fan · Wenxian Yu
Editors

China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II

Springer

Editors
Jiadong Sun
Academician of CAS
China Aerospace Science and Technology Corporation
Beijing
China

Jingnan Liu
Wuhan University
Wuhan
China

Yuanxi Yang
National Administration of GNSS and Applications
Beijing
China

Shiwei Fan
China Satellite Navigation Office
Beijing
China

Wenxian Yu
Shanghai Jiao Tong University
Shanghai
China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-10-4590-5 ISBN 978-981-10-4591-2 (eBook)
DOI 10.1007/978-981-10-4591-2

Library of Congress Control Number: 2017937524

Printed on acid-free paper

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

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 (GNSS) worldwide. The BDS will provide highly reliable and precise positioning, navigation and timing (PNT) services as well as short-message communication for all users under all-weather, all-time and worldwide conditions.

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

The 8th China Satellite Navigation Conference (CSNC 2017) is held during May 23–25, 2017, Shanghai, China. The theme of CSNC2017 is Positioning, Connecting All, including technical seminars, academic exchanges, forums, exhibitions and lectures. The main topics are as follows:

Conference Topics

S01 Satellite Navigation Technology
S02 Navigation and Location Service
S03 Satellite Navigation Signals and Signal Processing
S04 Satellite Orbit and Satellite Clock Error
S05 Precise Positioning Technology
S06 Atomic Clock and Time-frequency Technology
S07 Satellite Navigation Augmentation Technology
S08 Test and Assessment Technology
S09 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 (Lecture Notes in Electrical Engineering) have 188 papers in twelve topics of the conference, which were selected through a strict peer-review process from 599 papers presented at CSNC2017. In addition, another 272 papers were selected as the electronic proceedings of CSNC2017, 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 249 referees and 48 session chairmen who are listed as members of editorial board. The assistance of CNSC2017 organizing committees and the Springer editorial office is highly appreciated.

Beijing, China — Jiadong Sun
Wuhan, China — Jingnan Liu
Beijing, China — Yuanxi Yang
Beijing, China — Shiwei Fan
Shanghai, China — Wenxian Yu

Editorial Board

Topic: S5: Precise Positioning Technology

Qile Zhao, Wuhan University
Jianwen Li, Information Engineering University
Songshu Li, Shanghai Astronomical Observatory, Chinese Academy of Sciences
Feng Yanming, Queensland University of Technology Brisbane, Australia

Topic: S6: Atomic Clock and Time-frequency Technology

Lianshan Gao, The 203th Research Institute of China Aerospace Science & Industry Corporation
Chunhao Han, Beijing Satellite Navigation Center
Xiaohui Li, National Time Service Center, Chinese Academy of Sciences
Rochat Pascal, Spectratime & T4Science

Topic: S7: Satellite Navigation Augmentation Technology

Junlin Zhang, OLinkStar Co., Ltd., China
Jinping Chen, Beijing Satellite Navigation Center
Rui Li, Beihang University
Shaojun Feng, Imperial College London

Topic: S8: Test and Assessment Technology

Xiaolin Jia, Xi'an Institute of Surveying and Mapping
Jun Yang, National University of Defense Technology
Jianping Cao, Air Force Equipment Research Institute
Yang Gao, University of Calgary, Canada

Topic: S9: User Terminal Technology

Haibo He, Beijing Satellite Navigation Center
Baowang Lian, Northwestern Polytechnical University
Hong Li, Tsinghua University
Yong Li, University of New South Wales, Australia

Topic: S10: Multi-source Fusion Navigation Technology

Zhongliang Deng, Beijing University of Posts and Telecommunications
Hong Yuan, Academy of Opto-electronics, Chinese Academy of Sciences
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

Mingquan Lu, Tsinghua University
Wei Wang, The 20th Research Institute of China Electronics Technology Group Corporation
Yin Xu, Academy of Opto-Electronics, Chinese Academy of Sciences
Xiangzhen Li, Chungnam National University, Korea

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

Daiping Zhang, China Defense Science and Technology Information Center
Yonggang Wei, China Academy of Aerospace Standardization and Product Assurance
Junli Yang, Beihang University
Huiying Li, Electronic Intellectual Property Center, Ministry of Industry and Information Technology, PRC

Scientific Committee and Organizing Committee

The 8th China Satellite Navigation Conference (CSNC 2017)

Scientific Committee

Chairman
Jiadong Sun, China Aerospace Science and Technology Corporation

Vice-Chairman
Rongjun Shen, China
Jisheng Li, China
Qisheng Sui, China
Changfei Yang, China
Zuhong Li, China Academy of Space Technology
Shusen Tan, Beijing Satellite Navigation Center, China

Executive Chairman
Jingnan Liu, Wuhan University
Yuanxi Yang, China National Administration of GNSS and Applications
Shiwei Fan, China

Committee Members: (By Surnames Stroke Order)
Xiancheng Ding, China Electronics Technology Group Corporation
Qingjun Bu, China
Liheng Wang, China Aerospace Science and Technology Corporation
Yuzhu Wang, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences
Guoxiang Ai, National Astronomical Observatories, Chinese Academy of Sciences
Shuhua Ye, Shanghai Astronomical Observatories, Chinese Academy of Sciences
Zhaowen Zhuang, National University of Defense Technology
Qifeng Xu, PLA Information Engineering University

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

Organizing Committee

Secretary General
Haitao Wu, Navigation Headquarters, Chinese Academy of Sciences

Vice-Secretary General
Weina Hao, Navigation Headquarters, Chinese Academy of Sciences

Under Secretary
Wenhai Jiao, China Satellite Navigation Office Engineering Center
Zhao Wenjun, Beijing Satellite Navigation Center
Wenxian Yu, Shanghai Jiao Tong University
Wang Bo, Academic Exchange Center of China Satellite Navigation Office

Members: (In Surname Stroke Order)
Qun Ding, The 20th Research Institute of China Electronics Technology Group Corporation
Miao Yu, China Academy of Space Technology
Li Wang, International Cooperation Research Center, China Satellite Navigation Engineering Office
Liu Peiling, Shanghai Jiao tong University
Ying Liu, China Satellite Navigation Office Engineering Center
Lu Mingquan, Tsinghua University
Xiuwan Chen, Peking University
Ouyang Guangzhou, Academy of Opto-Electronics, Chinese Academy of Sciences
Zhao Qile, Wuhan University

Xiangan Zhao, China Defense Science and Technology Information Center
Gang Hu, Beijing Unicore Communications, Inc.
Min Shui, National Remote Sensing Centre of China

Contents

Part II Satellite Navigation Augmentation Technology

Part I
Satellite Navigation Signals and Signal Processing

Research on Communication Delay Model for Narrow Beam Inter-satellite Links in a TDMA System

Daoning Yang, Ying Liu, Gang Li, Jun Yang, Qian Zhao and Zhixi Yang

Abstract Navigation constellation inter-satellite link (ISL) is one of the efficient ways to improve the service performance of global navigation satellite system (GNSS). The four major GNSSs have been equipped with ISLs or plan to equip in near future. Narrow beam inter-satellite link in a time division multiple access (TDMA) system have advantages of fast beam switching, flexible networking, high communication speed and good security. For navigation constellation to improve service performance, the measurement values have to transfer back to the ground facilities in a short time, at the same time, facilities have to transfer telecommand and uplink ephemeris to satellites which are invisible to ground facilities as soon as possible. These all made a request to shorter communication delay. And it is very important to improve the performance of ISLs by rapidly assessing the quality of the connection and adjusting the unreasonable connection in time. In this paper, a communication time delay model for ISLs based on directed graphs is studied. By abstracting the ISL timeslot table into a directed graph, and based on the shortest path algorithm, the communication time delay can be obtained quickly. Simulation verifications have been carried out for different scenes. Results show that the proposed model works well.

Keywords Inter-satellite link · Time division multiple access (TDMA) · Delay · Directed graph

D. Yang (✉) · J. Yang · Z. Yang
College of Mechatronics Engineering and Automation,
National University of Defense Technology, Changsha 410073, China
e-mail: Yang_daoning@hotmail.com

Y. Liu · G. Li
Beijing Institute of Tracking and Telecommunications Technology,
Beijing 100094, China

Q. Zhao
National Time Service Center, CAS, Xi'an 710600, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_1

1 Introduction

The Navigation constellation ISL is one of the efficient ways to improve GNSS service performance, the four major GNSSs have been equipped with ISLs or plan to equip in near future. Global Navigation System (GPS) which has already been equipped wide beam ISL in UHF band is planning to be equipped with V or Ka band narrow beam ISL in the next stage as well as Galileo system [1–4]. Ka band narrow beam ISL is also planned to equip BeiDou system which will supply global navigation service in 2020 [5]. Research aiming to the implement of the ISL technology is going on with the recent launched new generation satellites of BeiDou system.

ISL is one of the important infrastructure of GNSS in the future. Time-divided narrow beam ISL is the development direction for GNSS because of its fast beam switching, flexible networking, high communication speed and good security [6, 7]. Navigation constellation can complete the inter-satellite measurement and communication quickly in this system.

In order to improve the service performance, measurement data need to be transmitted to ground facilities in a short time. At the same time, facilities have to transfer telecommand and uplink ephemeris to satellites which are invisible to ground facilities as soon as possible. So these all made a request to shorter communication delay. Timeslot is a kind of important resource in time-divided system, and the design of timeslot and route decides the performance of satellite communication, so it is need to be estimated to optimize the system performance. In this paper, a communication time delay model for ISLs based on directed graphs is studied. By abstracting the ISL timeslot table into a directed graph, and based on the shortest path algorithm, the communication time delay can be obtained quickly.

2 Timeslot and Route

2.1 Timeslot and Superframe

In the time-divided system, continuous time is divided into timeslots, a satellite may build different link with different target in different timeslots, so that it can achieve multi-access connection and realize communication with different targets. Timeslot is the basic unit that TDMA ISL can build and maintain.

The visibility to others of a satellite varies real time, and it is the basic restraint because ISL can only be built between two satellites which are visible to each other. To simplify our analysis, we divide the continuous time into superframes, and there are several timeslots in every superframe. During a superframe, the visibility relationship in the whole satellite constellation is considered to be invariable. Only if two satellites are visible to each other in the whole superframe, we consider the two satellites are visible in this superframe.

2.2 Timeslot Table

Timeslot tables describe the slot scheme. The row of a table represents satellite index in the constellation and the rank is timeslot counter in a superframe. Consider that the number of satellites in the constellation is m, and there are n timeslots in a superframe, then an $m \times n$ orders matrix S is used to represents a timeslot table. The element $S(i, j)$ represents the target that satellite i establishes the inter-satellite link with in slot j. The ISL is dual-directional, so when

$$S(i,j) = r \tag{1}$$

It is obvious that

$$S(r,j) = i \tag{2}$$

A timeslot table is shown in Fig. 1 and three satellites within 4 time slots are listed.

2.3 Design of Route

The route of our ISL scheme used is minimum delay route because the connection target may varies in every timeslot. When a satellite A has the requirement to communicate with the target satellite P in timeslot T, it judges whether the current

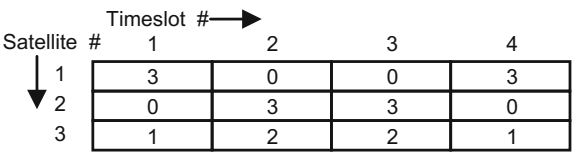


Fig. 1 Example of timeslot table

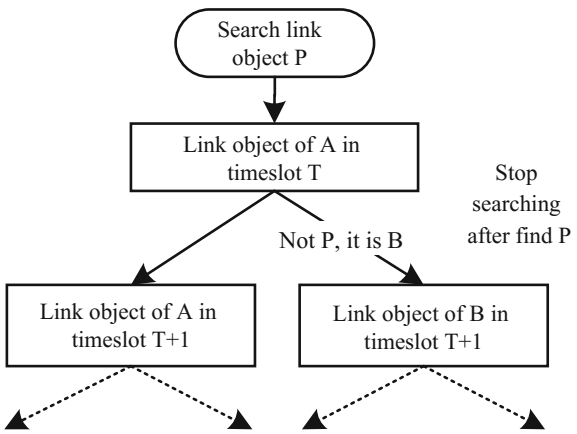


Fig. 2 Example of the route searching process

target B is the destination node P. If is, it will transport to B directly, if not, it will look up the timeslot table for the T + 1 target and the next target of B to find the destination P. Repeat the step until P is reached. The process is shown in Fig. 2.

3 Time Delay Model

3.1 Directed Graph Model

Obviously, time delay of two satellites can be obtained from the timeslot table because the table describes the connection relationship in every timeslot. Directed graph is abstracted from timeslot table. The nodes of the directed graph represent the satellites in timeslots, dual-directional vectors represent the communication links established with other satellites and single-directional vectors represent time the data spent in a satellite when the transmission is not implemented.

As mentioned above, directed graph shown in Fig. 3 is abstracted from the timeslot table shown in Fig. 1. It is notice that there are no vectors come from node 2 to 4, that is to say that satellite node 2 does not establish a link with other satellites in slot 4. So the transport has to be deferred to the next slot and dotted line is used to describe this.

3.2 Shortest Path Algorithm

Shortest path algorithm is a classic algorithm in graph theory. The original purpose of the algorithm is to find the shortest distance of two nodes. There are two main algorithms that are usually used, Dijkstra algorithm and Bellman-Ford algorithm. We will use Dijkstra algorithm to calculate the shortest delay between two satellite nodes. The basis of the algorithm is optimum factor character which is a theorem that if the shortest path from node i to node j is $P(i,j) = \{V_i, \ldots V_k, \ldots V_s, \ldots V_j\}$, then path $P(k,s)$ is the shortest path between node k and node s who are in the path i to j. We can calculate the shortest path between two nodes base on the theorem.

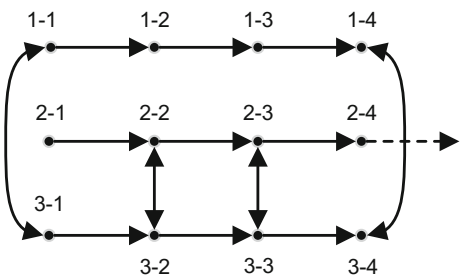


Fig. 3 Directed graph time delay model

In the directed graph, every timeslot of a satellite is a node. The source node and timeslot is confirmed when we want to obtain the communication time delay of two satellites and the destination node is unknown because of the unknown time delay. The adjacent nodes of a same satellite is connected, so if there is a shortest path between source node $i - m$ and destination node $j - n$ passing $j - (n - 1)$ nodes, then the path from node $i - m$ to node $j - (n - 1)$ is a shorter one that node i passes to j. The first step to find the shortest path between i and j is to set the destination as the node j at current timeslot, if the path is not exist, set the node j at next timeslot as the destination node. Repeat the step until the shortest path is found. The number of directional-vectors that the path goes through is the communication time-delay.

4 Simulation and Analysis

In this section, we simulate the navigation satellite constellation and analyze the result to verify the practicability of the model.

4.1 Simulation Scene

We simulate the BeiDou Navigation System which will contains 24 Medium Earth Orbit (MEO) satellites, 3 Geosynchronous Orbit (GEO) satellites and 3 Inclined Geosynchronous Orbit (IGSO) satellites when completely established. The orbit parameters are listed in Table 1 [8]. There are 3 ground facilities in China.

Every satellite equips an ISL narrow beam antenna. The ISL is dual-directional link. The simulation time is set 7 days (10,080 min). Every timeslot is 3 s and every superframe is 60 s, that is to say there are 20 timeslots in a superframe [9, 10].

Firstly, we generate 10,080 timeslot tables randomly based on the visibilities between satellites. Directed graphs which are abstracted from timeslots connect together to one. We analyze the time delay that the invisible satellites invisible to the ground facilities downlink telemetry data back to the ground in 7 days. In the simulation we consider the satellite-ground time delay is 0.

After that, we adjusted the link relationship with large time delay and regenerated the optimized timeslot tables. Then we analysed the performance of the optimized timeslot tables with same method.

Table 1 Orbit parameters of satellites in constellation

Satellite type	Orbit parameters
MEO	Walker 24/3/1 satellite constellation, orbit height 21,528 km, incline 55°
GEO	Orbit height 35,786 km, longitude 80°, 110.5° and 140°
IGSO	Orbit height is 35,786 km, orbit incline 55°, phase interval 120°

4.2 Simulation Result

The time delay of random-generated timeslot tables are shown in Fig. 4. The result shows that there are less than 60% links delay within one timeslot and more than 30% links delay beyond 3 timeslot.

The time delay of the optimized timeslot tables are shown in Fig. 5. The result show that links whose delays beyond 3 timeslot are no longer exist and the number of links whose delays within one timeslot are up to 80%.

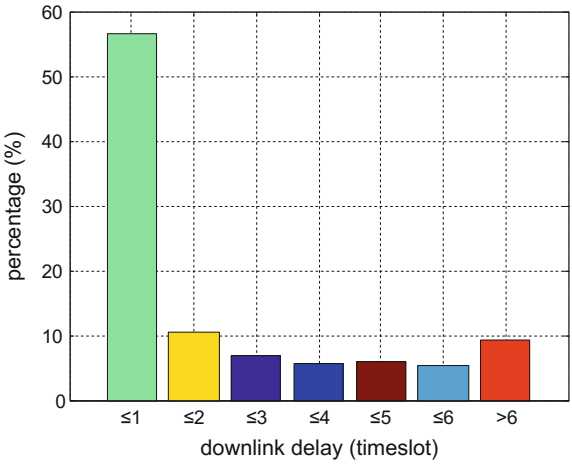


Fig. 4 The time delay of random scheme

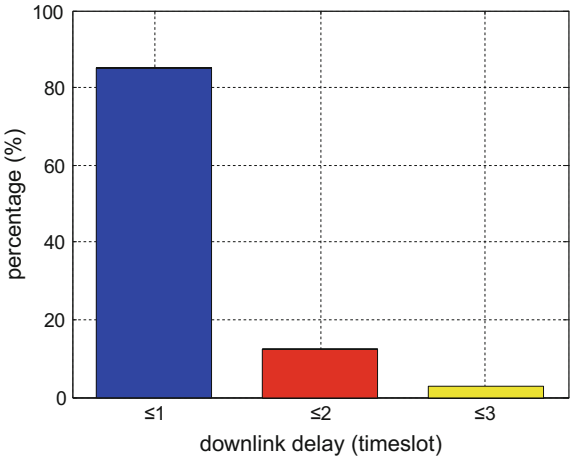


Fig. 5 The time delay of optimized scheme

5 Conclusion

This paper studies and simulates the inter-satellite link communication delay model based on directed graph. The result show that this model can obtain the communication delay of any two satellite nodes rapidly and accurately. It is also indicated that the model have the reference to estimate and adjust the connection relationship of ISL.

References

1. Fisher SC, Ghassemi K (1999) GPS IIF-the next generation. Proc IEEE 87(2):24–47
2. Maine KP, Anderson P, Langer J (2003) Crosslinks for the next-generation gps. In: IEEE aerospace conference proceedings
3. Fernández FA (2011) Inter-satellite ranging and inter-satellite communication links for enhancing GNSS satellite broadcast navigation data. Adv Space Res 47(5):786–801
4. Yang D, Yang J, Xu P (2016) GPS Solut. doi:10.1007/s10291-016-0587-0
5. Meng Y, Fan S, Yang Q et al (2015) Analysis of spacecraft orbit determination method using GNSS crosslink. Spacecraft Eng
6. Zheng K, Dong X, Yang Y et al (2014). A new time synchronization system based on the GEO ISLs. Geomatics Inf Sci Wuhan Univ 39(10):1163–1167
7. Li L, Geng G, Li Z (2016) Study of the development of the inter-satellite links in foreign GNSS. J Geomatics Sci Technol 33(2):133–138
8. BeiDou Navigation Satellite System Signal in Space Interface Control Document Open Service Signal (Version 2.0). China Satellite Navigation Office, Dec 2013
9. Wu G, Chen J, Guo X et al (2014) Design and simulation of time-slot allocation of inter-satellites links based on TDMA system. Comput Measur Control 22(12):4087–4090
10. Xu P, Chen J, Tang Y et al (2016) Time-slot allocation of inter-satellite links with low communication delay simulation and analysis. In: China satellite navigation conference

Satellite-Induced Multipath Analysis on the Cause of BeiDou Code Pseudorange Bias

Hailong Xu, Xiaowei Cui and Mingquan Lu

Abstract Data from previous observation have shown that the BeiDou satellite navigation system (BDS) has been suffering from satellite-induced elevation-dependent code pseudorange bias, which can be as much as 1 m and degrade the performance of high-precision applications such as Precise Point Positioning (PPP). Mechanism analysis on the cause of the bias has been absent in previous literature, which is essential to avoid the issue from happening again in following satellites. In this paper, both theoretical analyses and simulations are performed, proving that the bias can be caused by the multipath effect induced by the transmitting array on the satellite. Based on a proper model of this satellite-induced multipath effect, the bias values can be calculated through mathematical derivation. Under certain model parameters, the simulation results agree with the observed bias values very well. Methods and conclusions of this paper are useful in both investigating the current BeiDou bias issue and the process of design, development and test of following satellites by satellite manufactures.

Keywords BeiDou satellite navigation system · High-precision application, precise point positioning · Code pseudorange bias · Satellite-induced multipath · Correlation function

1 Introduction

With the system construction advancing and the application areas expanding, the BeiDou satellite navigation system (BDS) is becoming more and more important in both military and civilian fields. All of this is based on the condition that the system provides code pseudorange measure precise enough to end receivers. However, through the transmitting, propagating and receiving processes of the navigation signals, it is inevitable that biases will be induced to the pseudorange

H. Xu · X. Cui (✉) · M. Lu
Department of Electronic Engineering, Tsinghua University, Beijing, China
e-mail: cxw2005@mail.tsinghua.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_2

measurements. The bias sources include the satellite clock and orbit error, the ionosphere and troposphere delay, as well as the multipath effect on the receiver end. Features and correcting methods of these biases have been studied extensively in the past. However, in recent years, a new kind of pseudorange bias is observed in BDS, whose values are dependent on the elevation of the satellite from the view of the receiver and can be as large as 1 m. This bias is independent of the type, location, observation interval of the receiver, but holds differences between the two satellite groups, which are in medium earth orbits (MEO) and inclined geostationary orbits (IGSO), respectively, thus is supposed to be caused by the satellites [1]. This bias has little effect on absolute positioning applications due to the limited accuracy requirements, but can degrade the performance of some high-precision applications using code pseudorange measurements but without differencing techniques, such as Precise Point Positioning (PPP), thus restricts the application extension of the system.

Towards this issue, one countermeasure is to determine the bias values under different satellite elevations from historical observations, then perform compensation in the receiver [1]. Another countermeasure is to adjust the broadcasted ephemeris, and use the orbit shift calculated from the ephemeris with respect to the true orbit to counteract the bias [2]. However, these solutions cannot eliminate the bias from the source. To solve the issue thoroughly, the cause of the bias must be found, and then be avoided strictly in the design and manufacture of following satellites.

This paper proves through theoretical analyses and simulations that, signal delays between different elements of the satellite antenna can cause bias similar to the observed one. These delays can be modelled as the satellite-induce multipath effect, in which the direct and multipath signals result in their respective correlation functions with different delays in the receiver, causing the peak of the composite correlation function to be distorted. Based on an accurate model of the satellite-induced multipath effect, the bias values can be calculated through mathematical derivation. Under certain parameters of the model, the simulated bias agree with the observed one very well. The methods and conclusions of this paper are helpful in the investigation of the BeiDou bias issue, as well as in the design, development and test of the following satellites by satellite manufactures.

2 Transmitting Array Modelling

Navigation satellites transmit signals towards the earth from orbits of certain altitudes, just as presented in Fig. 1. The off-boresight angle is defined as the angle between the line from the satellite to the earth center and the line from the satellite to the receiver. For a receiver on the earth's surface, the larger the satellite elevation, the smaller the off-boresight angle, and so is the distance between the satellite and the receiver. Thus, the propagation loss of the signal increases with the off-boresight angle. In order to let the signal reach the earth's surface with nearly

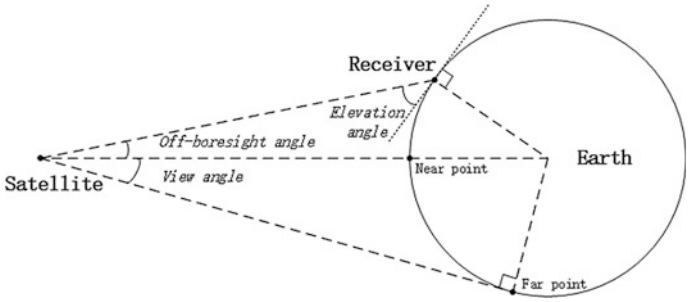


Fig. 1 Geometrical relationship of the satellite, the earth and the receiver

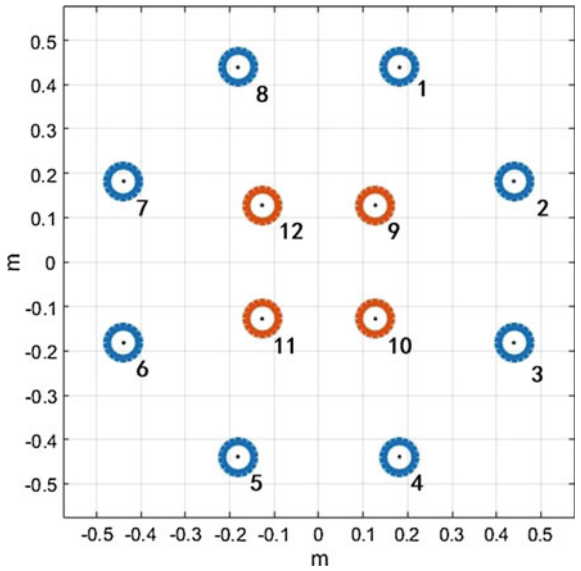


Fig. 2 Geometry of the GPS Block IIR transmitting array

equal power, the beam pattern of the transmitting antenna on the satellite should be shaped to form a dimple towards the earth's center [3].

In order to achieve this goal, multi-antenna arrays are used to transmit signals on the satellites [3, 4]. For example, Fig. 2 presents the array geometry of the GPS Block IIR satellite, including 8 equally distributed elements in the outer ring and 4 equally distributed elements in the inner ring [3]. The signal transmitting flow on the satellite is presented in Fig. 3. For beam pattern shaping, the navigation signal of a particular frequency is generated by the signal generation payload before being divided into 12 streams by the power dividing and phase shifting network, each of which is transferred to one array element to be transmitted. In the division, most of the power is allocated to the inner ring, forming a strong wide beam, while the rest is allocated to the outer ring, forming a weak narrow beam. Besides, the carrier

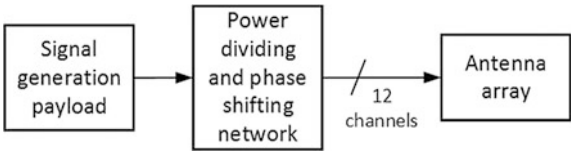


Fig. 3 Signal transmitting flow on the satellite

Table 1 BeiDou MEO satellite transmitting array parameters used in this paper

Parameters	Values
Radius of inner ring r_{inner}	18.03 cm
Radius of outer ring r_{outer}	47.50 cm
Power allocation of inner ring P_{inner}	95%
Power allocation of outer ring P_{outer}	5%
Phase shift of inner ring ϕ_{inner}	0°
Phase shift of outer ring ϕ_{outer}	180°

phase of the signal transmitted by the outer ring is shifted by 180° with respect to the inner ring. Thus, the composite beam pattern is obtained by subtracting the narrow beam from the wide beam, with a dimple towards the earth's center being formed.

The parameters of the BeiDou satellite transmitting antenna are not available in open literature, so it is assumed in this paper the BeiDou MEO satellite has an array with similar geometry and parameters with the GPS Block IIR satellite. Through a "try and see" method, it concludes that under the parameters presented in Table 1, the desired beam pattern can be obtained approximately. Based on these parameters, the weighting value corresponding to each element can be calculated by following equations:

$$w_{outer} = -\sqrt{P_{outer}/8}, \quad w_{inner} = \sqrt{P_{inner}/4}. \tag{1}$$

Then, under the isotropic element assumption, the transfer function of the array can be written by [5]

$$H(f,\theta) = \sum_{i=1}^{12} w_i \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\mathbf{p}_i\mathbf{a}(\theta)}, \tag{2}$$

where f denotes the frequency, λ denotes the corresponding wavelength, and θ denotes the off-boresight angle. Due to the rotational symmetry of the array geometry, the beam pattern variations along the same off-boresight angle are omitted. Index i ranging from 1 to 8 corresponds to the outer ring with $w_i = w_{outer}$, while index i ranging from 9 to 12 corresponds to the inner ring with $w_i = w_{inner}$. Vector $\mathbf{p}_i$ denotes the position of the ith element, and vector $\mathbf{a}(\theta)$ denotes the signal transmitting direction. Fixing f to be B1 (1561.098 MHz), B2 (1207.14 MHz) and B3 (1268.52 MHz), respectively, the beam pattern of the three frequencies can be

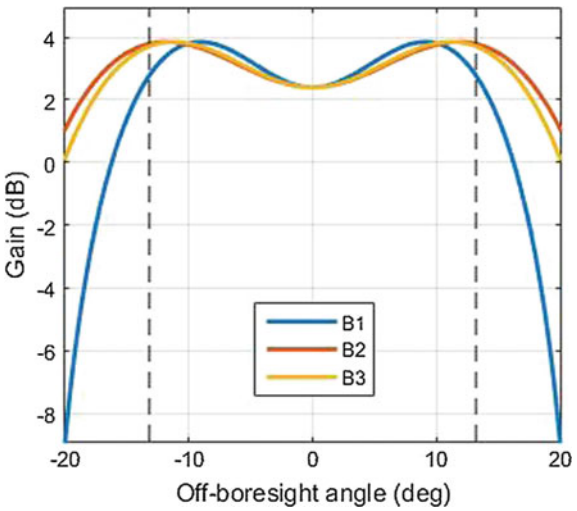


Fig. 4 Simulated beam patterns of the BeiDou MEO satellite transmitting array antenna. The *dotted lines* depict the view angle of the satellite towards the earth

obtained, which are presented in Fig. 4. It can be seen that the desired dimples are indeed formed, the depths of which are nearly identical to the propagation loss differences between the near point and far point in Fig. 1, proving the rationality of the parameters used. Following simulations will be performed under these parameters.

3 Satellite-Induced Multipath and Pseudorange Bias

In practice non-ideal factors always exist in the power dividing and phase shifting network depicted in Fig. 3, which can cause signal transfer delays between the different elements. These delays can be modeled as the satellite-induced multipath effect. Considering the two-ring array geometry, it is reasonable to assume that there is no delay difference between the elements within the same ring, but the signal transmitted by the outer ring is delayed by τ_d with respect to the inner ring. Thus, the signal out of the inner ring can be regarded as the direct path signal, while the signal out of the outer ring can be regarded as the multipath signal. The complex envelopes of them can be written as follows, respectively:

$$\tilde{s}_{inner}(t) = d(t)c(t)\mathrm{e}^{\mathrm{j}\phi_{inner}}, \tag{3}$$

$$\tilde{s}_{outer}(t) = \alpha d(t-\tau_d)c(t-\tau_d)\mathrm{e}^{\mathrm{j}\phi_{outer}}, \tag{4}$$

where $d(t)$ denotes the ephemeris bits, $c(t)$ denotes the pseudo-random ranging codes, ϕ_{inner} and ϕ_{outer} denote the carrier phases of the direct and the multipath signals, respectively, while α denotes the amplitude ratio of the multipath signal to

the direct signal. Assuming that within the observation interval, the ephemeris bit of the two signals are the same, then the complex envelope of the composite signal is given by

$$\tilde{s}(t) \approx d(t)(c(t) + \alpha c(t - \tau_d)\mathrm{e}^{\mathrm{j}\Delta\phi})\mathrm{e}^{\mathrm{j}\phi_{inner}}, \tag{5}$$

where

$$\Delta\phi = \phi_{outer} - \phi_{inner}. \tag{6}$$

At the receiver end, a delayed local code $c(t - \tau)$ is generated to correlate with the received signal, resulting in the correlation function to be

$$R'(\tau) = ||R(\tau) + \alpha\mathrm{e}^{\mathrm{j}\Delta\phi}R(\tau - \tau_d)||_2, \tag{7}$$

where $||\cdot||_2$ denotes the 2-norm operation, and $R(\tau)$ is the ideal correlation function of the pseudo-random ranging codes which is given by

$$R(\tau) = \begin{cases} (1 - |\frac{\tau}{T_c}|), & |\tau| \le T_c \\ 0, & |\tau| > T \end{cases}, \tag{8}$$

with T_c denoting the period of one code chip.

Figure 5 presents an instance of the correlation functions of the direct, multipath and composite signals when $\Delta\phi$ equals zero. Although the ideal correlation function $R(\tau)$ is symmetric and reaches its peak at $\tau = 0$, the composite correlation function $R'(\tau)$ is no longer symmetric and suffers a peak shift. In a typical receiver the Delay-Lock-Loop (DLL) is usually used to estimate and track the code phase through finding the peak of the correlation function [6]. With D denoting the early-minus-late interval of the DLL, let

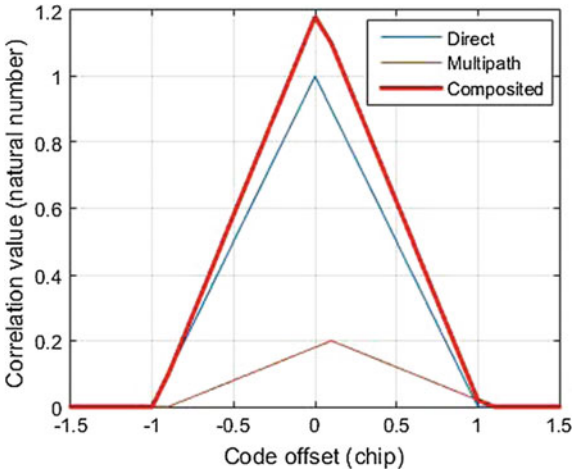


Fig. 5 An example of the direct, multipath and composite correlation functions

$$R'(\tau_e - D/2) = R'(\tau_e + D/2), \tag{9}$$

then τ_e is the code phase bias. As a result, the code pseudorange bias is calculated by

$$\rho_e = \mathrm{c} \cdot \tau_e, \tag{1.10}$$

where c is the speed of light.

The amplitude ratio α can be calculated from the array model. For this purpose, Eq. (2) is divided into two parts for the two rings, respectively:

$$H_{outer}(f, \theta) = \sum_{i=1}^{8} w_{outer} \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\mathbf{p}_i\mathbf{a}(\theta)}, \tag{11}$$

$$H_{inner}(f, \theta) = \sum_{i=9}^{12} w_{inner} \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\mathbf{p}_i\mathbf{a}(\theta)}. \tag{12}$$

Under the narrowband presumption [5], λ can be fixed to λ_c which is the wavelength at the central frequency of the signal, so the argument f can be omitted. If the center of the array geometry is at the antenna coordinate origin, then

$$\mathbf{p}_i = -\mathbf{p}_{i+4} (i = 1, 2, 3, 4), \quad \mathbf{p}_j = -\mathbf{p}_{j+2} (j = 9, 10). \tag{13}$$

Thus the transfer functions of the outer and inner rings can be reformed as

$$H_{outer}(\theta) = 2w_{outer} \sum_{i=1}^{4} \cos\left(\frac{2\pi}{\lambda_c}\mathbf{p}_i\mathbf{a}(\theta)\right), \tag{14}$$

$$H_{inner}(\theta) = 2w_{inner} \sum_{i=9}^{10} \cos\left(\frac{2\pi}{\lambda_c}\mathbf{p}_i\mathbf{a}(\theta)\right). \tag{15}$$

Therefore, α is obtained by

$$\alpha(\theta) = \frac{H_{outer}(\theta)}{H_{inner}(\theta)} \tag{16}$$

Figure 6 presents the transfer function variations of the inner and outer rings, as well as the amplitude ratio α over the off-boresight angle.

From equations from (7) to (1.10) it can be seen that the code pseudorange bias is determined by α, $\Delta\phi$ and τ_d. Parameter α is dependent on θ, while θ and the satellite elevation *el* are in one-to-one correspondence. Thus it concludes that the code pseudorange bias is dependent on the satellite elevation.

If α is small enough, then τ_e can be calculated by Braasch's approximate formula [7]:

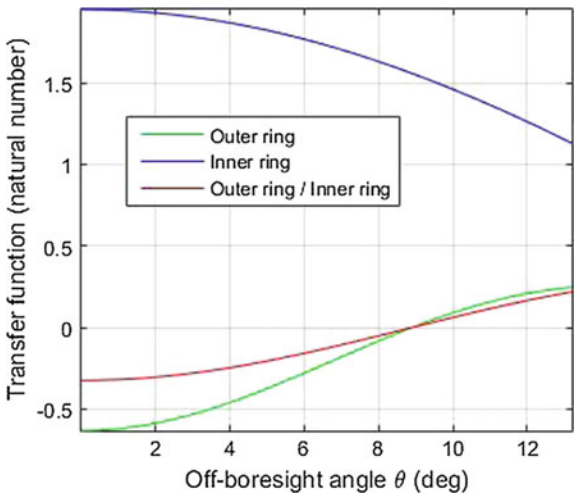


Fig. 6 Transfer function variations over the off-boresight angle for the inner and outer rings of the transmitting antenna array

$$\tau_e = \begin{cases} \frac{\beta\tau_d}{\beta+1} & 0 \le \tau_d \le \frac{D}{2}(\beta+1) \\ \frac{\beta D}{2} & \frac{D}{2}(\beta+1) \le \tau_d \le T_c + \frac{D}{2}(\beta-1) \end{cases}, \tag{17}$$

with

$$\beta = \alpha\cos(\Delta\phi). \tag{18}$$

The chip period of the ranging code in BDS is usually hundreds of nanoseconds, while D is usually set to $T_c/2$ or a bit smaller. Besides, it is reasonable to assume τ_d is about tens of nanoseconds or less. Then, the condition of the first branch in Eq. (17) is always satisfied, leading the calculation formula to be

$$\rho_e(\theta) = c \cdot \frac{\alpha(\theta)\cos(\Delta\phi)\tau_d}{\alpha(\theta)\cos(\Delta\phi)+1}. \tag{19}$$

After $\Delta\phi$ and τ_d are fixed, the code pseudorange bias under any elevation can be calculated using this equation.

4 Simulation

The B1 signal of BeiDou's MEO satellites is taken for example in this section, where simulation results of how the code pseudorange bias varies with the satellite elevation angle are given, and comparisons with the observed results given in [1] are made. The goal is to find proper parameter combinations in the parameter

domain of $\Delta\phi$ and τ_d, with which the simulated bias values agree with the observations very well.

Here we define a parameter called the bias dynamic as below

$$\Delta\rho_e = \rho_e|_{el=0^\circ} - \rho_e|_{el=90^\circ} \tag{20}$$

When the simulated and observed $\Delta\rho_e$ are the same, then it is highly possible that the simulated and observed bias values are very close to each other for each satellite elevations.

Figure 7 presents how $\Delta\rho_e$ varies with τ_d, with the observed $\Delta\rho_e$ depicted by the dotted red line. It can be seen that under different values of $\Delta\phi$, the "proper" values of τ_d under which the simulated and observed $\Delta\rho_e$ meet are also different. As a step further, Fig. 8 presents how the proper values of τ_d vary with $\Delta\phi$. Considering the even symmetry of $\cos(\Delta\phi)$ in Eq. (19) and assuming τ_d is positive, only the interval $0 \leq \Delta\phi < \pi/2$ needs to be considered. It can be seen that the proper τ_d increases with $\Delta\phi$. When $\Delta\phi$ is zero, τ_d achieves the minimum, which is 7.69 ns.

In the end, Fig. 9 presents how the code pseudorange bias varies over the satellite elevation under a few randomly selected parameter combinations on the curve of Fig. 8. It can be seen that under all these parameter combinations, the simulated results agree with the observations very well. Similar results can be obtained for the B1 and B2 signals of the BeiDou MEO satellite, as well as the B1, B2 and B3 signals of the BeiDou IGSO satellite, thus proving the effectiveness of the method proposed in this paper.

It is worth emphasizing that, due to the real parameters of the BeiDou satellite transmitting antennas are not available in open literature, the parameters of GPS Block IIR satellite are used as demonstration in this paper's simulations. Although with this compromise, the results are still persuasive enough to illustrate the

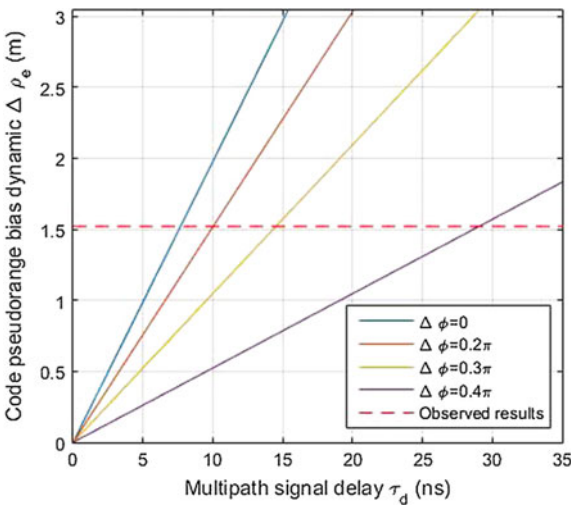


Fig. 7 Code pseudorange bias dynamic variations over the multipath signal delay

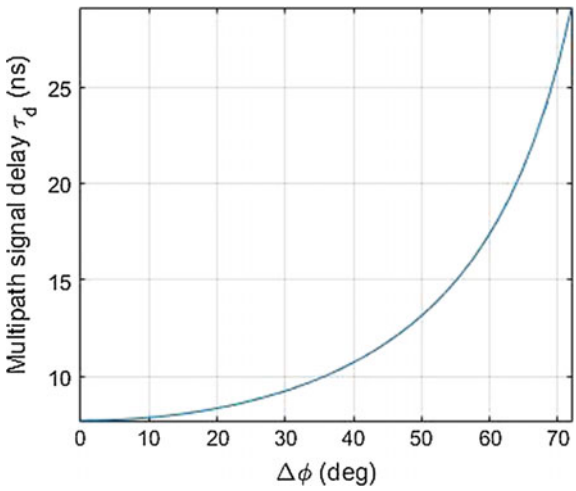


Fig. 8 Multipath signal delay variations over the phase difference between the direct and multipath signals

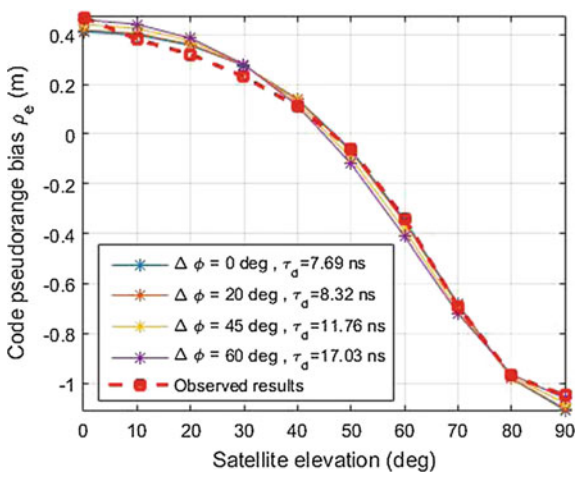


Fig. 9 Simulated code pseudorange bias variations over the satellite elevation

principle. Once the real parameters are given, more accurate simulation results can be obtained using the same method given above.

5 Conclusion

In order to analyse the cause of the code pseudorange bias anomaly of the BeiDou system, this paper models the satellite-induced multipath effect and proposes a method of calculating the bias value. Simulation results prove that transfer delays between different antenna elements can indeed cause elevation-dependent code

pseudorange bias, and under certain parameter combinations of the multipath model, simulated results agree with the observations very well. Therefore, it concludes that the anomaly may be caused by the satellite-induced multipath effect. Accordingly, we suggest that in the investigation of the BeiDou bias anomaly, special attention should be paid to see if the transfer delays motioned above exist. In addition, in order to avoid the satellite-induced multipath effect in the manufacture of following satellites, the transfer delay inconsistency of different antenna elements should be strictly restricted in the design, development and test processes.

References

1. Wanninger L, Beer S (2015) BeiDou satellite-induced code pseudorange variations: diagnosis and therapy. GPS Solutions 19:639–648
2. Springer T, Dilssner F (2009) SVN49 and other GPS anomalies. Inside GNSS 2009:32–36
3. Parkinson B, Spilker J (1996) Global position system: theory and applications, vol 1. AIAA, New York, pp 234–239
4. Marquis A, Reigh L (2015) The GPS block IIR and IIR-M broadcast L-band antenna panel: its pattern and performance. Navigation 62(4):329–347
5. Trees V (2002) Optimum array processing: Part IV of detection, estimation and modulation theory. Wiley, New York
6. Enge P (2001) The global positioning system: signals, measurements, and performance. Ganga-Jamuna Press, Lincoln, Massachusetts
7. Braasch S, Dibenedetto F (2001) Spread-spectrum ranging multipath model validation. IEEE Trans Aerosp Electron Syst 37(1):298–304

Single Receiver Against Repeater Deception Jamming Technology Research

Ke Liu, Wenqi Wu, Kanghua Tang, Shihao Zhang and Zhijia Wu

Abstract Intentional deception jamming brings great threat to satellite navigation receiver and seriously influence GNSS application security. For repeater deception jamming, an anti-spoofing algorithm based on a single receiver is proposed. Using the adjacent two times measurement pseudo-range does double differences, and Taylor expansion is applied to the position relationship in equations. Under the condition of the previous time's position information is known, the authenticity of the satellite signal of the next time can be verified according to the position information when it is combined with the least square iterative solution for the double differences equations. The algorithm is simple, and only a single receiver and pseudo-range measurement data is needed. NovAtel receiver is adopted for the actual experiments, and the result validates the feasibility and effectiveness of the algorithm.

Keywords Single receiver · Repeater deception jamming · Anti-spoofing · Pseudo-range differences · Taylor expansion

1 Introduction

As the global satellite navigation system plays more and more important roles in people's life and production and in-depth application, its security is becoming more and more seriously. Intentional or unintentional deception and interference affect the normal use of the navigation terminals and timing terminals and results in wrong positioning and timing information. Deceptive interference to the satellite

K. Liu · W. Wu (✉) · K. Tang · S. Zhang · Z. Wu
College of Mechatronics Engineering and Automation,
National University of Defense Technology, Deya Rd,
Changsha 410073, Hunan, China
e-mail: wenqiwu_lit@hotmail.com

K. Liu
e-mail: kevin880205@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_3

navigation receiver is mainly to make the receiver to capture the jamming signal, and then produce the wrong location. In general, it has two kinds of types: generated and repeated. C/A code of the satellite signal is open and transparent, and its structure is well known to the public, therefore the spoofing signals which are so similar to the real satellite navigation signals can be simulated and mislead receiver from the correct position. P code is encrypted, so it cannot be simulated easily. Compared to generated, repeated deception jamming has lower cost, and realized more easily. Normally, we are faced with repeated deception jamming.

For deception jamming, the jammer first need to destroy the connection between the receiver and the real signal, and then make the receiver capture the jamming signal preferentially by a higher signal power than the normal. In [1], the author pointed out that the receiver's tracking for the real signal can be destroyed as long as the deception jamming signal's power higher 4 dB than the real signal. In order to realize deception jamming suppression, the first thing is to detect the existence of the deception jamming signal accurately. The original detection methods are mainly concentrated on the signal distortion detection, including the signal power [2, 3], the spatial distribution properties [4, 5], etc. The current research focuses on using antenna array [6], the receiver pseudo-range or carrier phase differences [7, 8], correlation [9, 10], inertial aided [11, 12], etc. Different methods can realize deception jamming signal detection in a certain extent, but it also has the limitations and application scopes. For example, the signal arrival angle information can't be get when the ordinary antenna is used, then the method of direction of arrival monitoring will invalid; when there is only one single receiver or single antenna, dual receiver or dual antenna difference calculation method cannot be used, and the performance of the pseudo-range single difference is affected by the baseline length and the number of satellites. If single receiver can be used to implement detection, the disadvantages of the dual receivers can be overcome.

The purpose of this paper is to use a single receiver for repeated jamming signal detection. Through the pseudo-range differences calculation at the two time points and the least square iterative method is used to get the position, the authenticity of the signal can be confirmed. The paper is organized as follows: Sect. 2 gives the theoretical and method of single receiver against repeated deception jamming signal, in Sect. 3 signal transponder is used to build simulation environment for the actual test, and the conclusion is summarized in Sect. 4.

2 Single Receiver Pseudo-range Difference

2.1 Real Signal Difference

Suppose the measurement pseudo-range of the ith satellite at t_k is ρ_k^i, and t_{k+1} is ρ_{k+1}^i, then we have

$$\rho_k^i = r_k^i + \left[\delta t_r(k) - \delta t^i(k)\right] \cdot c + \left[\delta t_{ion}(k) + \delta t_{trop}(k)\right] \cdot c + \varepsilon_k^i \tag{1}$$

$$\rho_{k+1}^i = r_{k+1}^i + \left[\delta t_r(k+1) - \delta t^i(k+1)\right] \cdot c + \left[\delta t_{ion}(k+1) + \delta t_{trop}(k+1)\right] \cdot c + \varepsilon_{k+1}^i \tag{2}$$

where, r is the real distance, δt_r is the receiver's clock bias, δt^i is the satellite clock bias, δt_{ion} is the ionosphere delay, δt_{trop} is the troposphere delay, and ε^i is measurement noise and other non-model error.

We assume the ionosphere delay and troposphere delay at two times are consistent, and the non-model error such as measurement noise are not taken into account, then the pseudo-range single difference of Eqs. (1) and (2) is

$$\Delta\rho_{k+1,k}^i = \left[r_{k+1}^i - r_k^i\right] + \left[\delta t_r(k+1) - \delta t_r(k)\right] \cdot c - \left[\delta t^i(k+1) - \delta t^i(k)\right] \cdot c \tag{3}$$

For the satellite clock bias, we have

$$\begin{aligned} \delta t^i(k) &= a_0 + a_1(t_s(k) - t_{oc}) + a_2(t_s(k) - t_{oc})^2 \\ &= a_2 t_s^2(k) + (a_1 - 2a_2 t_{oc})t_s(k) + a_0 - a_1 t_{oc} + a_2 t_{oc}^2 \end{aligned} \tag{4}$$

In the general case, $a_2 = 0$, so Eq. (4) can be rewritten as

$$\delta t^i(k) = a_1 t_s(k) + a_0 - a_1 t_{oc} \tag{5}$$

further more, Eq. (5) can be written as

$$\delta t^i(k) = a_1\left[t_k - \rho_k^i/c\right] + a_0 - a_1 t_{oc} \tag{6}$$

substituting Eq. (6) into Eq. (3), and known that $\rho_{k+1}^i - \rho_k^i = \Delta\rho_{k+1,k}^i$, $t_{k+1} - t_k = 1$, then

$$\left(1 - a_1^i\right)\Delta\rho_{k+1,k}^i = \left[r_{k+1}^i - r_k^i\right] + \left[\delta t_r(k+1) - \delta t_r(k)\right] \cdot c - a_1 \cdot c \tag{7}$$

Similarly, for the jth satellite, we also have

$$\left(1 - a_1^j\right)\Delta\rho_{k+1,k}^j = \left[r_{k+1}^j - r_k^j\right] + \left[\delta t_r(k+1) - \delta t_r(k)\right] \cdot c - a_1 \cdot c \tag{8}$$

Pseudo-range double differences is calculated between Eqs. (7) and (8)

$$\left(1 - a_1^i\right)\Delta\rho_{k+1,k}^i - \left(1 - a_1^j\right)\Delta\rho_{k+1,k}^j = \left[r_{k+1}^i - r_k^i\right] - \left[r_{k+1}^j - r_k^j\right] \tag{9}$$

2.2 Spoofing Signal Difference

Suppose the measurement pseudo-range of the ith satellite at t_k is ρ_k^i, and t_{k+1} is ρ_{k+1}^i, then we have

$$\rho_k^i = d_{is,k} + d_{sr,k} + \left[\delta t_r(k) - \delta t^i(k)\right] \cdot c + \left[\delta t_{ion}(k) + \delta t_{trop}(k)\right] \cdot c + \varepsilon_k^i \tag{10}$$

$$\begin{aligned}\rho_{k+1}^i = {} & d_{is,k+1} + d_{sr,k+1} + \left[\delta t_r(k+1) - \delta t^i(k+1)\right] \cdot c \\ & + \left[\delta t_{ion}(k+1) + \delta t_{trop}(k+1)\right] \cdot c + \varepsilon_{k+1}^i\end{aligned} \tag{11}$$

where, d_{is} is the distance between the satellite and the transponder, d_{sr} is the distance between the transponder and the receiver, other symbols are defined as the former.

Pseudo-range single difference is made between Eqs. (10) and (11), and Eq. (6) is substituted in, then we have

$$\left(1 - a_1^i\right)\Delta\rho_{k+1,k}^i = \left[d_{is,k+1} - d_{is,k}\right] + \left[d_{sr,k+1} - d_{sr,k}\right] + \left[\delta t_r(k+1) - \delta t_r(k)\right] \cdot c - a_1 \cdot c \tag{12}$$

Similarly, for the jth satellite, we also have

$$\left(1 - a_1^j\right)\Delta\rho_{k+1,k}^j = \left[d_{js,k+1} - d_{is,k}\right] + \left[d_{sr,k+1} - d_{sr,k}\right] + \left[\delta t_r(k+1) - \delta t_r(k)\right] \cdot c - a_1 \cdot c \tag{13}$$

Pseudo-range double differences is calculated between Eqs. (12) and (13)

$$\left(1 - a_1^i\right)\Delta\rho_{k+1,k}^i - \left(1 - a_1^j\right)\Delta\rho_{k+1,k}^j = \left[d_{is,k+1} - d_{is,k}\right] - \left[d_{js,k+1} - d_{js,k}\right] \tag{14}$$

2.3 Taylor Expansion

For the real distance r_k^i in the real single, we have

$$r_k^i = \sqrt{\left(x_k^i - x_{r,k}\right)^2 + \left(y_k^i - y_{r,k}\right)^2 + \left(z_k^i - z_{r,k}\right)^2} \tag{15}$$

where, $\left(x_k^i, y_k^i, z_k^i\right)$ is the ith satellite position in ECEF coordinates at t_k which can be obtained by navigation messages, $\left(x_{r,k}, y_{r,k}, z_{r,k}\right)$ is the receiver's position in ECEF coordinates at t_k.

Taylor expansion of Eq. (15) is done at (x_0, y_0, z_0), so

$$r_k^i \approx r_{0,k}^i + \frac{\partial F}{\partial x}\Delta x_{r,k} + \frac{\partial F}{\partial y}\Delta y_{r,k} + \frac{\partial F}{\partial z}\Delta z_{r,k} \tag{16}$$

where

$$r_{0,k}^i = \sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2} \tag{17}$$

$$\frac{\partial F}{\partial x} = \frac{-x_k^i + x_0}{\sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2}} = e_{x,k}^i \tag{18}$$

$$\frac{\partial F}{\partial y} = \frac{-y_k^i + y_0}{\sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2}} = e_{y,k}^i \tag{19}$$

$$\frac{\partial F}{\partial z} = \frac{-z_k^i + z_0}{\sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2}} = e_{z,k}^i \tag{20}$$

$$\begin{cases} x_{r,k} = x_0 + \Delta x_{r,k} \\ y_{r,k} = y_0 + \Delta y_{r,k} \\ z_{r,k} = z_0 + \Delta z_{r,k} \end{cases} \tag{21}$$

In Eq. (9), the position at t_k is known when we detect the signal at t_{k+1}, so Taylor expansion is only needed to make for r_{k+1}^i and r_{k+1}^j.

$$\begin{aligned} \left[r_{k+1}^i - r_k^i\right] - \left[r_{k+1}^j - r_k^j\right] = & \sqrt{\left(x_{k+1}^i - x_0\right)^2 + \left(y_{k+1}^i - y_0\right)^2 + \left(z_{k+1}^i - z_0\right)^2} \\ & - \sqrt{\left(x_k^i - x_{r,k}\right)^2 + \left(y_k^i - y_{r,k}\right)^2 + \left(z_k^i - z_{r,k}\right)^2} \\ & - \sqrt{\left(x_{k+1}^j - x_0\right)^2 + \left(y_{k+1}^j - y_0\right)^2 + \left(z_{k+1}^j - z_0\right)^2} \\ & + \sqrt{\left(x_k^j - x_{r,k}\right)^2 + \left(y_k^j - y_{r,k}\right)^2 + \left(z_k^j - z_{r,k}\right)^2} \\ & + \left(e_{x,k+1}^i - e_{x,k+1}^j\right)\Delta x_{r,k} + \left(e_{y,k+1}^i - e_{y,k+1}^j\right)\Delta y_{r,k} \\ & + \left(e_{z,k+1}^i - e_{z,k+1}^j\right)\Delta z_{r,k} \end{aligned} \tag{22}$$

Similarly, for $d_{is,k}$ in the spoofing signal, we have

$$d_{is,k} = \sqrt{\left(x_k^i - x_{s,k}\right)^2 + \left(y_k^i - y_{s,k}\right)^2 + \left(z_k^i - z_{s,k}\right)^2} \tag{23}$$

where, $\left(x_k^i, y_k^i, z_k^i\right)$ is the ith satellite position in ECEF coordinates at t_k which can be obtained by navigation messages, $\left(x_{s,k}, y_{s,k}, z_{s,k}\right)$ is the transponder's position in ECEF coordinates at t_k.

Taylor expansion of Eq. (23) is made at (x_0, y_0, z_0), so

$$d_{is,k} \approx d_{is0,k} + \frac{\partial F}{\partial x}\Delta x_s + \frac{\partial F}{\partial y}\Delta y_s + \frac{\partial F}{\partial z}\Delta z_s \tag{24}$$

where

$$d_{is0,k} = \sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2} \tag{25}$$

$$\frac{\partial F}{\partial x} = \frac{-x_k^i + x_0}{\sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2}} = e_{x,k}^i \tag{26}$$

$$\frac{\partial F}{\partial y} = \frac{-y_k^i + y_0}{\sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2}} = e_{y,k}^i \tag{27}$$

$$\frac{\partial F}{\partial z} = \frac{-z_k^i + z_0}{\sqrt{\left(x_k^i - x_0\right)^2 + \left(y_k^i - y_0\right)^2 + \left(z_k^i - z_0\right)^2}} = e_{z,k}^i \tag{28}$$

$$\begin{cases} x_{s,k} = x_0 + \Delta x_{s,k} \\ y_{s,k} = y_0 + \Delta y_{s,k} \\ z_{s,k} = z_0 + \Delta z_{s,k} \end{cases} \tag{29}$$

In Eq. (14), the distance $d_{is,k}$ and $d_{js,k}$ at t_k is known when we detect the signal at t_{k+1}, so Taylor expansion is only needed to make for $d_{is,k+1}$ and $d_{js,k+1}$.

$$\begin{aligned} &\left[d_{is,k+1} - d_{is,k}\right] - \left[d_{js,k+1} - d_{js,k}\right] \\ &= +\sqrt{\left(x_{k+1}^i - x_0\right)^2 + \left(y_{k+1}^i - y_0\right)^2 + \left(z_{k+1}^i - z_0\right)^2} \\ &- \sqrt{\left(x_k^i - x_{s,k}\right)^2 + \left(y_k^i - y_{s,k}\right)^2 + \left(z_k^i - z_{s,k}\right)^2} \\ &- \sqrt{\left(x_{k+1}^j - x_0\right)^2 + \left(y_{k+1}^j - y_0\right)^2 + \left(z_{k+1}^j - z_0\right)^2} \\ &+ \sqrt{\left(x_k^j - x_{s,k}\right)^2 + \left(y_k^j - y_{s,k}\right)^2 + \left(z_k^j - z_{s,k}\right)^2} \\ &+ \left(e_{x,k+1}^i - e_{x,k+1}^j\right)\Delta x_{s,k+1} + \left(e_{y,k+1}^i - e_{y,k+1}^j\right)\Delta y_{s,k+1} \\ &+ \left(e_{z,k+1}^i - e_{z,k+1}^j\right)\Delta z_{s,k+1} \end{aligned} \tag{30}$$

Substituting Eq. (22) into Eq. (9), and using the least square method for iterative solution. When the termination condition $|x_{k+1} - x_k| < \varepsilon_x$ is satisfied, the real signal calculating result at t_{k+1} is output as $(x_{r,k+1}, y_{r,k+1}, z_{r,k+1})$ and this position is the receiver's real positioning. Therefore, when the receiver is at the stationary state, we have $(x_{r,k}, y_{r,k}, z_{r,k}) \approx (x_{r,k+1}, y_{r,k+1}, z_{r,k+1})$, and when the receiver is at the moving state, we have $(x_{r,k}, y_{r,k}, z_{r,k}) \neq (x_{r,k+1}, y_{r,k+1}, z_{r,k+1})$.

Substituting Eq. (30) into Eq. (14), and using the least square method for iterative solution. When the termination condition $|x_{k+1} - x_k| < \varepsilon_x$ is satisfied, the spoofing signal calculating result at t_{k+1} is output as $(x_{s,k+1}, y_{s,k+1}, z_{s,k+1})$ and this position is the transponder's positioning. So, no matter the state of the receiver is stationary or moving, we all have the approximate equality $(x_{s,k}, y_{s,k}, z_{s,k}) \approx (x_{s,k+1}, y_{s,k+1}, z_{s,k+1})$ as long as the transponder keeps still.

So we can draw a conclusion that the vehicle's real position information can be obtained from both of the least square method and the algorithm in this paper when the real satellite signals are acquired, and the results of them are approximate. While for the spoofing signals, the least square method get the wrong positioning result, and the algorithm in this paper get the transponder's position, and both of them are not equality. Furthermore, from Eqs. (22) and (9), Eqs. (30) and (14) we can see that equations solving accuracy depends on the position precision of the previous time and pseudo-range measurement precision.

In the course of the receiver for the real signal is lock-lose and capture the deception jamming signal, u-blox receiver can implement with no interrupt, but NovAtel receiver will have no satellite signal or less than four satellite signal in 2–3 s. Due to the position information $(x_{r,k}, y_{r,k}, z_{r,k})$ or $(x_{s,k}, y_{s,k}, z_{s,k})$ at t_k is unknown, we cannot continue to determine the authenticity of the satellite signal according to Eq. (22) or Eq. (30). Aimed at this situation, we can increase the time span between two signals when unlock occurs and the receiver capture signal again, and the form of the equations remain the same. For example, it is real signal at t_k and spoofing signal at t_{k+4}, and $t_{k+1} \sim t_{k+3}$ have no satellite signal or less than four satellite signal, then the pseudo-range double differences form like Eqs. (9) and (14) is

$$\left(1 - a_1^i\right)\Delta\rho_{k+4,k}^i - \left(1 - a_1^j\right)\Delta\rho_{k+4,k}^j = \left[d_{is,k+4} - r_k^i\right] - \left[d_{js,k+4} - r_k^j\right] \tag{31}$$

where the vehicle's position at t_k is known, that is to say, r_k^i and r_k^j can be determined. Then we can make Taylor expansion for $d_{is,k+4}$ and $d_{js,k+4}$, and Eq. (31) can be rewritten as

$$\begin{aligned}\left[d_{is,k+4}-r_k^i\right]-\left[d_{js,k+4}-r_k^j\right]=&+\sqrt{\left(x_{k+4}^i-x_0\right)^2+\left(y_{k+4}^i-y_0\right)^2+\left(z_{k+4}^i-z_0\right)^2}\\&-\sqrt{\left(x_k^i-x_{r,k}\right)^2+\left(y_k^i-y_{r,k}\right)^2+\left(z_k^i-z_{r,k}\right)^2}\\&-\sqrt{\left(x_{k+4}^j-x_0\right)^2+\left(y_{k+4}^j-y_0\right)^2+\left(z_{k+4}^j-z_0\right)^2}\\&+\sqrt{\left(x_k^j-x_{r,k}\right)^2+\left(y_k^j-y_{r,k}\right)^2+\left(z_k^j-z_{r,k}\right)^2}\\&+\left(e_{x,k+4}^i-e_{x,k+4}^j\right)\Delta x_{s,k+4}+\left(e_{y,k+4}^i-e_{y,k+4}^j\right)\Delta y_{s,k+4}\\&+\left(e_{z,k+4}^i-e_{z,k+4}^j\right)\Delta z_{s,k+4}\end{aligned}\tag{32}$$

Repeated deception jamming signal transponder's position at t_{k+4} can be obtained through the iterative solution of the above equations.

3 Experimental Analysis

To verify the feasibility and effectiveness of the algorithm, NovAtel receiver is adopted for the actual tests, and the process is divided into static actual tests and dynamic simulation tests. Receiver is placed at the top of the college building and receives the real signal. Don't change the receiver's position and open the signal transponder, and then we use its receiving antenna receives the real satellite signal and place transmitting antenna at the place where with the receiver two meters to simulate the satellite signal's relay. This kind of relay is by natural distance without humane latencies. The least square method and the above algorithm are used for data analysis and using $(-2.1969e6, 5.1775e6, 2.9981e6)$ as a reference point to the calculating result which is shown in Fig. 1 for coordinate transformation.

The sub-graph on the top of Fig. 1, the blue line represents the real position which from the real satellite signal calculated by the least square method, the green line represents the positioning result when the transponder is opened and the signal is transmitted by the signal transponder, that is to say, it is the spoofing positioning, the red line shows the result with the true signal using the algorithm in this paper, the black line shows the result with the repeated deception jamming signal using the algorithm in this paper. The sub-graph on the bottom of Fig. 1, the left is the position deviation when the real signal calculated by two kinds of algorithms and the right is the position deviation when the deception jamming signal calculated by two algorithms.

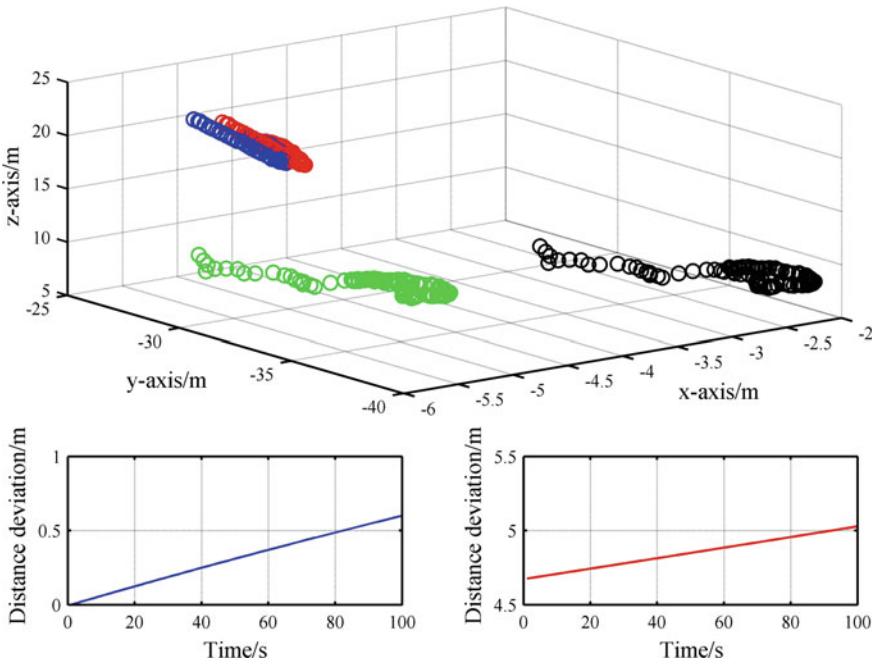


Fig. 1 Experimental result with no delay

From the figure we can see that, for the real satellite signal, the blue line and the red line keep consistent and the position deviation within 1 m. This means the position solved by the least square method and the algorithm in this paper are basically the same, and both of them are real position. While, for repeated deception jamming signal, the green line and the black line are not in the same and the distance deviation is greater than 4 m. The green line is the fault positioning comes from the measurement pseudo-range in spoofing signal, and the black line is the repeater deception jamming signal transponder's position.

In order to further verify the validity of the algorithm in the existing limited experimental conditions, we do the following things with measurement data:

(1) Added a fixed value to all the pseudo-range in the whole time of the deception jamming signals which the value is $c \times 1e-3$, where c is the speed of light. This process is used to simulate the same fixed time delay in the whole time, and the result is shown in Fig. 2;

(2) Added a fixed value to all the pseudo-range at the same epoch of the deception jamming signals which the value is $c \times (1+0.5k)e-4$. The value of k is different at different time and along with time increasing at the rate of unit 1. This process is used to simulate the same fixed time delay at the same epoch, different time with different delay, and the result is shown in Fig. 3;

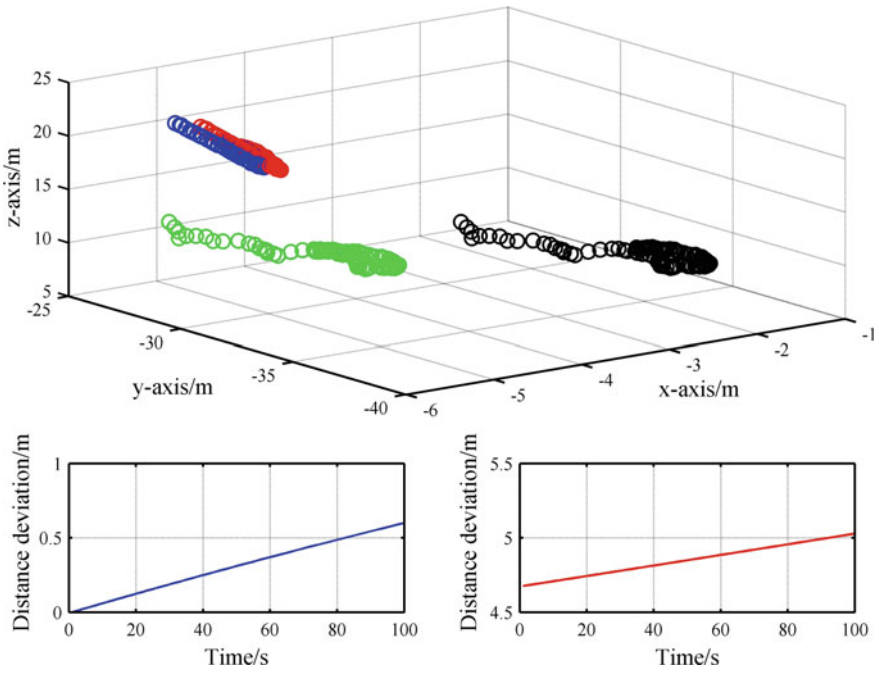


Fig. 2 Experimental result with the same fixed delay in the whole time

(3) Through data analysis, we found that the satellite which with carrier noise power density ratio greater than 45 dB have four. Then corresponding four satellite signals are selected and used for evaluating algorithm. For four satellite corresponding pseudo-range of the deception jamming signals at the same time added the value of $(1.38e-7, 0.78e-7, 1.1e-7, 1.53e-7)$, respectively. In addition, for the whole pseudo-range at the same time added a fixed value $c \times k \times 1e-8$. The value of k is different at different time and along with time increasing at the rate of unit 1. This process is used to simulate all pseudo-range delay are different in the whole time, and the result is shown in Fig. 4.

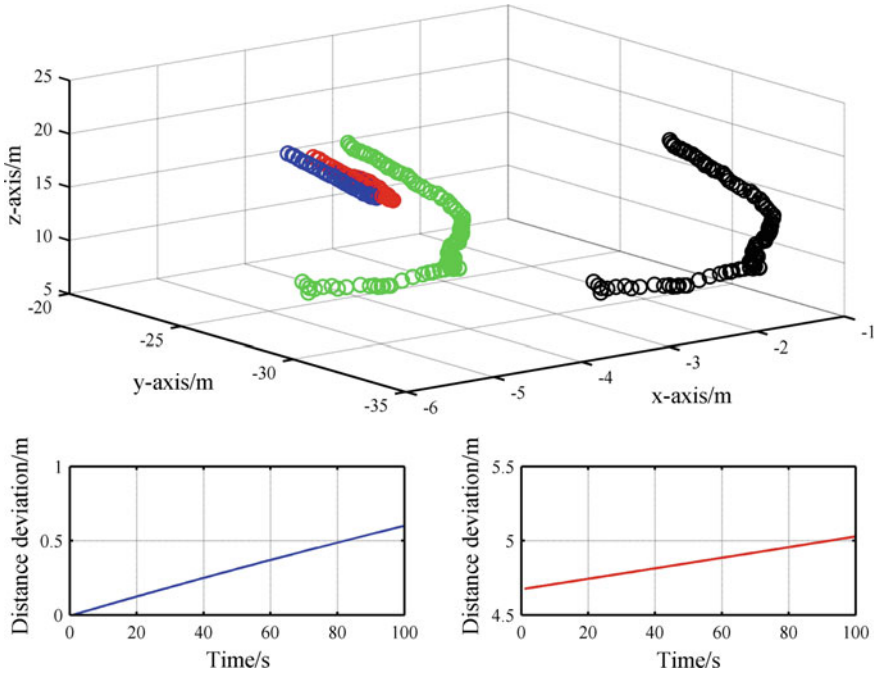


Fig. 3 Experimental result with the same delay at the same epoch

Lines color definition in Figs. 2, 3 and 4 are the same as the definition in Fig. 1. From the figures we can see that the algorithm positioning results are changed with the time delay. However, the distance deviation from two algorithms calculating result is basically remains unchanged. The calculating result in Figs. 1 and 2 are basically identical, and this is because the deception jamming signal added the same fixed delay in all the time, and it is offset each other in pseudo-range differences calculation according to the deception jamming signal model in Eqs. (10) and (11).

Experimental results agree with the theoretical analysis, therefore it can be used to identify whether the satellite signal true or not.

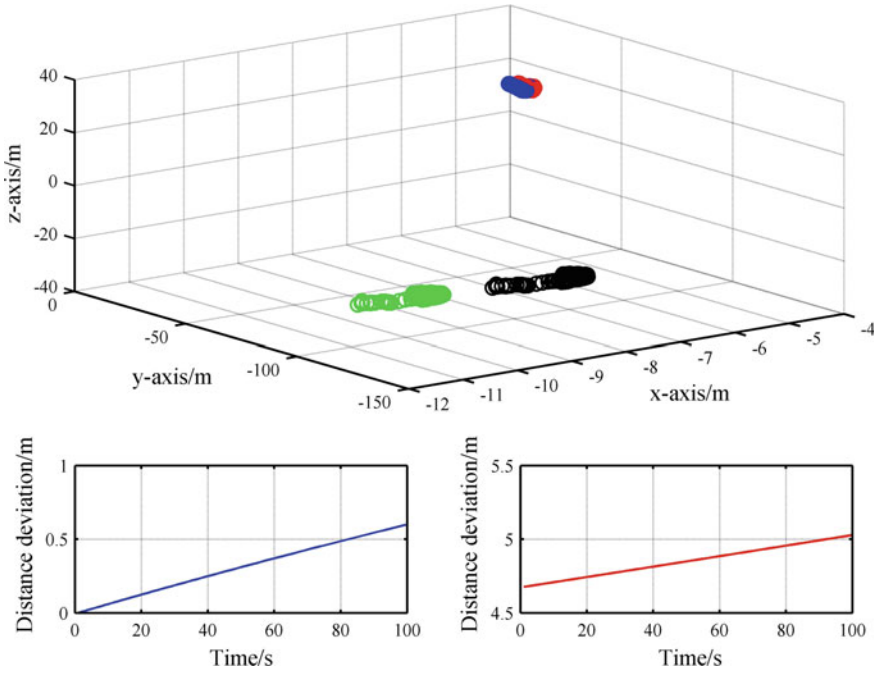


Fig. 4 Experimental result with different delay at different epoch and different pseudo-range

4 Conclusion

To the common problem of repeated deception jamming, this paper proposes a single receiver against repeated deception jamming algorithm. The realization of the algorithm requires only a single receiver or a single antenna, pseudo-range double differences is adopted between two adjacent times. The Taylor series approximation is carried out to solve the position relationships in equations, and then use the least square method to solve it. Finally the authenticity of the satellite signal is identified according to the calculating results.

The advantages of this algorithm are: (1) the requirement of the equipment and measurement information is low, and it only needs a single receiver and pseudor-ange information. Therefore, it can avoid the receiver detection performance affected by the baseline length; (2) it is not affected by satellite signal lock-lose. From unlock to recapture, it is only need to increase the time span between two signals and can realize the signal discrimination.

References

1. Long H, Zhicheng L, Feixue W (2012) Spoofing pattern research on GNSS receivers. J Astronautics 33(7):884–890
2. Jahromi AJ, Broumandan A, Nielsen J et al (2014) Pre-despreading authenticity verification for GPS L1 C/A signals. J Navig 61(1):1–11
3. Dehghanian V, Nielsen J, Lachapelle G (2012) GNSS spoofing detection based on signal power measurements: statistical analysis. Int J Navig Obs 7:1–8
4. Yaotian Z, Lu W, Wenyi W et al (2014) Spoofing jamming suppression techniques for GPS based on DOA estimating. In: The 6th China satellite navigation conference
5. Borio D (2013) PANOVA tests and their application to GNSS spoofing detection. IEEE Trans Aero Electron Syst 49(1):381–394
6. Wenyi W, Geng C, Renbiao W et al (2015) A low-complexity spoofing detection and suppression approach for ADS-B. In: Integrated communications navigation and surveillance (ICNS) conference
7. Borio D, Gioia C (2016) A sum-of-squares approach to GNSS spoofing detection. IEEE Trans Aero Electron Syst 52(4):1756–1768
8. Radin DS, Swaszek PF, Seals KC et al (2015) GNSS Spoof detection based upon pseudoranges from multiple receivers. In: International technical meeting of the Institute of Navigation
9. Psiaki ML, O'Hanlon BW, Bhatti JA et al (2013) GPS spoofing detection via dual-receiver correlation of military signals. IEEE Trans Aero Electron Syst 49(4):2250–2267
10. Broumandan A, Jahromi AJ, Lachapelle G (2015) Spoofing detection, classification and cancelation (SDCC) receiver architecture for a moving GNSS receiver. GPS Solut 19(3): 475–487
11. White NA, Maybeck PS (1998) Detection of interference/jamming and spoofing in a DGPS-aided inertial system. IEEE Trans Aero Electron Syst 34(4):1208–1217
12. Lee J-H, Kwon K-C, An D-S et al (2015) GPS spoofing detection using accelerometers and performance analysis with probability of detection. Int J Control Autom 13(4):951–959

The Effect of Cross-Correlation Items on the Intersystem Interference Between GPS, BDS and Galileo

Bingxue Chen, Xufang Huang, Xuyang Wang and Tangchao Li

Abstract The spectrum of four global satellite navigation systems allocated on the L1 band are severely overlapped, which makes signal interfere inevitably. Effective carrier power-to-noise density ratio ($(C/N_0)_{eff}$) and its degradation are important criteria to evaluate the degree of interference between Global Navigation Satellite Systems (GNSS). In order to simplify the analysis process and computational complexity, the cross-correlation items between desired signal and interference signals outputted from the receiver are ignored when calculating $(C/N_0)_{eff}$. In order to estimate how much deviation would be caused by ignoring the cross-correlation items, the power spectrum density (PSD) of cross-correlation items have been derived, and a new calculation formula of $(C/N_0)_{eff}$ considering the influence of cross-correlation items have been given, then the simulated results of intarsystem interference of GPS, BDS and Galileo were presented, and the results of intersystem interference between GPS, BDS and Galileo were also presented.

Keywords GNSS · Radio frequency compatibility · Cross-correlation function · Carrier power-to-noise density ratio

B. Chen (✉) · X. Huang · X. Wang · T. Li
Guangxi Colleges and Universities Key Laboratory of Satellite Navigation and Position Sensing, Guilin, China
e-mail: 804664595@qq.com

B. Chen · X. Huang · X. Wang · T. Li
Guangxi Key Laboratory of Multimedia Communications and Network Technology, Nanning, China

B. Chen · X. Huang · X. Wang · T. Li
Nanning College of Computer and Information Engineering, Guangxi University, Guilin, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_4

1 Introduction

Over the past a decade, there are some research about the interference between GPS, BDS and Galileo [1–10]. The effective carrier power-to-noise density ratio ($(C/N_0)_{eff}$) is an important criterion to assess the in-band interference from other systems. Some assessment of these interference effects has generally been overly conservative for the sake of simplicity in performing analyses [1–7]. Other some researchers noticed that simplifying may cause estimation errors. Dr. Dierendonck had considered some relatively short repeating codes, such as the GPS C/A code signals, and evaluated the effects of signal Doppler on short code spectral lines, data and symbol modulation effects [8]. Dr. Soualle analyzed the influence of the data rate and the spreading codes of the navigation signals onto intra-system interference for the GALILEO and GPS navigation systems [9]. Dr. Wei Liu presented a methodology combing the spectral separation coefficient (SSC) and code tracking spectral sensitivity coefficient (CT-SSC) for GNSS radio frequency compatibility assessment [10].

To date, these interference analysis methods have neglected the influence of the cross-correlation items onto intrasystem interference and intersystem interference for GNSS. We would focus on the effects of the cross-correlation items onto GNSS's mutual interference.

Firstly, based on the matched filtering principle, the power and power spectral density of cross-correlation items have been gained from the frequency domain, and the cross spectral separation coefficient (CSSC) is defined by copying the definition of spectral separation coefficient (SSC). The modified expressions of the $(C/N_0)_{eff}$ was presented in the end of this part, and the detailed calculating formula of equivalent power spectral density of intrasystem interference and intersystem interference were also presented. Then the numerical results of intrasystem interference of civil signals of GPS, BDS and Galileo were showed, and intersystem interference between GPS, BDS and Galileo were also displayed. During the simulation process, the real power spectral of C/A, L1C, B1 and E1OS are considered. Finally, the last section draws conclusions.

2 Signal Model of Cross-Correlation Items

2.1 Power of the Cross-Correlation Items

The expression of received signals from all visible satellites can be written as:

$$\begin{aligned} r(t) &= s(t) + u(t) + n(t) \\ s(t) &= \mathrm{Re}\{\alpha_o s(t-\tau_o)\exp(j2\pi(f_c+f_d)(t-\tau_o))\} \\ u(t) &= \mathrm{Re}\{\alpha_u s(t-\tau_u)\exp(j2\pi(f_{uc}+f_{ud})(t-\tau_u))\} \end{aligned} \tag{1}$$

where, $n(t)$ denotes thermal white noise with two-side spectral density of $N_0/2$. $s(t)$ denotes the desired signal, $s(t) = c(t)d(t)$, $d(t)$ denotes the data code, $c(t)$ denotes the pseudorandom noise (PRN) of the desired satellite signal. $u(t)$ denotes the interference signals come from the same satellite navigation system and come from the different navigation systems. f_c and f_{uc} denote the carrier frequency of the desired and the interference signals, respectively. α_0 and α_u denote amplitude of the desired and the interference signals, respectively. τ_0 and τ_u denote the propagation delay of the desired and the interference signals, respectively. f_d and f_{ud} denote the Doppler shift of the desired and the interference signals, respectively.

The spectral density function of the pseudo code $c(t)$ will be denoted as $S(f)$, according to the matched filtering theory, the transfer function of matched filter should be the complex conjugate of $S(f)$, which is denoted as $S^*(f)$. When the impact of noise and front-end filter is ignored, the spectral density function of output signals from the matched filter, which is denoted as $s_Y(t)$, would be expressed as:

$$Y(f) = (S_o(f) + S_u(f))S^*(f) \tag{2}$$

where, $S_o(f)$ denotes the Fourier transform of the desired signal, $S_u(f)$ denotes the Fourier transform of the interference signal.

Then the power spectral density of $s_Y(t)$ can be expressed as:

$$\begin{aligned} G_Y(f) &= |Y(f)|^2 = ((S_o(f) + S_u(f))S^*(f))((S_o(f) + S_u(f))S^*(f))^* \\ &= \left[G_o(f) + G_u(f) + \left(S_o(f)S_u^*(f) + S_o^*(f)S_u(f)\right)\right]G(f) \end{aligned} \tag{3}$$

Consequently, the power output from the matched filter can be obtained by calculating the integral of its PSD within the bandwidth:

$$\begin{aligned} P_Y = \int_{-\infty}^{\infty} G_Y(f)df = \int_{-\infty}^{\infty} G_o(f)G(f)df + \int_{-\infty}^{\infty} G_u(f)G(f)df \\ + \int_{-\infty}^{\infty} \left(S_o(f)S_u^*(f) + S_o^*(f)S_u(f)\right)G(f)df \end{aligned} \tag{4}$$

where, $G_o(f)$ is the power spectral density (PSD) of $s(t)$, $G(f)$ is the PSD of $c(t)$, $G_u(f)$ is the PSD of $u(t)$. $S_o^*(f)$ is the complex conjugate of $S_o(f)$, $S_u^*(f)$ is the complex conjugate of $S_u(f)$.

Now, considering the effects of the front-end filtering and noise, the power of $s_Y(t)$ should be expressed as:

$$
\begin{aligned}
P_Y &= \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 G_Y(f) df + N_0 \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 G(f) df \\
&= \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 G_o(f) G(f) df + N_0 \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 G(f) df \\
&\quad + \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 \left(G_u(f) + \left(G_{u,o}(f) + G_{o,u}(f) \right) \right) G(f) df
\end{aligned} \tag{5}
$$

where, $G_{u,o}(f) = \sqrt{C_s C_u} S_o(f + \Delta f) S_u^*(f + \Delta f)$ and $G_{o,u}(f) = \sqrt{C_s C_u} S_o^*(f + \Delta f) S_u(f + \Delta f)$, which are the cross power spectrum density between desired signal and interference signals. $H(f)$ is the transfer function of the receiver filter with bandwidth β_r.

2.2 Cross Power Spectral Density

Assuming that the desired C/A signal emits from one of visible satellites, denoted as PRN_i, and the interference signals may come from other visible satellites which are the same system or the different systems. For example, the C/A PRN_i signal will be suffered from intrasystem interference, such as Code Division Multiple Access (CDMA) interference, denoted as $PRN_j (j \neq i)$ code, P code, M code and L1C code, these interferences come from GPS. At the same time, the C/A PRN_i signal will be suffered from intersystem interference, such as BDS's signals and Galileo's signals. In the case of intrasystem interference, $S_u(f)$ can be rewritten as:

$$
S_u(f) = S_{CA\_j}(f) + S_p(f) + S_M(f) + S_{L1C}(f) \tag{6}
$$

The cross power spectrum $G_{u,o}(f)$ and $G_{o,u}(f)$ in Eq. (5) can be written as:

$$
G_{u,o}(f) = \left(S_{CA\_j}(f) + S_P(f) + S_M(f) + S_{L1C}(f) \right)^* S_{CA\_i} \tag{7}
$$

$$
G_{o,u}(f) = S_{CA\_i}^*(f) S_u(f) \tag{8}
$$

In the case of intersystem interference, $S_u(f)$ can be rewritten as:

$$
S_u(f) = S_{E1A}(f) + S_{E1B}(f) + S_{E1C}(f) \tag{9}
$$

So the cross power spectrum $G_{u,o}(f)$ and $G_{o,u}(f)$ can be rewritten as:

$$G_{u,o}(f)=(S_{E1A}(f)+S_{E1B}(f)+S_{E1C}(f))^{*}S_{CA\_i} \tag{10}$$

$$G_{o,u}(f)=S^{*}_{CA\_i}(f)S_{u}(f) \tag{11}$$

2.3 Cross Spectral Separation Coefficient

Spectral Separation Coefficients (SSC) provides a measure of the interference power output from a receiver when certain signals with given spectra are incident at its input, which is defined as:

$$SSC(\Delta f)=\frac{\int_{-\beta_r/2}^{\beta_r/2}|H(f)|^2 G_{u\_N}(f+\Delta f)G_N(f)df}{\int_{-\beta_r/2}^{\beta_r/2}|H(f)|^2 G_N(f)df} \tag{12}$$

The terms, $G_{u\_N}(f)$ and $G_N(f)$ denote normalized PSD of the interference signal and the local reference signal in the infinite bandwidth, respectively. That is, $\int_{-\infty}^{\infty} G_{u\_N}(f)df=1$ and $\int_{-\infty}^{\infty} G_N(f)df=1$. Δf represents the frequency shift, which comprises of Doppler frequency shift and the difference of central frequency between the desired signal and the interference. The unit of SSC is 1/Hz.

The theory of SSC is extended to measure of the cross-correlation items power, defined as:

$$CSSC(\Delta f)=\frac{\int_{-\beta_r/2}^{\beta_r/2}|H(f)|^2\left(G_{u,o\_N}(f+\Delta f)+G_{o,u\_N}(f+\Delta f)\right)G_N(f)df}{\int_{-\beta_r/2}^{\beta_r/2}|H(f)|^2 G_N(f)df} \tag{13}$$

where, $G_{u,o\_N}(f)$ and $G_{o,u\_N}(f)$ are normalized cross power spectral density, that means, $\int_{-\infty}^{\infty} G_{u,o\_N}(f)df=1$ and $\int_{-\infty}^{\infty} G_{o,u\_N}(f)df=1$.

For example, $CSSC(\Delta f)$ between the desired PRN_i code and the interference $PRN_j(j\neq i)$ code, can be calculated by the following equation, PRN_i codes belong to C/A signal:

$$CSSC_{CA\_ij}(\Delta f)=\frac{\int_{-\beta_r/2}^{\beta_r/2}|H(f)|^2\left(S^{*}_{CA\_j}(f+\Delta f)S_{CA\_i}(f+\Delta f)+S_{CA\_j}(f+\Delta f)S^{*}_{CA\_i}(f+\Delta f)\right)G_N(f)df}{\int_{-\beta_r/2}^{\beta_r/2}|H(f)|^2 G_N(f)df} \tag{14}$$

The $CSSC(\Delta f)$ between the PRN_i code and P code, or M code, or L1C code can be written as:

$$CSSC_{CA\_P}(\Delta f) = \frac{\int_{-\beta_r/2}^{\beta_r/2} |H_{(f)}|^2 \left(S_P^*(f+\Delta f) S_{CA\_i}(f+\Delta f) + S_P(f+\Delta f) S_{CA\_i}^*(f+\Delta f) \right) G_N(f) df}{\int_{-\beta_r/2}^{\beta_r/2} |H_{(f)}|^2 G_N(f) df} \tag{15}$$

2.4 Modified Effective Carrier Power-to-Noise Density Ratio

The traditional effective carrier power-to-noise density ratio has been defined as:

$$\left(\frac{C}{N_0}\right)_{eff} = \frac{C}{N_0 + I_{tot}} \tag{16}$$

$$I_{tot} = C_u \int_{-\beta_r/2}^{\beta_r/2} |H_{(f)}|^2 G_{u\_N}(f+\Delta f) G_N(f) df \tag{17}$$

When considering the cross-correlation items, Eq. (16) should be rewritten as:

$$(C/N_0)_{eff\_m} = \frac{C}{N_o + I_{tot\_m}} \tag{18}$$

$$C = C_s \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 G_{o\_N}(f+\Delta f) G_N(f) df \tag{19}$$

$$I_{tot\_m} = I_{tot} + \sqrt{C_s C_u} \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 \left(S_o(f+\Delta f) S_u^*(f+\Delta f) + S_o^*(f+\Delta f) S_u(f+\Delta f) \right) G_N(f) df \tag{20}$$

where, C is the received power of the desired signal, N_0 is the PSD of the thermal noise, I_{tot} and I_{tot_m} are the equivalent power spectra of the traditional and modified total interference, respectively. C_s and C_u are the received power of the desired and the interference signal, respectively.

Substituting (19) and (20) into (18) yields the modified $(C/N_0)_{eff\_m}$

$$\left(\frac{C}{N_0}\right)_{eff\_m} = \frac{C_s \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 G_{o\_N}(f+\Delta f) G_N(f) df}{\left(\begin{array}{l} N_0 \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 G(f) df + \sqrt{C_s C_u} \int_{-\beta_r/2}^{\beta_r/2} |H(f)|^2 \left(S_o(f+\Delta f) S_u^*(f+\Delta f) + S_o^*(f+\Delta f) S_u(f+\Delta f)\right) G_N(f) df \\ + C_u \int_{-\beta r/2}^{\beta r/2} |H_{(f)}|^2 G_{u\_N}(f+\Delta f) G_N(f) df \end{array} \right)} \tag{21}$$

Comparing with the traditional $(C/N_0)_{eff}$ [1–5], the traditional calculation formula neglects the cross power spectrum and Doppler frequency shift.

The degradation of $(C/N_0)_{eff}$ due to the cross-correlation items (in dB) should be modified as:

$$\left(\frac{C}{N_0}\right)_{\deg\_tot} = \left(\frac{C}{N_0}\right)_{\deg\_\text{intra}} + \left(\frac{C}{N_0}\right)_{\deg\_\text{inter}} \tag{22}$$

where,

$$\begin{aligned} \left(\frac{C}{N_0}\right)_{\deg\_\text{intra}} &= 10\log_{10}\left(1 + \frac{I_{\text{intra}}}{N_0}\right) \\ \left(\frac{C}{N_0}\right)_{\deg\_\text{inter}} &= 10\log_{10}\left(1 + \frac{I_{\text{inter}}}{N_0 + I_{\text{intra}}}\right) \end{aligned} \tag{23}$$

$$\begin{aligned} I_{\text{intra}}(t) = &\sum_{i=1}^{I-1} C_M^{(i)} SSC_M^{(i)}(\Delta f) + \sum_{u=1}^{U-1}\sum_{i=1}^{I} C_u^{(i)} SSC_u^{(i)}(\Delta f) \\ &+ \sum_{u=1}^{U-1}\sum_{i=1}^{I} \sqrt{C_M^{(i)} C_u^{(i)}}\, CSSC_u^{(i)}(\Delta f) \end{aligned} \tag{24}$$

$$I_{\text{inter}}(t) = \sum_{q=1}^{Q-1}\sum_{k=1}^{K_q}\sum_{j=1}^{J_p} C_{q,k}^{(j)} SSC_{q,k}^{(j)}(\Delta f) + \sum_{q=1}^{Q-1}\sum_{k=1}^{K_q}\sum_{j=1}^{J_p} \sqrt{C_M^{(j)} C_{q,k}^{(j)}}\, CSSC_{q,k}^{(j)}(\Delta f) \tag{25}$$

where, $(C/N_0)_{\deg\_tot}$ is the total degradation, $(C/N_0)_{\deg\_\text{intra}}$ is the degradation due to intrasystem interference. $(C/N_0)_{\deg\_\text{inter}}$ is the degradation due to intersystem interference. $SSC_M^{(i)}$ is the SSC between ith CDMA interference and the desired signals. $SSC_u^{(i)}$ is the SSC between the uth service on the ith satellite interference and the desired signals, $CSSC_u^{(i)}$ as the cross spectrum between signals separation coefficient system. $SSC_{q,k}^{(j)}$ is the SSC of the qth satellite of the pth system to transmit the jth interference signal, $CSSC_{q,k}^{(j)}$ is the cross spectral separation coefficient between the desired signal and interference. $C_M^{(i)}$, $C_u^{(i)}$, $C_{q,k}^{(j)}$ are the received power of CDMA interference signal, intrasystem and intersystem interference signal, respectively.

3 Numerical Results

3.1 The Real PSD of Civil Signals

Table 1 summaries the technical characteristics of GPS, Galileo and BDS signals in L1 band, detailed parameters can refer to the literature [11–13]. Table 2 summaries the calculation parameters which are calculated for the $(C/N_0)_{eff}$ of all civil signals

The GPS C/A codes, as well as the codes on the future GPS, BDS and Galileo signals, such as L1C, E1OS and B1, actually produce line spectrums. These codes are modulated with navigation data that spread the lines over the data/symbol bandwidth. The following simulations will account for these facts. A plot of a typical C/A code spectrum, B1 spectrum and E1B spectrum (relative to 1 W) in Fig. 1, we assumed that they have the same central carrier frequency in order to observe the difference. A close-in view of the real spectral density is shown in Fig. 2.

Table 1 Signal parameters of GPS, Galileo and BDS in L1 band

System	Service type	Carrier frequency/MHz	Modulation type	Chip rate/Mcps
GPS	C/A	1575.42	BPSK(1)	1.023
	P(Y)	1575.42	BPSK(10)	10.23
	M	1575.42	BOCsin(10,5)	5.115
	L1C	1575.42	MBOC(6,1,1/11)	1.023
Galileo	E1B	1575.42	MBOC(6,1,1/11)	1.023
	E1C	1575.42	MBOC(6,1,1/11)	1.023
	E1A	1575.42	BOCcos(15,2.5)	2.5575
BDS	B1	1561.098	QPSK	2.046
	B1-2	1589.742	QPSK	2.046

Table 2 Calculation parameter

Parameter	Value
Time period	7 day
Time step	60 s
Grid resolution	$5^0 \times 50$
Elevation angle	5^0
Emission bandwidth	30.69 MHz
Front-end bandwidth	24.552 MHz
N0	−201.31 dBW/Hz

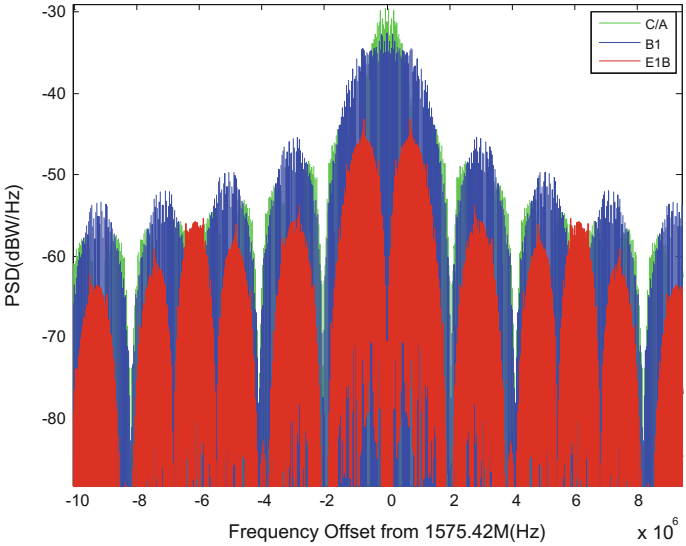


Fig. 1 Real PSD of C/A code, B1 code and E1B code

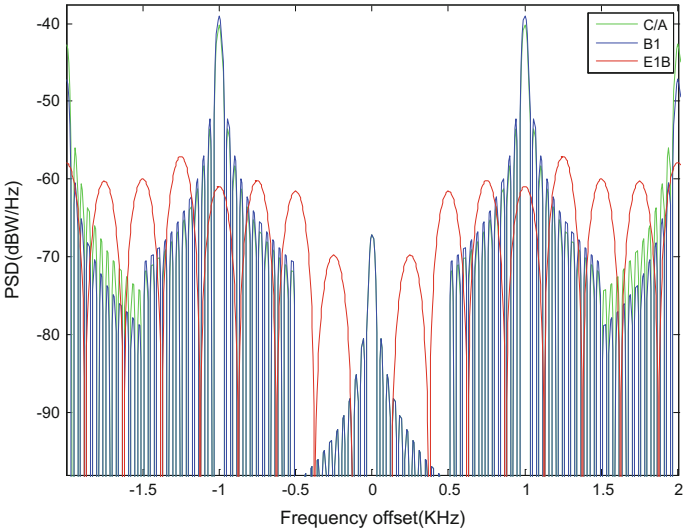


Fig. 2 Amplified real PSD of C/A code, B1 code and E1B code

3.2 *Spectral Separation Coefficient and Cross Spectral Separation Coefficient*

Figure 3 shows the single-sided convolution evaluation of Eq. (12) for two typical GPS C/A code spectral densities, as illustrated in Fig. 2, as a function of Doppler frequency offset, illustrating the attenuation of the spectral separation coefficient versus that offset, covering multiple 1-kHz lines. The minimum attenuation at zero-Doppler difference has an approximate value of −37.6 (dB/Hz). This attenuation is 20 (dB) less than the −57.6 (dB/Hz) average attenuation computed using a long-code (with the C/A code chipping rate) continuous spectral density. It is also more correct. Note that the attenuation at the odd-multiple of 500-Hz Doppler offsets is as much as 70 dB, much more than the long-code approximation would yield.

Figure 4 shows the CSSC versus the Doppler frequency offset. From the red curve, it can be seen that CSSC of desired C/A code and interference C/A code has the largest change, which is −40 to −80 dB. As the P code is also used BPSK modulation, the spectrum of the main lobe overlap is relatively serious, cross power spectrum interference followed by −70 to −110 dB. Moreover, the L1C code uses MBOC modulation, with the impact of the subcarrier modulation, so the cross power spectrum interference is slightly smaller. M code is military code, but the cross interference can not be ignored. While the cross spectral separation coefficients of P code, M code and L1C code, we can see that they are almost constant.

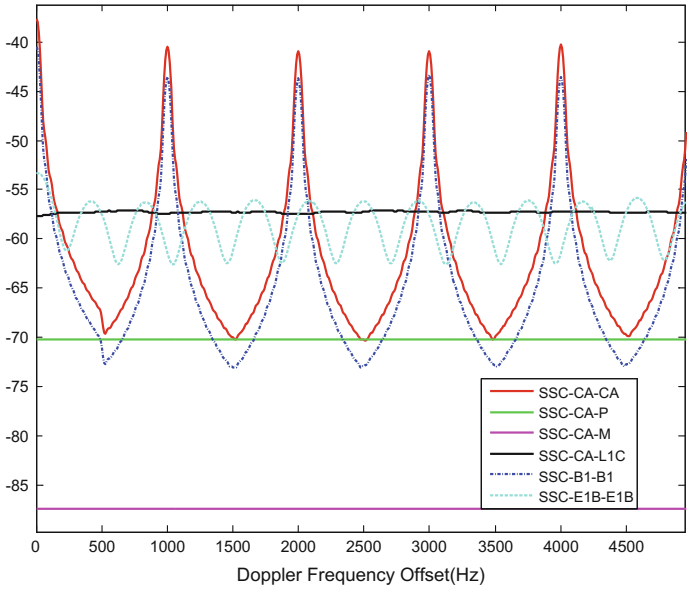


Fig. 3 Spectrum separation coefficients

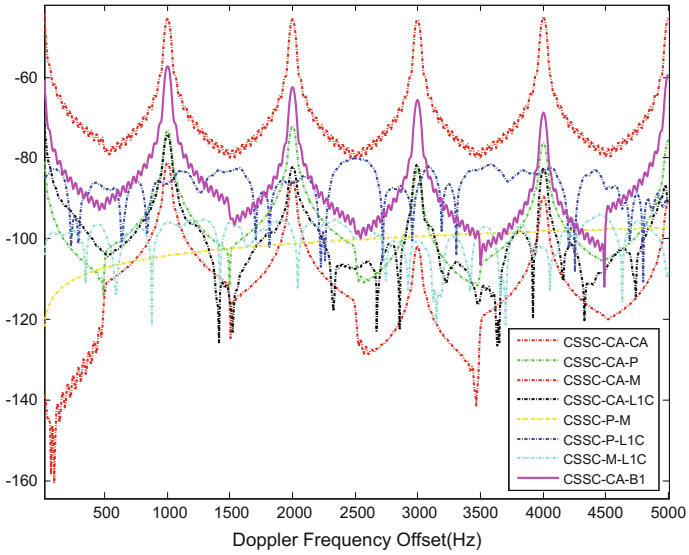


Fig. 4 Cross spectrum separation coefficients

3.3 Intrasystem Interference of GPS, BDS, Galileo

In this section, all the interference simulation results refer to the worst scenarios. The worst case assumes the minimum emission power of the desired signal, maximum emission power of the interference signals, maximum aggregate visible satellite over a grid of user locations on the earth's surface and over specified time intervals, and maximum $\Delta(C/N_0)$ over all the steps.

(1) Research on GPS C/A code in the worst scenarios, that is, the desired C/A code with the minimum power, other interference signals with maximum received power, and the maximum number of satellites received, the front-end bandwidth is 20.46 MHz. $\Delta(C/N_0)$ denotes a conventional value that does not take account of the cross-correlation items, $\Delta(C/N_0)_M$ denotes a modified value.

In Table 3, the real PSD of C/A code is used to obtain the results of the effective carrier power-to-noise density ratio's degradation with Doppler shift when the cross

Table 3 Maximum intrasystem interference of the real PSD of C/A code

Real PSD	$\Delta(C/N_0)$	$\Delta(C/N_0)_M$
Δf = 0 Hz	20–21.4	20.8–22.2
Δf = 500 Hz	0.95–1.25	1.05–1.35
Δf = 1000 Hz	17.2–18.6	18–19.4
Δf = 1500 Hz	0.85–1.1	0.95–1.2

power spectrum is considered. In lateral view, $\Delta(C/N_0)$ is larger than $\Delta(C/N_0)_M$ about 0.8–1.2 dB. In the longitudinal direction, the Doppler frequency shift has a great effect on the result, which has the smallest at 500 Hz, about 1.05–1.35 dB, and the maximum is 18–19.4 dB at 1000 Hz. This is also consistent with the results in Figs. 1 and 2.

The envelope PSD of C/A code is used in Table 4, and the data of $\Delta(C/N_0)$ is the result of the traditional value. It can be seen that $\Delta(C/N_0)_M$ which consider the effect of cross power spectrum increases by 2.1–2.4 dB.

(2) Research on BDS B1 in the worst case, that is, the desired B1 code with the minimum power, other interference signals with maximum received power, and the maximum number of satellites received, the front-end bandwidth is 20.46 MHz. It should be noted that the BDS system 1–5 satellite transmit a navigation message at 500 bps while 6–37 satellite navigates at 50 bps. It is now assumed that the desired signal B1 code is transmitted by one of the satellites 6–37.

In Table 5, $\Delta(C/N_0)$ is larger than $\Delta(C/N_0)_M$ about 0.5–1 dB, and the change in Table 6 is more obvious, which is about 1.25–1.7 dB. It can be seen from this, the effect of cross power spectrum is very large when the envelope PSD is calculated.

(3) Research on Galileo E1B code in the worst case, namely the desired E1B code with the minimum power, other interference signals with maximum received power, and the maximum number of satellites received, the front-end bandwidth is 24 MHz.

Table 4 Maximum intrasystem interference of the envelope of C/A code

Envelope PSD	$\Delta(C/N_0)$	$\Delta(C/N_0)_M$
$\Delta f = 0$ Hz	1.9–2.4	4.6–5.6

Table 5 Maximum intrasystem interference of the real PSD of B1 code

Real PSD	$\Delta(C/N_0)$	$\Delta(C/N_0)_M$
$\Delta f = 0$ Hz	13–15	13.5–16
$\Delta f = 500$ Hz	0.45–0.65	0.45–0.65
$\Delta f = 1000$ Hz	10–12	11–13
$\Delta f = 1500$ Hz	0.45–0.6	0.45–0.65

Table 6 Maximum intrasystem interference of the envelope of B1 code

Envelope PSD	$\Delta(C/N_0)$	$\Delta(C/N_0)_M$
$\Delta f = 0$ Hz	0.65–1	1.9–2.7

Table 7 Maximum intrasystem interference distribution of the envelope of E1B code

Real PSD	$\Delta(C/N_0)$	$\Delta(C/N_0)_M$
Δf = 0 Hz	6.2–6.8	6.3–6.8
Δf = 500 Hz	5.2–5.7	5.3–5.8
Δf = 1000 Hz	5.1–5.6	5.2–5.7

In Table 7, the effect of cross-correlation on Galileo E1B code is minimal compared with GPS and BDS, only about 0.1 dB. This has the certain relation with the anti-interference of the MBOC modulation of E1B code.

3.4 Intersystem Interference of GPS, BDS, Galileo

The data under the envelope PSD in Table 8 represents the traditional results without regard to cross-correlation effects. Analysis of the data in Table 8, we can see that: (1) Galileo cross power spectrum is always better than no cross situation, either on GPS or BDS; (2) the effect of GPS on BDS is greater than the effect of BDS on GPS, the interference of BDS B1 signal suffered from GPS signal is lager than that suffered from BDS. (3) The effect of C/A suffered from Galileo is greater than that from BDS. The center frequency of Galileo and GPS are the same, so the spectrum overlap is more serious, BDS has 14.322 MHz center frequency offset

Table 8 Maximum intersystem interference between GPS, BDS and Galileo

	Envelope PSD	Real PSD	Δf = 0 Hz	Δf = 500 Hz	Δf = 1000 Hz
C/A ← BDS	2–3.2 ($\times 10^{-3}$)	$\Delta(C/N_0)$	0.014–0.032	3.2–5 ($\times 10^{-3}$)	0.025–0.055
		$\Delta(C/N_0)_M$	0.012–0.026	0.022–0.036	0.025–0.055
C/A ← Galileo	0.22–0.27	$\Delta(C/N_0)$	0.2–0.35	3.55–4.05	0.4–0.65
		$\Delta(C/N_0)_M$	0.16–0.3	3.55–4.05	0.3–0.5
C/A ← BDS + Galileo	0.22–0.27	$\Delta(C/N_0)$	0.18–0.32	3.55–4.05	0.35–0.55
		$\Delta(C/N_0)_M$	0.18–0.32	3.55–4.05	0.35–0.55
B1 ← GPS	0.09–0.12	$\Delta(C/N_0)$	0.26–0.38	0.34–0.46	0.26–0.4
		$\Delta(C/N_0)_M$	0.32–0.48	0.48–0.62	0.32–0.5
B1 ← Galileo	0.03–0.036	$\Delta(C/N_0)$	0.085–0.13	0.18–0.215	0.115–0.16
		$\Delta(C/N_0)_M$	0.085–0.13	0.115–0.16	0.17–0.205
B1 ← GPS + Galileo	0.09–0.11	$\Delta(C/N_0)$	0.34–0.5	0.52–0.64	0.34–0.52
		$\Delta(C/N_0)_M$	0.4–0.6	0.64–0.8	0.4–0.6
E1B ← GPS	1.3–1.65	$\Delta(C/N_0)$	1.7–2.5	2.2–3.1	2.1–3
		$\Delta(C/N_0)_M$	1.7–2.6	2.3–3.2	2.2–3
E1B ← BDS	8.5–13.5 ($\times 10^{-3}$)	$\Delta(C/N_0)$	0.02–0.04	0.025–0.05	0.03–0.055
		$\Delta(C/N_0)_M$	0.02–0.04	0.025–0.05	0.03–0.055
E1B ← GPS + BDS	1.3–1.65	$\Delta(C/N_0)$	1.7–2.6	2.2–3.1	2.1–3
		$\Delta(C/N_0)_M$	1.8–2.6	2.3–3.2	2.2–3.1

away from them. (4) E1B interference is generally larger than other signals. Figure 2 can be more clearly seen in the actual power spectrum of C/A code is much larger than the E1B code, which is why the GPS signal will give E1B signal carrier to noise caused by more than decay value. (5) Compared with the traditional $\Delta(C/N_0)$, the values calculated by the cross power spectrum are almost slightly larger than the traditional values.

4 Conclusion

In this paper, the effect of the cross-correlation items to the $(C/N_0)_{eff}$ has been fully considered. We first modify the traditional $(C/N_0)_{eff}$ calculation model, and give a more accurate analytical formula considering the effect of cross-correlation items. Based on the modified effective carrier power-to-noise density ratio calculation model, the simulation results of intersystem interference and intrasystem interference between GPS, BDS and Galileo are discussed. It points out that when the envelope PSD is used to calculate the intrasystem interference, the interference value is about 1.5–3 dB larger than the unoptimized method; the intersystem interference is about 0.5–1 dB when using the cross power spectrum calculation, and the average value is 0.05–0.1 dB when calculating the intersystem interference; using the cross power spectrum calculation after correction, the intersystem interference is about 1–1.5 dB larger than the traditional value. It can be seen that the influence of cross-correlation items can not be neglected on the calculation of effective carrier power-to-noise density ratio.

Acknowledgements Funded by National Natural Science Foundation of China (Grant No. 61563004, 61004123)

Appendix (附录)

See Figs. 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, and 16

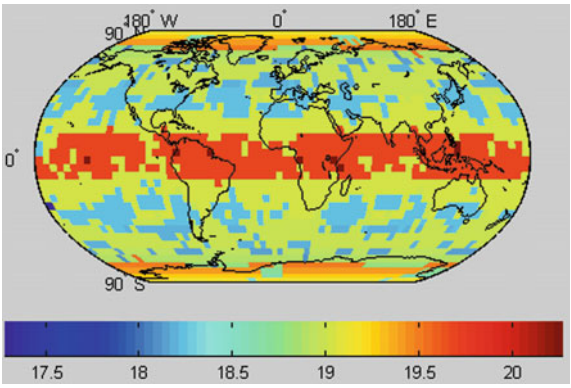


Fig. 5 Maximum intrasystem interference of GPS C/A

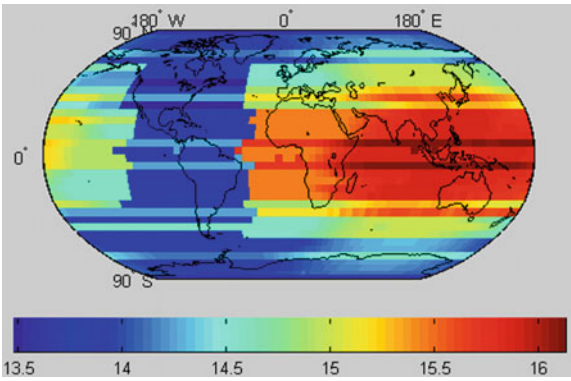


Fig. 6 Maximum intrasystem interference of BDS B1

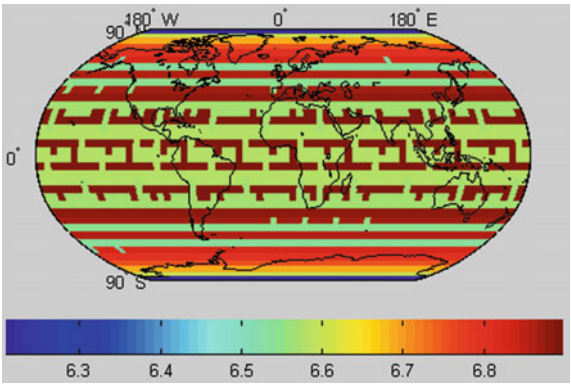


Fig. 7 Maximum intrasystem interference of Galileo E1B

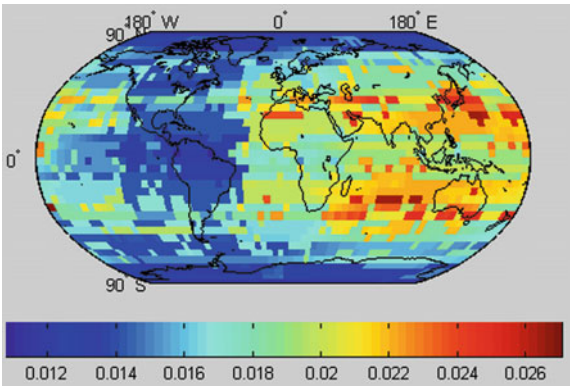


Fig. 8 Maximum intersystem interference between CA and BDS

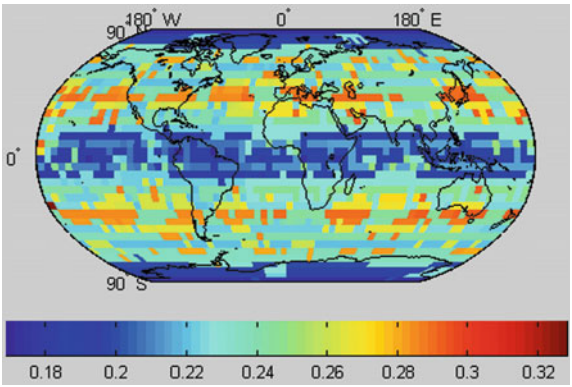


Fig. 9 Maximum intersystem interference between CA and BDS, Galileo

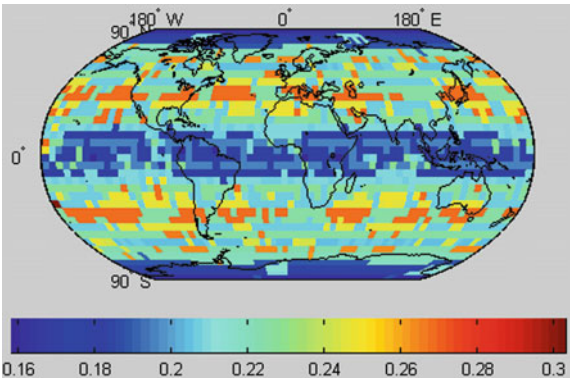


Fig. 10 Maximum intersystem interference between CA and Galileo

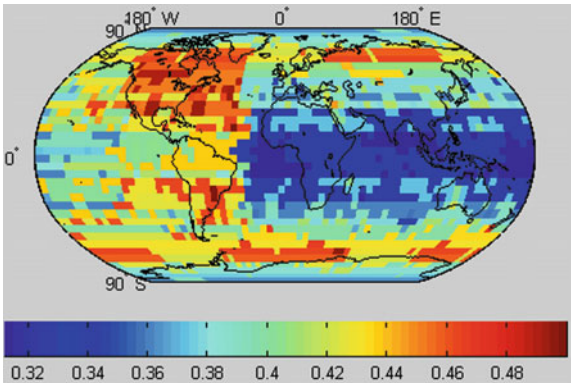


Fig. 11 Maximum intersystem interference between B1 and GPS

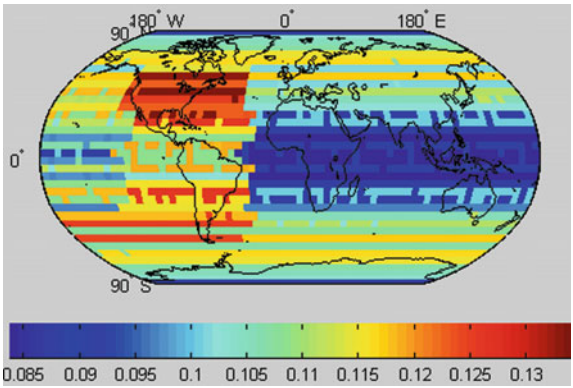


Fig. 12 Maximum intersystem interference of B1 and Galileo

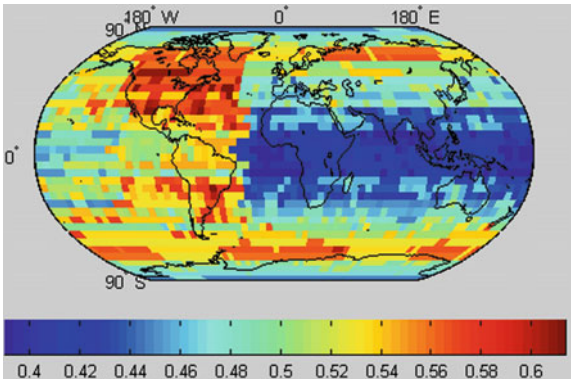


Fig. 13 Maximum intersystem interference between B1 and GPS, Galileo

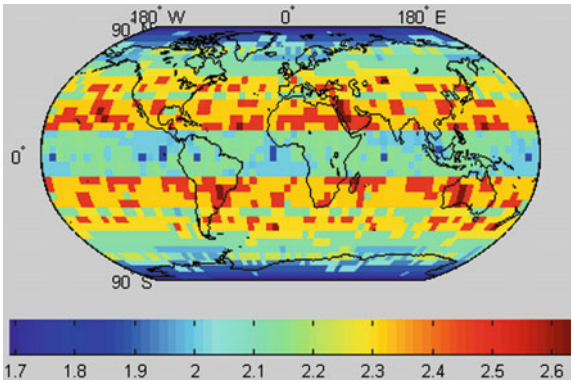


Fig. 14 Maximum intersystem interference between E1B and GPS

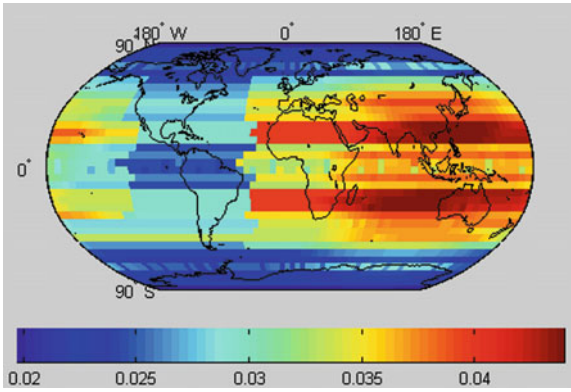


Fig. 15 Maximum intersystem interference between E1B and BDS

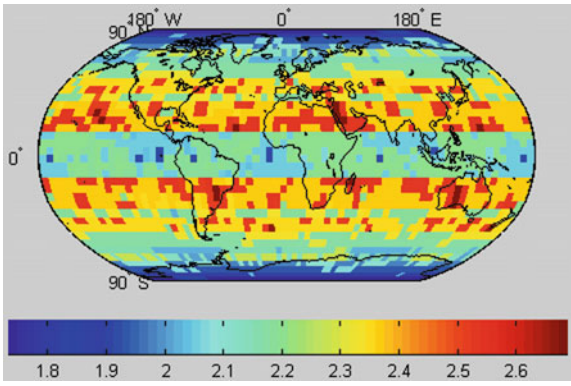


Fig. 16 Maximum intersystem interference between E1B and GPS, BDS

References

1. Betz JW (2000) Effect of narrowband interference on GPS code tracking accuracy. In: Proceedings of the National Technical Meeting of the Institute of Navigation, pp 16–27
2. Betz JW, Corporation TM (2001) Effect of partial-band interference on receiver estimation of C/N0: theory. In: Proceedings of the National Technical Meeting of the Institute of Navigation, pp 22–24
3. Godet J (2000) GPS/GALILEO radio frequency compatibility analysis. In: Proceedings of the ION GPS, Sept 2000, pp 1782–1790
4. Van Dierendonck AJ, Hegarty C (2002) Methodologies for assessing intrasystem and intersystem interference to satellite navigation systems. In: Proceedings of the ION GPS, Jan 2002, pp 1241–1250
5. Titus BM, Betz J, Hegarty C et al (2003) Intersystem and intrasystem interference analysis methodology. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, pp 2061–2069
6. Liu W, Du G, Zhan X et al (2010) Assessment of radio frequency compatibility relevant to the Galileo E1/E6 and compass B1/B3 bands. J Navig 63(3):419–434
7. Liu W, Zhai CR, Zhan XQ et al (2011) Assessment and analysis of radio frequency compatibility among several global navigation satellite systems. Iet Radar Sonar Navig 5 (2):128–136
8. Dierendonck AJV, Hegarty C, Pullen S (2003) A more complete and updated methodology for assessing intrasystem and intersystem interference for GPS and galileo, pp 1484–1493
9. Soualle F, Burger TB (2003) Impact of Galileo spreading code selection and data rate onto navigation signal interference. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, pp 1035–1043
10. Liu W, Li S, Liu L et al (2010) A comprehensive methodology for assessing radio frequency compatibility for GPS, Galileo and compass. ION GNSS Portland, OR, US 943–954
11. Lassiter EM (1993) NAVSTAR GPS space segment/navigation user interfaces, (Rev. C). ICD-GPS-200 C
12. Code T (2010) European GNSS (Galileo) open service: signal in space: interface control document. European Gnss
13. Zhao XD (2006) Submission of the updated information of compass system to the fourth resolution 6009 consultation meeting (WRC'03). People's Republic of China, Radio Regulatory Department

An Improved High-Sensitivity Acquisition Algorithm for BDS B2 Signal

Xue Wang, Jianchao Du and Weibin Li

Abstract The NH code modulation of Beidou B2 satellite navigation signal causes the problem of high bit reversal rate, which results in difficulty to directly apply the coherent integration over periods longer than 1 ms as capturing weak Beidou B2 satellite signal. Besides, it needs to search again to determine the acquisition of NH code chip position. Based on half-bit acquisition algorithm, an improved acquisition algorithm is proposed which combines the feature sequence and half-bit algorithm according to the characteristics of B2 signal. The proposed algorithm handles both of high bit reversal rate problem and NH code chip positioning problem. Experimental result shows that the performance of the proposed algorithm is not only improved about 2 dB compared to the non-coherent algorithm, but also does not need additional steps to determine the NH code phase.

Keywords Beidou B2 signal · Half-bit algorithm · Characteristic sequence · NH code

1 Introduction

The Beidou Navigation Satellite System (BDS) is a global navigation system compatibly with other navigation systems that has been developed and operated independently by China. It has been designed to provide positioning, navigating and timing services for Southeast Asia region in 2012 and extend to global users in 2020 [1]. By the end of 2013, the signal context of B2 frequency point is added to documents *Beidou navigation satellite signal in space interface control document open service signal Version 2.0* [2]. It means that the Beidou has become the world's first satellite system to use two frequency points to achieve civil positioning

X. Wang · J. Du (✉)
School of Telecommunications Engineering, Xidian University, Xi'an, China
e-mail: jcdu@xidian.edu.cn

W. Li
Shaanxi Jinkong Compass Information Service Co., LTD, Xi'an, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_5

service, which the positioning accuracy has improved to meter level even decimeter level comparing to the single frequency signal.

The Beidou B2 satellite navigation signal of NH code modulation improves the ability to resist narrow band interference and the performance of cross-correlation. However, the rate of NH code is 1 kbps, which causes the problem of high bit reversal rate. Besides, it is difficult to directly apply the coherent integration over periods longer than 1 ms and to determine the NH code position as capturing weak Beidou satellite signal. It is not sufficient to meet the requirements of the software receiver for real-time acquisition.

To overcome this problem, this paper studies a high-sensitive B2 signal capturing algorithm. First, it introduces the composition of the Beidou B2 signal and analyses the feature of navigation message. Secondly, it brings the concept of feature sequence and the half-bit acquisition. Then, an improved acquisition algorithm is proposed that combines the feature sequence and the improved half-bit algorithm according to the characteristics of B2 signal. Finally, the proposed algorithm is verified with simulated BDS B2 signal in Matlab. The performance of the proposed algorithm is compared with that of the non-coherent integration algorithm [3].

2 Analysis of BDS B2 Signal

2.1 The Composition of the BDS B2 Signal

The B2 frequency signal of Beidou can be written as

$$\begin{aligned} S_{B2}^{j}(t) = {} & A_{B2I} C_{B2I}^{j}(t) D_{B2I}^{j}(t) \cos(2\pi f_{B2} t + \varphi_{B2I}^{j}) \\ & + A_{B2Q} C_{B2Q}^{j}(t) D_{B2Q}^{j}(t) \sin(2\pi f_{B2} t + \varphi_{B2Q}^{j}) \end{aligned} \tag{1}$$

where j is satellite number. I and Q denote the in-phase channel and quadrature channel. A is the amplitude of signal and C is pseudo-random noise (PRN) code. D is the navigation data. f_B is the B2 frequency which is 1207.140 MHz. φ is the initial phase.

It can be found the B2 frequency signal of Beidou is similar to the L1 frequency signal of GPS. Both of them adopt the CDMA and right-hand circular polarization modulation. The difference is that the Beidou adopts the QPSK modulation and its ranging code rate and code length are twice than the ranging code rate and code length of GPS [4]. So the B2 frequency signal of Beidou has high spectrum utilization and good anti-interference ability.

In the paper, we only explore the in-phase channel because the quadrature channel is licensed and used for military rather than civil use. The received in-phase channel signal is defined as the following expression.

$$S_{B2}^{j}(t) = A_{B2I}C_{B2I}^{j}(t)D_{B2I}^{j}(t)\cos(2\pi f_{B2}t + \varphi_{B2I}^{j}) \tag{2}$$

The satellite radio frequency signal is transmitted to the ground by the satellite aerials. Then the signal passes through the atmosphere and is received by the receiver. The receiver carries out a serial of operation such as amplifying power, down-conversion, filtering and sampling-quantization to get IF signal. The navigation satellite captured and processed signal is the digital IF signal. The B2 digital IF signal of Beidou can be written as

$$S_k = A_I C_I(t_k - \tau)D_I(t_k - \tau)\cos(2\pi(f_{IF} + f_d)(t_k - \tau) + \varphi_0) + n_k \tag{3}$$

where t_k is the kth sample time, τ is the delay of signal. The f_d is the Doppler frequency shift of satellite signals. The n_k is the noise value at the sample time. The S_k is the B2 IF signal value at the sample time.

2.2 NH Code and Navigation Message

The B2 signal is modulated with NH code. The NH code contains 20 chips. Its length is presented by N_S. The NH code is the fixed value storage code. The values are that 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0. The modulated NH code can improve the ability to resist narrow-band interference, spread the spectrum. It benefits to bits and signal synchronization and improves mutual correlation. The B2 satellite signal of Beidou spreads the navigation message, which has a frequency of 50 Hz in D1 message and 500 Hz in D2 message. D1 message is modulated with NH code and the NH code rate is 1 kbit. Both of code carry out the operation of mod 2 and load to the B2 I channel signal of *MEO*/*IGSO* satellites. The D1 navigation message with modulated NH code has high bit reversal rate, so it is difficult to capture the weak B2 signal of Beidou. Therefore, this paper focuses on the D1 navigation message broadcast on high-sensitivity acquisition technology. The composition of D1 navigation is illustrated in Fig. 1.

According to Fig. 1, the NH code period is 20 ms and the D1 period is 20 ms. One period NH code matches one bit D1 navigation message and the starting position of both is strictly aligned. The length of NH code is 20, which means one period has 20 code chips and the chipping width is 1 ms. The period of bit reversal is 1 ms so it is difficult to directly apply the coherent integration over periods longer than 1 ms. It has no effect on no-coherent integration over periods longer than 1 ms. Therefore the no-coherent algorithm can be used to capture the weak B2 signal of Beidou. However, it needs to search again to determine the position of NH code chip using the no-coherent algorithm.

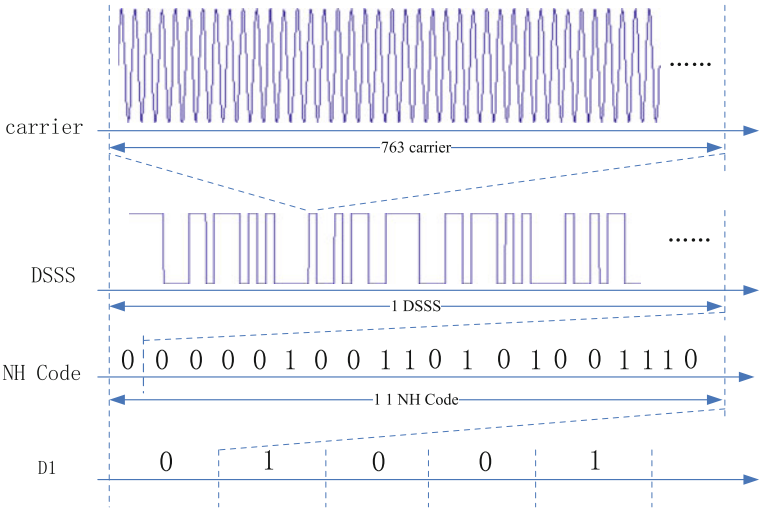


Fig. 1 The composition of D1 navigation

3 Acquisition Algorithm

3.1 FFT Acquisition Algorithm

Signal fast acquisition in the receiver is implemented by a FFT-based parallel code phase search acquisition method [5]. The algorithm is illustrated in Fig. 2.

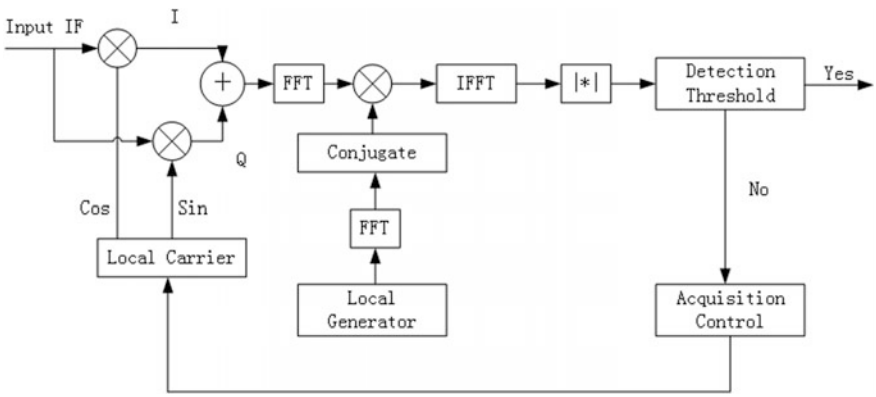


Fig. 2 Parallel code phase search acquisition algorithm

3.2 Half-Bit Algorithm

Previous work related to BDS weak signal acquisition considers only the I channel signal for coherent integration, and then the result is integrated non-coherently for longer integration periods to improve the sensitivity of acquisition. The longer the data for coherent integration, the more the extra gain of coherent integration, the higher the sensitivity of acquisition is [6]. The improved half-bit method is that each 1 ms IF signal is divides into two parts. The former 0.5 ms data is odd signal and the rest of 0.5 ms is zero-padding to combine 1 ms data. The latter 0.5 ms data is even signal and the rest of 0.5 is also zero-padding to combine 1 ms data. According to the principle of base 2 FFT, the padding data carries out radix-2 Fast Fourier Transform. In this way, one of the two groups of data doesn't contain the bit reversal. The improved half bit algorithm is illustrated in Fig. 3.

4 Improved Half Bit-Feature Sequence Acquisition Algorithm

4.1 Concept of Feature Sequence

The section introduces the concept feature sequence [7]. Firstly, we introduce the linear generated space of linear algebra. The longest linear feedback shift register sequence is called m sequence with periodic and copied. The length n of m sequence generate pseudo-random sequences, whose cycle is $2^n - 1$ chips [7].

$$c_k = f(c_{k-1}, c_{k-2}, \ldots, c_{k-n}) \tag{4}$$

The minimum length which is sufficient to determine the chip position is called the linear span space of the code sequence. The NH code which modulates the B2 I channel signal of Beidou are 20 storage code chips with fixed value not

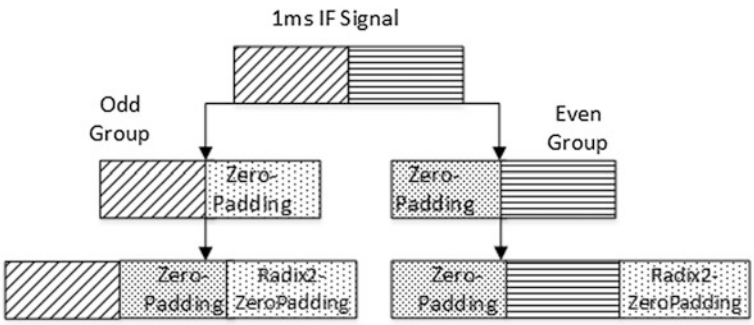


Fig. 3 The half bit algorithm

pseudo-random code generated by the shift register. We extend the above mentioned concept of the linear span space to the storage codes. We call the linear span space of NH code the Feature Length. The calculation steps of the feature length are as follows [7].

(1) Let k be the max number of contiguous 'zeros' or 'ones' in the sequence whose length.
(2) The size of M matrix is which the k is the length of sub sequence. Each row is shifted by one bit based on the previous row. Mathematically,

$$M = \begin{bmatrix} c_1 & c_2 & \cdots & c_k \\ c_2 & c_3 & \cdots & c_{k+1} \\ \vdots & \vdots & \cdots & \vdots \\ c_{N_s-k} & c_{N_S-K-1} & \cdots & c_{N_S} \\ c_{N_S} & c_1 & \cdots & c_{k-1} \end{bmatrix} \tag{5}$$

(3) Calculate the correlation coefficient of M matrix. Detect the matrix for identical rows. If there are any rows are identical then auto increment k and repeat the step 2 until no two rows of the M matrix are identical.
(4) The feature length L of sequence is the minimum values of k which satisfies the condition in step 3.

According to the Sect. 2.2, the NH code sequence which modulates the D1 navigation message is $(0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1, 0)$ and it can generate 20 sub sequences. The feature length of NH code is 7 using the aforementioned procedure calculation. Only the length of sub sequence is greater than or equal to 7, any two sub sequence is not identical. Hence the method can uniquely determine the NH code chip position. And call those sub sequences are feature sequence of NH code.

4.2 *Improved the Half-Bit and Feature-Sequence Acquisition Algorithm*

According to Sects. 3.2 and 4.1, the improved half-bit algorithm can avoid the reversal of length the correlation operation and the feature sequence can uniquely determine the NH code chip position. In this paper, we propose an improved acquisition algorithm which combines the feature sequence and half-bit algorithm according to the characteristics of B2 signal. The proposed algorithm principle is introduced as follows. Use the 80 ms length of B2 IF signal to carry out weak satellite acquisition. According to Sect. 3.2, the IF signal is divided into two parts using aforementioned procedure. The data of odd group and even group can be expressed as:

$$S_{odd} = \{D_1, D_2, D_3, \ldots, D_n\} \tag{6}$$

$$S_{even} = \{E_1, E_2, E_3, \ldots, E_n\} \tag{7}$$

where S_{odd} is the data of odd group and S_{even} is the data of even group. The D_n is the data of nth odd block and E_n is the data of nth even block.

If $k \geq 7$, the 20 sub sequences which is generated by NH code is not identical. Hence

$$C_k = \{C_{k1}, C_{k2}, \ldots, C_{ki}, \ldots, C_{kn}\} \tag{8}$$

where C_{ki} is the length k sub sequence of the ith NH code chip. Considering the signal amplitude and anti-interference ability, the value of k is set 10.

Since the D1 navigation message is modulated synchronously by NH code, the initial edge of first NH code chip is corresponding to the initial edge of the navigation message. Hence calculate sum of result which the first sequence is multiplied with the continuous k data blocks. Then first sequence is multiplied with the next 1 ms data and is summed. In this case we need to continue these operation such as shift, multiply and sum until 20 times because the length of NH code is 20 ms. When the phases between the NH code and the multiplied feature sequence is same, the NH code is right stripped. The data of stripped NH code carried out the sum of data block which equal to coherent integration. It means that do first accumulation and then carry out correlation calculation. Due to strip the NH code, the longest length of coherent integration reach to 20 data blocks. Hence it can not only improve the sensitivity of the acquisition but also avoid the problem of bit reversal in the presence of NH code.

$$S_{oi} = \sum_{n-1}^{k} D_{i+n-1} * C_{k1}(n) \tag{9}$$

$$S_{ei} = \sum_{n-1}^{k} E_{i+n-1} * C_{k1}(n) \tag{10}$$

where S_{oi} is the result that the feature sequence is multiplied and sum with the ith data of odd group. The S_{ei} is the result that the feature sequence is multiplied and sum with the ith data of even group.

The sum results carry out zero-padding and radix-2 operation. Then the output carries out FFT parallel correlation calculation with local copy PRN (S_1). Mathematically,

$$Z_{oi} = \mathit{ifft}[\mathit{fft}(S_{oi}). * \mathit{fft}(S_1)] \tag{11}$$

$$Z_{ei} = \mathit{ifft}[\mathit{fft}(S_{ei}). * \mathit{fft}(S_1)] \tag{12}$$

where S_1 is the complex conjugate of the local copy PRN based Fast Fourier Transform. The * is the point-multiplication.

Hence, the odd group and even group have 20 result matrix of coherent operation which correspond to 20 the positions of NH code chip. Find the peak value in the 20 result matrix. The NH code chip position depends on the maximum peak result matrix. Comparing the peak between the odd group and even group, there is greater peak group which does not contain the bit reversal. Then the selected greater peak value is compared to the set capture threshold. If the value exceeds the set capture threshold, it means the capture has completed successfully. Otherwise it is failed. We can set the value of k is 20 when the signal power is very low. Then a number of coherent integration is integrated no-coherently. Finally, carry out the aforementioned peak judgment procedures. The specific procedures for the improved half-bit and feature-sequence algorithm are as follows.

(1) Set the k value and obtain the C_k.
(2) The input of IF signal is divided into odd group and even group.
(3) The contiguous k data blocks of odd group and even group is demodulated with NH code. It means that the ith data block is multiplied with the ith NH code of the first sequence.
(4) Sum the data blocks of demodulated kth group.
(5) The sum results carry out zero-padding and radix-2 operation, which obtains 1 ms data of odd group and even group.
(6) Then the output carries out FFT parallel correlation calculation with local copy signal.
(7) The results of different data blocks of odd arrays or even groups is integrated non-coherently.
(8) Compare the peak between the odd group and even group, there is greater peak group which does not contain the bit reversal.
(9) Calculate the capture detection value. If the value exceeds set capture threshold, it means the capture has completed successfully. If not, the input data is delayed 1 ms and repeat steps 3–8.

Figure 4 shows the flow chart.

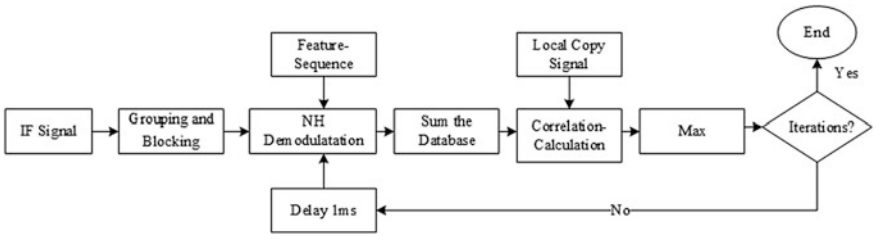


Fig. 4 Flow chart of proposed algorithm

5 Simulation Results and Analysis

In order to verify the validity of proposed the half-bit and feature-sequence algorithm, make simulation in Matlab. Set up simulation parameters. The sampling frequency is 4.092 MHz. The IF frequency of B2 signal of Beidou is 1.25 MHz. The D1 navigation message is a serial of random sequence including 1 and −1 generated by random function. The Doppler frequency of signal is 1000 Hz. The phase of Ranging-codes is 500. The predictive time of acquisition algorithm is 1 ms. The search Doppler frequency change is in the range of (−10 kHz to +10 kHz). The search frequency bandwidth is 500 Hz. In the experiment, the input IF signal of $B2_I$ is set consistently 80 ms. In order to avoid the influence of different threshold settings on the detection results, it means the capture has completed successfully when the search unit of result matrix peak corresponds to correctly the phase of Ranging-codes and Doppler frequency.

Firstly, under the same acquisition probability, the proposed half-bit and feature-sequence algorithm is compared to the traditional non-coherent integration algorithm. The feature sequence length value k of proposed algorithm is set 15 and the integration length of non-coherent integration algorithm is set 15 ms. The two algorithms are evaluated using emulated experiments based on the data of different signal-to-noise ratios and calculate their acquisition probability. Figure 5 shows the result of experiments and "Pd" is the abbreviation of "Probability of detection".

According to Fig. 5, compared with the traditional non coherent integration algorithm, the improved half-bit and feature-sequence acquisition algorithm is better in capturing the sensitivity. The simulation results show that the performance of proposed algorithm improves above 2 dB compared with non-coherent integration algorithm under same acquisition probability.

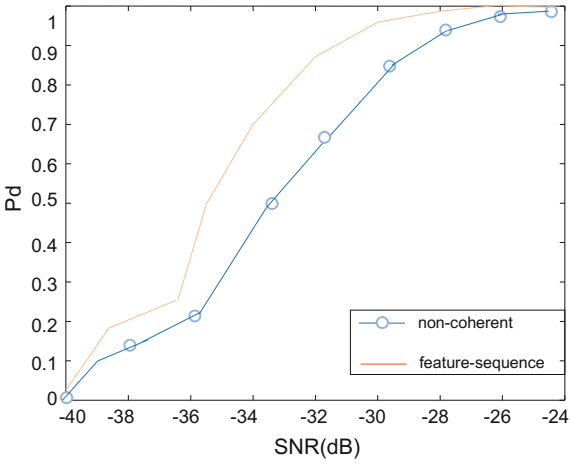


Fig. 5 Acquisition probability of under different SNR

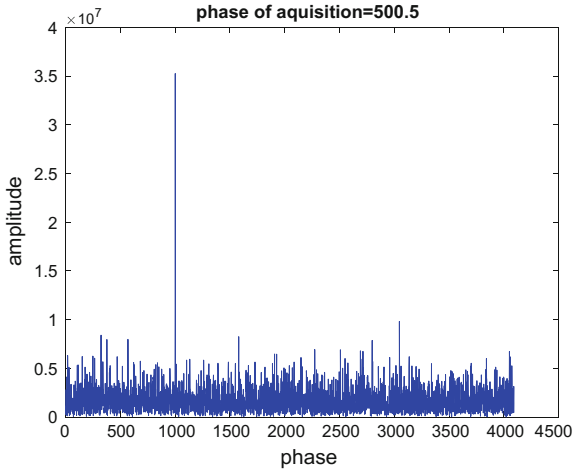


Fig. 6 SNR = −34 dB k = 15 the captured code phase of proposed algorithm

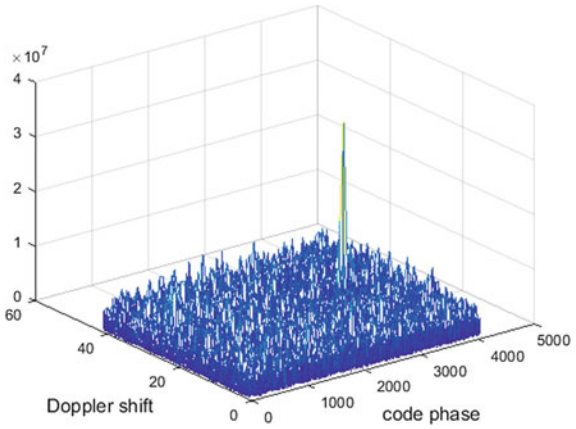


Fig. 7 SNR = −34 dB k = 15 the captured 3D result of proposed algorithm

In indoor environment, the SNR of weak signal $B2_I$ is about −34 dB. Hence, the feature sequence length value k of proposed algorithm set 15 carries out the SNR −34 dB weak $B2_I$ signal. The captured results are showed Figs. 6 and 7.

According to Figs. 6 and 7, in indoor environment the proposed algorithm captures very well the $B2_I$ signal. Besides in the process of acquisition, the noise interference is small, and the correlation value is prominent at the capture point, and the preference of acquisition is good.

The 15 ms integration length of non-coherent integration algorithm carries out t SNR −34 dB weak $B2_I$ signal. The captured results are showed Figs. 8 and 9.

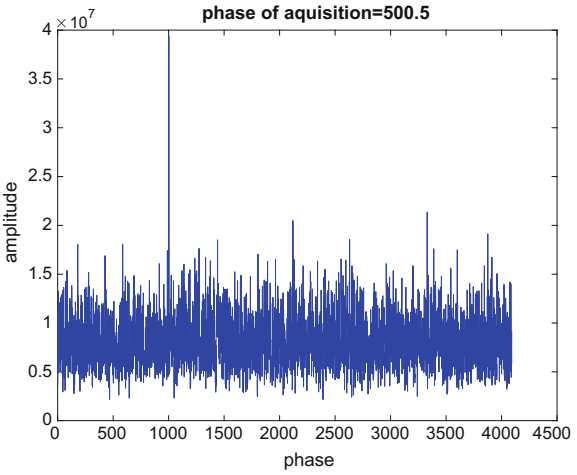


Fig. 8 SNR = −34 dB 15 ms the captured code phase of non-coherent algorithm

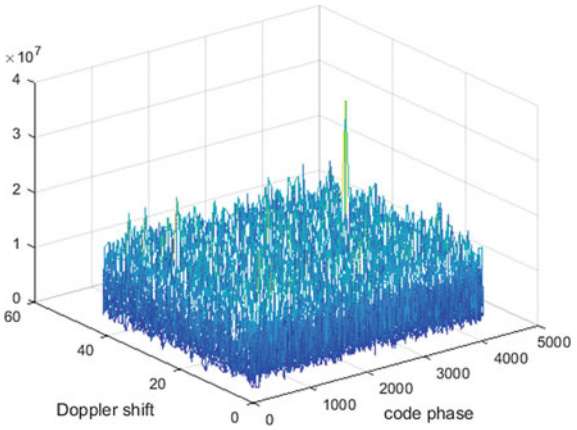


Fig. 9 SNR = −34 dB 15 ms the captured 3D result of non-coherent algorithm

Besides, the improved algorithm can determine the NH code position by the iteration of the flow chart while the phase of the ranging-code and Doppler frequency are captured. However, the non-coherent integration algorithm can't determine the NH code position. After capturing the phase of ranging-code and Doppler frequency, it is needed to other step to search the NH code position.

6 Conclusions

The NH code modulation of Beidou B2 satellite navigation signal causes the problem of high bit reversal rate, so it is difficult to directly apply the coherent integration over periods longer than 1 ms to capturing weak Beidou B2 satellite signal. In the paper, an improved acquisition algorithm is proposed which combines the feature sequence and half-bit algorithm according to the characteristics of B2 signal. The concept of NH code feature sequence is introduced and the flow chart of the improved algorithm is described in detail. Then the algorithm is verified with simulated $B2_I$ signal and compared to the non-coherent integration algorithm. The experimental result indicates that the performance of the proposed algorithm is not only improved about 2 dB compared to the non-coherent algorithm, but also does not need other steps to search the NH code phase in the same detection probability.

Acknowledgements The authors are grateful for the financial support from the Shaanxi Province Natural Science Foundation (2015JM6334).

References

1. Yang Y (2010) Progress, contribution and challenges of compass/Beidou satellite navigation system. Acta Geodaet Cartographica Sinica 39(1):1–6
2. China Satellite Navigation Office (2013) Beidou navigation satellite signal in space interface control document open service signal. Version 2.0. China Standardization, Beijing
3. Qin X, Han C, Xie Y (2010) A high sensitive fast acquisition algorithm suitable to implement in software GPS receiver. Acta Electronica Sinica 38(1):99–104
4. Liu T, Mao X (2013) Method of BD2 B1 signal generating. Comput Simul 30(9):70–73
5. Xu X, Fan J (2009) Acquisition algorithm of GPS software receiver. J Chin Inertial Technol 17(2):16–19
6. O'Driscoll C, Petovello MG, Lachapelle G (2008) Software receiver strategies for the acquisition and re-acquisition of weak GPS signals. J Guidance. Proceedings of the institute of navigation, national technical meeting, pp 843–854
7. Chan A, Games R (1990) On the linear span of binary sequences obtained from Q-ary M sequences, Q odd. In: IEEE transactions on information theory, vol 36

An Implementation of Navigation Message Authentication with Reserved Bits for Civil BDS Anti-Spoofing

Muzi Yuan, Zhicheng Lv, Huaming Chen, Jingyuan Li and Gang Ou

Abstract Since navigation message authentication (NMA) requires a modification in space segment and the signal broadcasted by satellites, negligent implementations of NMA may cause a failure in legacy GNSS receiver to process civil GNSS signal. This paper proposes an implementation of NMA applying reserved bits as signature area. Digital signature and key chain algorithms are employed to improve security and efficiency of the implement. In additional, performance of this implementation is investigated.

Keywords Anti-spoofing · Digital signature · Reserved bits · Navigation message authentication

1 Introduction

GNSS spoofing attack refers to the attack performed by broadcasting a counterfeit GNSS signal to the target receiver. Without extra information, the receiver is not able to distinguish an authorized signal from its fake replica. NMA is an effective method to identify authorized signal by authenticating the signature attached to the navigation message. Since NMA may include a modification in the structure of GNSS signal, the implementation should be designed carefully to maintain the compatibility for legacy receivers.

The spoofing attack neutralized by NMA can be detailed as the security code estimation and replay (SCER) attack [1]. Several methods of NMA based on ECDSA and TESLA were proposed recent years [1–7]. In latest research, performance of NMA was measured by two indicators as time to first authenticated fix (TTFAF) and time between authentications (TBA) in term of efficiency [8]. These

M. Yuan · Z. Lv · H. Chen · J. Li · G. Ou (✉)
School of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan, China
e-mail: ougangcs@139.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_6

works mainly concentrate on safety, feasibility and efficiency of NMA and ignore the affects to legacy receivers.

Compatibility is a significant characteristic of GNSS. Modifications in navigation message should guarantee the capability for legacy receivers. This work provides an implementation of NMA for BDS. The implementation attaches the signature to the reserved bits in navigation message. Legacy receivers need no upgrade to maintain the capability and other receivers can achieve the anti-spoofing ability by a modification in software.

The paper is organized as follows: model of SCER attack and basic principles of NMA are discussed in Sect. 2; the proposed strategy of NMA is demonstrated in Sect. 3; design of modified navigation message structure is introduced in Sect. 4; an analysis of performance is performed in Sect. 5.

2 Model of SCER Attack and Algorithms for NMA

Meaconing attack can be mitigated because there is always a delay in the spoofing signal to the authorized signal. Receivers can examine the time mark of all validate signal and pick up the signal with the minimum delay. However, SCER signal can be advance in phase than authorized signal. NMA is needed to detect and mitigate this kind of spoofing attack.

A SCER attacker can manipulate the PVT solution of the target receiver by broadcasting a well-designed GNSS signal. This counterfeit signal is composed by two elements: delayed pseudorange code and fake navigation message. The combination of authorized signal and its fake replica can be modeled as

$$Y_k = \alpha \hat{w}_{k-d} c_{k-d} + w_k c_k + N_k \tag{2.1}$$

Here in the model, $\hat{w}_{k-d}$ is the fake navigation message, c_{k-d} is the delayed pseudorange code, w_k and c_k are navigation message and pseudorange code broadcasted by authorized satellite. α is the energy gain of spoofing signal and N_k is an IID Gaussian noise. While authorized signal is jammed or fake signal is higher in power, the receiver will capture and track the fake signal. Thus $\hat{w}_{k-d}$ will be adopted as the navigation message, which may end up with a fake positioning result.

NMA works as a tool to identify whether the navigation message is generated by an authorized source. By adopting specific validation algorithms to navigation message text in $\hat{w}_{k-d}$ or w_k, receivers can distinguish authorized signal form the fake one. Algorithms applied by NMA mainly include ECDSA and TESLA.

2.1 The Elliptic Curve Digital Signature Algorithm (ECDSA)

First introduced in 1985 [9], the ECDSA is a public key digital signature standard operates on groups associated with an elliptic curve space. The signature can be generated by private key and plaintext navigation message. The verification can be performed through a public key to validate the signature and message [10]. The length of signature generated by standard 233-bit Koblitz curve is 466 bits [11].

Since the signature is unpredictable to attackers, the SCER attackers are not able to attach a valid signature to $\hat{w}_{k-d}$. Receivers can verify the signature and plaintext navigation message by ECDSA to implement the NMA.

The advantage of ECDSA is its security. The algorithm has been developed for years and is adopted as cryptography method in many scenarios. However, compared with the length of a 300-bit main frame, the signature may occupy large amount of extra bandwidth. The performance of TTFAF and TBA will be limited due to message bit rate.

2.2 The Timed Efficient Stream Loss-Tolerant Authentication (TESLA)

Introduced in 2002, the TESLA authenticates a navigation message through keys in a one-way chain [12]. The chain holds a property that it is easy to generate keys in one direction while it is unfeasible to generate keys in the opposite direction. Satellites broadcast the key in the chain in the inversed order of the generation process. The verification can be performed through verifying the key chain and a message authenticating code (MAC) generated by message and key.

A SCER attacker is not able to generate the chain from broadcasted information. Receivers can verify the key and MAC to identify an authenticated signal.

The advantage of TESLA is its low communication overhead. The length of key is possible to be relatively short because of the short expiration duration. However, length of the chain is limited due to the storage and calculation capability. Receivers need a periodic update of the root key in the chain.

Since there are advantages and disadvantages in both algorithms, this paper provides an implementation of NMA combining both algorithms.

3 Proposed NMA Strategy

The proposed NMA is a scheme that transforms w_k into d_k where d_k is navigation message modified for authentication. In user segment, receivers can recover w_k form d_k and get the extra information from d_k to verify w_k. Moreover, for

compatibility requirement, the implementation should ensure that w_k and d_k behaviors exactly the same in both modern receiver and legacy receiver.

As discussed in Sect. 2, EDCSA is adopted to achieve the security demand and ensure a stable method for key distribution; while TESLA provides an improvement in efficiency. Thus, the proposed scheme contains two forms with different algorithm, mentioned as super frame authentication (SFA) and main frame group authentication (MFGA).

3.1 Super Frame Authentication (SFA)

This authentication form is the method to authenticate GNSS signal as well as distribute keys for NMA. This form is based on the ECDSA. Root keys for TESLA and public keys for ECDSA are supposed to be acquired from this authentication from.

In this form, a super frame is separated logically into three parts as navigation message, key sets and signature, which is shown below. Here in the key sets segment, operant key refers to those keys used for current authentication stage, backup keys are keys used for authentication in next authentication stage. An authentication stage refers to the life time of a single authentication key (Fig. 1).

For every super frame, each satellite calculates a digital signature of navigation message and key set through a unique private key generated and updated by control segment for each satellite. Based on the principle of ECDSA discussed in Sect. 2, a 466-bit signature will be generated for every super frame. The plaintext and signature can be verified by operant public key, which is supposed to be stored in receivers. The signature is attached to the super frame and broadcasted to receivers.

Receiver is supposed to perform an attempt of authentication after a super frame is transmitted. An operant public key is loaded into the receiver in the process of manufacture and can be updated by a wire connection to authorized devices. The operant public key stored in the receiver is adopted to verify the super frame and its

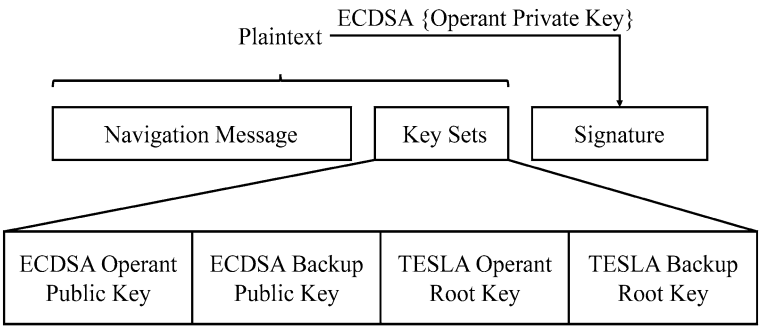


Fig. 1 Super frame structure of the proposed NMA scheme

signature. While the verification passes, public keys of ECDSA and root keys of TESLA are updated by key sets segment in navigation signal. The local time in receiver is loosely synchronized with the time broadcasted by satellite. When second of week (SOW) and week number (WN) in navigation message or the local time in receiver indicates the expiration of operant keys, the operant keys are supposed to be dropped and backup keys are employed to be operant keys. The lifetime of ECDSA key pairs are supposed to be relatively long (e.g. 1 year) because the algorithm is considered safe. Thus, the authentication stage duration of ECDSA is long enough that the communication overhead in public key update through wire connection can be ignored.

Receivers can perform one attempt of NMA for every duration of super frame, or 720 s in BDS. This authentication form is a fundamental authentication with long TBA. This authentication form supports the implementation of advanced authentication with short TBA. The advanced authentication form is main frame group authentication based on TESLA.

3.2 Main Frame Group Authentication (MFGA)

This authentication form is supposed to be performed every several durations of main frame. Those main frames which share the same authentication attempt are grouped as a main frame group. Every main frame group is separated logically into three parts as navigation message, message authentication code (MAC) and a key of previous MAC. The structure of main frame group is shown below (Fig. 2).

Authentication stage of MFGA is the time duration of a one-way chain. The chain is shared by all satellites and keys in chain are applied in the order arranged by PRN IDs. For every authentication stage, satellite i receives a random number s generated by control segment and is same for all satellites. Then the satellite calculates and truncates the hash result $F^{(i)}(s) = \{trunc[hash(s)]\}^{(i)}$ to designed key length. The result is set as the end of the key chain for TESLA. Here $F^{(i)}$ means

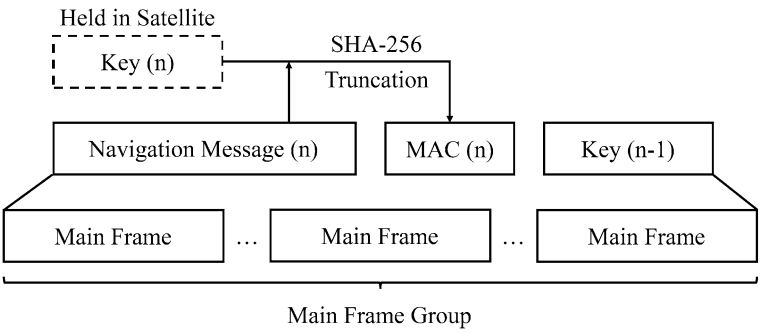


Fig. 2 Main frame group structure of the proposed NMA scheme

adopting hash and truncation function i times. Assuming the stage includes M complete duration of main frame group, the end of the chain $K_{M,i} = F^{(i)}(s)$. Based on $K_{M,i}$, the entire one-way chain can be generated by

$$K_{n-1} = F^{(N)}(K_n || SOW_e), n = M, M-1, \ldots, 2, 1 \tag{3.1}$$

Here N is a fixed number lager than the number of satellites in the constellation, marked as N_C. Operator $\|$ merges two string, SOW_e is the second of week in navigation message when the first bit of key K_n is released. The introduction of SOW_e is to add unpredictable part into the key chain. This structure can neutralize the attack performed by generating and storing all possible chains. $K_{0,i}$ is the operant root key for satellite i of the stage. A different chain generated by another random number s' is generated at the same time. The root key $K_{0,i}{}'$ is the backup root key. In ECDSA authentication from, keys broadcasted can be generated by $K_0 = F^{(N+1-i)}(K_{0,i})$ and $K_0{}' = F^{(N+1-i)}(K_{0,i}{}')$. These two keys are broadcasted to receivers in authentication form of ECDSA (Fig. 3).

In every duration of main frame group, the MAC is calculated by hashing and truncating the conjunction of navigation message and TESLA key for current duration. The generation of MAC can be described as $MAC_n = trunc[hash(M_n \| K_n)]$. Here M_n is the navigation message for current duration of main frame group. TESLA key for previous group, navigation message and MAC for current group are broadcasted to receivers in a main frame group. For the first group duration, $K_{0,i}$ is broadcast as the previous key.

After the transmission of a main frame group, the receiver first verifies the validation of TESLA key K_n. As the duration of key chain is supposed to be set as a fixed number for all authentication stages. The receiver can speculate the authentication stage and the location of $K_{n,i}$ in current chain by the local time which is loosely synchronized with satellite. While function $K_{n-1} = F(K_{n,i})$ to $K_{n,i}$ is repeated by $n * (N + 1) - i$ times, the validation of chain can be verified. If K_0 is the final result, current $K_{n,i}$ is validate.

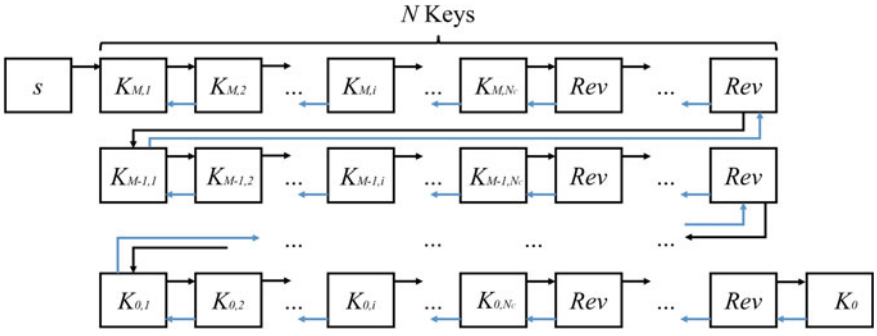


Fig. 3 Key chain of TESLA. Here in the figure Rev represents reserved key, *black arrows* are generation direction, *blue arrows* are release direction (Color figure online)

When the validation of TESLA key is verified, the authentication can be performed by calculating the hash result of the conjunction of navigation message in previous group duration and TESLA key K_n to get the MAC of previous group duration. If it matches the MAC received in previous group, the signal passes the authentication.

4 Design of Modification in Navigation Message

To maintain the capability of legacy receivers, all modifications in navigation message should not change significant information bits. Thus, the implementation of NMA is supposed to be arranged in the areas which are reserved.

4.1 Reserved Bits in D1 Navigation Message of BDS

The structure of BDS navigation message is documented in the interface control document (ICD). From the ICD, reserved bits in D1 navigation message can be located as Table 1 [13].

Here in Table 1 the slash mark (/) represents all available IDs. There are at least 25 bits reserved in every main frame and 2722 extra reserved bits in a super frame. For SFA scenario, key sets and signature are supposed to be arranged in those 2722 reserved bits. In MFGA form, MAC and TESLA key are supposed to be arranged in several 25-bit blocks.

Table 1 Reserved bits in BDS D1 navigation message

Frame ID	Page ID	Start bit	End bit	Length
/	/	12	15	4 bits
3	/	292	292	1 bit
4	/	43	43	1 bit
4	/	291	292	2 bits
5	/	43	43	1 bit
5	1-6	291	292	2 bits
5	7	286	292	7 bits
5	8	214	276	63 bits
5	9	219	276	58 bits
5	10	171	260	90 bits
5	11-24	51	228	178 bits

4.2 Proposed Structure of Navigation Message

Based on conclusion in Sect. 4.1, a practical arrangement of NMA functional bits can be proposed.

The scale of key sets in SFA is two ECDSA standard 233-bit Koblitz curve public keys (for each key up to 233 bits) and two TESLA root keys (for each up to 128 bits). Length of the signature to authenticate the super frame is 466 bits. Thus, reserved bits 51-228 in the fifth frame in page 11-24 can be applied for SFA.

For example, the operant key of ECDSA can be broadcasted in the reserved bits 51-228 in page 11 and 51-105 in page 12; the backup key of ECDSA can be broadcasted in the reserved bits 51-228 in page 13 and 51-105 in page 14; the operant key of TESLA can be broadcasted in bits 51-178 in page 15; the backup key of TESLA can be broadcasted in bits 51-178 in page 16; the signature can be broadcasted in bits 51-228 in page 16, 51-228 in page 17 and 51-160 in page 18. This arrangement ensures a rapid update of key sets through MFGA because keys are transmitted in several independent main frames.

The scale of MAC and key in MFGA is variable. The arrangement bases on the principle that all reserved bits in frame 1-4 in every page are supposed to be applied for TESLA authentication. Thus, total length of MAC and key is supposed to be integer multiples of 25 bits.

5 Performance Analysis

Performance of NMA can be indicated from the aspect of security and the aspect of efficiency. For security indicators, this paper takes the average attack time (AAT) and the possibility of false alarm (PFA) into consideration. For efficiency indicators, the TTFAF and TBA are taken into consideration.

5.1 Security

Introduced in a 1998 report, it takes 3 months of a network of 50,000 Pentium Pro 200 MHz machines to crack a ECCp-109 challenge in ECDSA [11]. Based on Moore's law and analysis of hardware associate [14], the estimating time to crack a ECDSA signature system by the network of 50,000 mainstream machines can be shown as below (Table 2).

For a normal expiration period (e.g. 1 year), ECDSA is secure for NMA in predictable feature.

Keys in TESLA key chain are operational only for a very short period (e.g. key chain designed for 32 satellites with 16,384 keys set to expire in one day, every key only stands for about 2 min). Crack of the chain requires series of crack of hash

Table 2 Crack time estimation for different length of key by different platforms

Key length	Pentium M	XC3S1000 [14]	ASIC
96	10 h	3.6 h	–
128	97 years	51 years	5 months
160	7.6×10^6 years	6.2×10^5 years	5.0×10^4 years
233 (estimate)	5.0×10^{16} years	4.1×10^{15} years	3.3×10^{14} years

function. Thus, the TESLA method would be secure if the estimate crack time for a single truncated hash function exceeds two times of the transmitting time of a main frame group [15]. Because in this scenario both operant key and backup key will expire before the attacker succeeds in crack process.

By Moore's Law, a mainstream processor may take 3.5×10^{23} attempts in 1 s in 2037 [14]. Thus, the length of key is supposed to meet the condition shown below.

$$2^{L_{key}} \geq 2R_a \left[ceil\left(\frac{L_{key} + L_{MAC}}{L_{Rev}} \right) L_{mf} T_b \right] \tag{5.1}$$

Here in 5.1 L_{key} is the length of key, L_{MAC} is the length of MAC, L_{Rev} is the length of reserved bits in every main frame (25 bits in BDS), $ceil(s)$ is the minimum integer not smaller than s, L_{mf} is the length of main frame (1500 bits in BDS) and T_b is the interval of chip (0.02 s in BDS). For example, an 88-bit key and 12-bit MAC can work securely ($k = 4$).

The PFA changes with the bit error ratio (BER) in the receive process. Every single error bit can lead to an FA. Thus, PFA of both two authentication forms can be shown as Fig. 4.

Form Fig. 4 MFGA has a much lower PFA than SFA. Thus, MFGA is the mainstream authentication method in this implementation.

5.2 *Efficiency*

The TTFAF of the proposed implementation is discussed in two scenarios: initial NMA and continuous NMA.

Initial NMA refers to the scenario where the receiver losses capture of authorized signal for over two expiration durations of TESLA key chain. In this scenario TESLA root keys stored in the receiver are not applicable. Hence the receiver is supposed to perform an entire process of SFA first to access the operant and backup key for TESLA. In this scenario, the TTFAF is

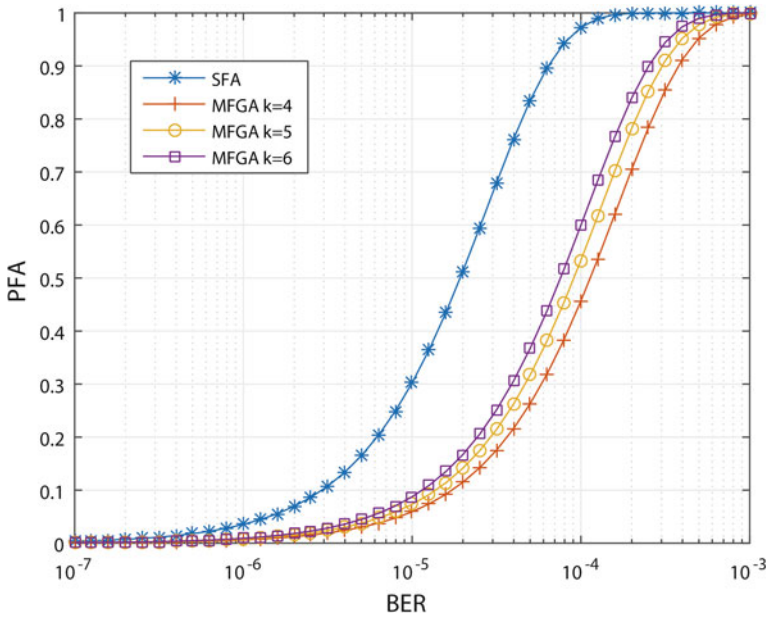


Fig. 4 PFA in different authentication forms under different BER conditions

$$\overline{TTFAF}_{SFA} = \frac{1}{L_{sf}} \cdot \sum_{n=1}^{L_{sf}} \frac{(n + L_{sf})T_b}{(1 - BER)^{L_{sf}}} \tag{5.2}$$

Here in 5.2 L_{sf} is the length of super frame (36,000 bits in BDS). For BER = 10^{-6} the average TTFAF is 1119 s (18.65 min).

Continuous NMA refers to the scenario where the root keys stored in the receiver are applicable. In this scenario, receiver can authenticate the signal through MFGA. The TTFAF is

$$\overline{TTFAF}_{MFGA} = \frac{1}{kL_{mf}} \cdot \sum_{n=1}^{kL_{mf}} \frac{(n + 2kL_{mf})T_b}{(1 - BER)^{k(L_{mf} + L_{Rev})}} \tag{5.3}$$

Here in 5.3 L_{mf} is the length of main frame (1500 bits in BDS), k is number of main frame in a group and L_{Rev} is the length of reserved bits applied for authentication message transmission in main frame (25 bits in BDS). For k = 4 and BER = 10^{-6} the average TTFAF is 311 s (5.18 min).

The TBA is the time interval between two MFGA, which is

$$\overline{TBA}_{MFGA} = \frac{kL_{mf}}{(1 - BER)^{kL_{mf}}} \tag{5.4}$$

Here k is number if main frame in a group. For k = 4 and BER = 10^{-6} the average TBA is 124 s (2.07 min).

6 Conclusion

This paper offers a practical implementation of NMA for BDS. The proposed implementation combines ECDSA and TESLA, maintains the capability of legacy receivers in NMA signal. The proposed strategy enables receivers to authenticate signal every 2 min. The implementation does not severely degrade the performance in time to first positioning fix of an authentication receiver with respect to a legacy receiver.

References

1. Wesson KD, Rothlisberger MP, Humphreys TE (2011) A proposed navigation message authentication implementation for civil GPS anti-spoofing. In: Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2011), pp 3129–3140
2. Humphreys TE (2013) Detection strategy for cryptographic GNSS anti-spoofing. IEEE Trans Aerosp Electron Syst 49:1073–1090
3. Hernaandez IF, Rijmen V, Granados GS et al (2014) Design drivers, solutions and robustness assessment of navigation message authentication for the Galileo Open Service. In: Proceedings of the 27th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2014), pp 2810–2827
4. Wullems C, Pozzobon O, Kubik K (2005) Signal authentication and integrity schemes for next generation global navigation satellite systems. In: Proceedings of the European navigation conference
5. Lo S, Enge P (2010) Authenticating aviation augmentation system broadcasts. In: Proceedings of IEEE/ION PLANS 2010, Indian Wells, CA, pp 708–717
6. Curran JT, Paonni M, Bishop J (2014) Securing the Open-Service: a candidate navigation message authentication scheme for Galileo E1 OS. In: European navigation conference ENC 2014, Rotterdam
7. Kerns AJ, Wessons K, Humphreys T (2014) A blueprint for civil GPS navigation message authentication. In: Proceedings of IEEE/ION PLANS 2014, Monterey, CA, pp 262–269
8. Hernandez IF, Rijmen V, Granados GS et al (2016) A navigation message authentication proposal for the Galileo Open Service. J Inst Navig 63:85–102
9. Miller VS (1985) Use of elliptic curves in cryptography. In: International cryptology conference
10. Menezes AJ, Oorschot PCV, Vanstone SA (2001) Handbook of applied cryptography. CRC Press
11. Johnson D, Menezes AJ, Vanstone SA (2001) The elliptic curve digital signature algorithm (ECDSA). Certicom Corporation

12. Perrig A, Canetti R, Tygar JD, Song D (2002) The TESLA broadcast authentication protocol. CryptoBytes 5(2):2–13
13. Anonymous (2013) BeiDou system signal in space interface control document. Management Office of China GNSS
14. Gueneysu T, Paar C, Pelzl J (2007) Attacking elliptic curve cryptosystems with special-purpose hardware. In: Field programmable gate arrays, pp 207–215
15. Caparra G, Sturaro S, Laurenti N, Wullems C (2016) Evaluating the security of one-way key chains in TESLA-based GNSS navigation message authentication schemes. In: Proceedings of 2016 international conference on localization and GNSS
16. Anonymous (2001) Vulnerability assessment of the transportation infrastructure relying on the global positioning system. John A. Volpe National Transportation Systems Center, US

Performance Analysis of Signal Diversity Reception for Large Aperture Array in Beidou RDSS System

Haodong Jiang, Jingyuan Li, Xiangwei Zhu and Gang Ou

Abstract Due to the fact that the double-satellite active positioning function of Beidou RDSS system needs the users are at least double-layer overlapped by GEO satellites, so the ground central station (GCS) of RDSS system can get several same copies of user instation signals through different large aperture antenna. The RDSS GCS can take the advantage of received redundant signal to meet the requirement of diversity reception. It can increase the signal-to-noise ratio (SNR) of the received signal and reduce the code error rate of the received instation signal. The diversity gain of instation signal is depended on the error of pseudo-code phase and carrier phase. Based on the algorithm of Maximum-Ratio-Combination (MRC) and without taking into account of the influence of SNR on the tracking accuracy, both theoretical analysis and simulation results show that when the pseudo-code phase error is less than 0.1 chip and carrier phase error is less than 0.1477 rad, the diversity gain of two signals can greater to 2 dB compared to the signal without tracking error.

Keywords Beidou · RDSS · MRC · Diversity gains · Code error rate

1 Introduction

Beidou RDSS business is an effective means of contact navigation and communication [1]. Since its completion, the cost-effectiveness ratio is satisfactory [2]. Beidou RDSS system consists of five GEO satellites and usually the instation beams will multiple coverage. The GCS utilizes the signals which received by different large aperture antenna in the overlapping area to realize the diversity reception.

Beidou RDSS short message service requires the bit error rate (BER) is below to 10^{-5} (SNR is 9.6 dB). However, we need data transmission which requiring the

H. Jiang · J. Li (✉) · X. Zhu · G. Ou
Satellite Navigation Research and Development Center,
National University of Defense Technology, Changsha, China
e-mail: crazyjianghaodong@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_7

BER is below to 10^{-7} (SNR is 11.3 dB). So we come up to the diversity gain is greater than 1.9 dB index based on this. The purpose of diversity reception is to overcome the fading effect in the signal transmission and reduce the BER in the data demodulation. Single-channel signal is easy to be affected by the sudden change of the transmission channel. However, the probability that multi-channel signals are affected by the transmission channel at the same time is much smaller. Therefore, taking the advantage of the redundancy between the signal to achieve diversity reception and reduce the BER has important research value and application prospect.

2 Theory of Beidou RDSS Business Diversity Reception

The analog diversity is usually used in the mobile communication since the arriving time of multipath signals is roughly equal. But the signal transmission time delay is large through different GEO satellite, so we adopt digital diversity method in this paper.

2.1 Flowchart of Beidou RDSS Diversity Reception

User diversity reception of instation signal process is shown in Fig. 1. Different user instation signal which transmitted by different GEO satellite is received by different ground large antenna. The antenna receiving link accomplishes the signal sampling, down-conversion, filtering, analog to digital conversion and signal synchronization. And the signal will be sent to the signal processing unit and after processing the results will be sent to the ground terminal business processing server.

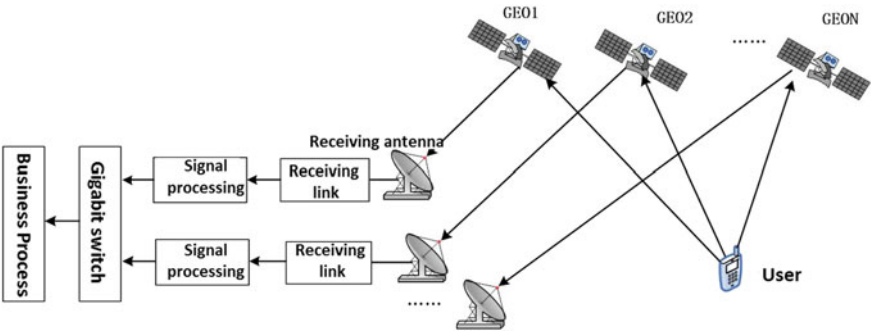


Fig. 1 The reception process of user instation signal

2.2 The Model of Instation Signal and the Principle of Diversity Algorithm

Beidou RDSS intermediate frequency (IF) discrete form signal can be represented as

$$r_i(k) = \sqrt{2}aD(k)C(k)\sin\left(2\pi f_i \frac{1}{f_s}k\right) \tag{2.1}$$

where k is the sampling point; a is the amplitude of the signal; D(k) is the data code; C(k) represents for the spread spectrum code [3]. In this paper, the IF signal was built for the instation signal. The intermediate frequency f_i = 1.5 MHz, the sampling rate fs = 4.6875 MHz, the spread spectrum code rate is 4.08 Mcps and the data rate is 500 bps. In order to check the stability of the test system, 30 bit short length data code (60 ms) is chosen and in each experiment the Monte Carlo simulation time is 10,000.

We use the playback to check out the validity and feasibility of the signal. And the signal spectrum is shown below (Fig. 2).

The diversity receiver receives the multipath signals can be represented as:

$$r = \sum_{i=1}^{i=N} \omega_i r_i \tag{2.2}$$

where r_i is the received signal of the large aperture antenna i. ω_i is the weighted coefficient for the channel i. When choosing different weighted coefficient, different way of merger is constituted. There are three common way of diversity merger [4].

Fig. 2 The IF signal spectrum

(1) Selective combining (SC).
Selective combining means detecting all the diversity branch and choosing a branch of the signal which has the highest SNR as the output of the combiner. According to this, the weighted coefficient of the signal branch which has the highest SNR is 1 and the other weighted coefficient is 0. After data demodulation, the two IF signal branch compare their SNR and select the branch with the higher SNR as output. Selective combining method is simple and easy to implement. However, the branch we not chosen is a loss of the signal information.
(2) Maximum ratio combining (MRC).
Maximum ratio combining is a best way to merge. The weighted coefficient of each branch is decided by the SNR. And for the MRC method, the output SNR is the sum of each branch's SNR. So, when multiple signals are very poor and there is no signal can be demodulated separately, MRC method is still likely to demodulate the right signal.
(3) Equal gain combining (EGC).
It is no need to weight for the signal for the EGC method because the weighted coefficient is equal. The EGC method is simple and its performance is close to the MRC.

In this paper, the MRC is adopted and the two branch signals' coherent accumulation value is merged. Since the simulation in this paper set the two branch signals have the same SNR, the weighted coefficient can be chosen as both 1. Therefore, under the premise, the MRC method is also the EGC method in this paper.

3 Demodulation Loss Analysis for Single-Channel Signal

RDSS instation signal adopt the way of I, Q coherent demodulation. And the coherent integration results are shown below [5 – 8].

$$I_p(k) = \frac{a}{2} D(k) R(\Delta\tau) \sin c(\Delta f T_{coh}) \times \cos(\pi \Delta f T_{coh} + \Delta\phi_0) + n_I \tag{3.1}$$

$$Q_p(k) = \frac{a}{2} D(k) R(\Delta\tau) \sin c(\Delta f T_{coh}) \times \sin(\pi \Delta f T_{coh} + \Delta\phi_0) + n_Q \tag{3.2}$$

where, $R(\cdot)$ is the autocorrelation function of the pseudo code. $\Delta\tau$, Δf, $\Delta\phi_0$ represent for the code phase mismatch, frequency mismatch, carrier initial phase mismatch between the local signal and the received signal. T_{coh} is the coherent integration time.

3.1 Pseudo-Code Phase Mismatch Loss Analysis

When the local carrier frequency is the same as the received signal frequency, $\Delta f = 0$, $\sin c(\Delta f T_{coh}) = 1$, ignoring the carrier initial phase difference and the influence of noise, the pseudo-code phase mismatch loss is [9]

$$L_R = 10\log 10([R(\Delta\tau)]^2) \tag{3.3}$$

In the tracking phase, code lock loop can take control of the code phase error in 0.5 chip. So when the pseudo-code phase error are 0.1, 0.3, 0.5 chip, the SNR loss for the code phase mismatch are 0.9151, 3.0980, 6.0206 dB.

3.2 Frequency Mismatch Loss Analysis

When the local pseudo-code phase and received signal pseudo-code phase are aligned, $\Delta\tau = 0$, $R(\Delta\tau) = 1$, ignoring the carrier initial phase difference and the influence of noise, the frequency mismatch loss is

$$L_f = 10\log 10(\sin c^2(\Delta f T_{coh}) \times \cos^2(\pi \Delta f T_{coh})) \tag{3.4}$$

As we can see, frequency mismatch relates to $\Delta f T_{coh}$, and usually $\Delta f T_{coh}$ is small. So the demodulation loss of frequency mismatch is minimal. But, on the other hand, the frequency mismatch will lead to the accumulation of phase, and we must use the method of phase compensation to offset the impact.

3.3 Carrier Phase Mismatch Loss Analysis

Ignoring the frequency, pseudo-code phase difference and the influence of noise, the carrier phase mismatch loss is

$$L_\theta = 10\log 10(\cos^2(\Delta\phi_0)) \tag{3.5}$$

In the accurate tracking phase, carrier loop can keep the carrier phase within 0.4 rad. When the carrier phase error is 0.1, 0.2, 0.3 rad, carrier phase mismatch loss is 0.0435, 0.1749, 0.3969 dB.

3.4 Summary

For single-channel signal demodulation loss is mainly caused by the pseudo-code phase mismatch and the carrier phase mismatch. Demodulation loss formula is

$$L = 10\log 10([R(\Delta\tau)]^2 \cos^2(\Delta\phi)) \tag{3.6}$$

4 Performance Analysis and Simulation for Two Channel Signals Diversity Reception Algorithm

Through the above analysis, we have made it clear that, for single-channel signal demodulation loss is mainly caused by the pseudo-code phase error and the carrier phase error. In the following question, we discuss two channel signals' diversity gain based on the MRC algorithm.

4.1 Performance Analysis and Simulation for Pseudo-Code Phase Mismatch Diversity Gain

4.1.1 Theoretical Analysis for Pseudo-Code Phase Mismatch Diversity Gain

Assume that the pseudo-code phase mismatch of the first channel signal is $\Delta\tau_2$ and the second is $\Delta\tau_2$. And the two pseudo-code phase mismatch are range from −0.5 chip to 0.5 chip. Assume that two signals have the same SNR and do not have frequency mismatch and carrier initial phase mismatch. From Sect. 3.1, the two signal diversity merger is the sum of two channel signals' coherent accumulation value. When we sum up the coherent value, the signal energy increment speed is squared but the accumulation of noise item is first power. So the SNR is increased. We concluded that, compared to the first signal, the diversity gain is

$$G_1 = 10\log 10\left(\frac{(\mathrm{R}(\Delta\tau_1) + \mathrm{R}(\Delta\tau_2))^2}{2\mathrm{R}(\Delta\tau_1)^2}\right) \tag{4.1}$$

Compared to the second signal, the diversity gain is

$$G_2 = 10\log 10\left(\frac{(\mathrm{R}(\Delta\tau_1) + \mathrm{R}(\Delta\tau_2))^2}{2\mathrm{R}(\Delta\tau_2)^2}\right) \tag{4.2}$$

And compared to the signal without pseudo-code phase mismatch, the diversity gain is

$$G_0 = 10\log 10\left(\frac{(\mathrm{R}(\Delta\tau_1)+\mathrm{R}(\Delta\tau_2))^2}{2}\right) \tag{4.3}$$

The theoretical diversity gain is given below.

From Fig. 3, we can conclude that when the first signal pseudo-code phase mismatch is getting greater, the diversity gain compared to the first signal is getting greater. But when the second signal pseudo-code phase mismatch is getting smaller, the diversity gain compared to the first signal is getting greater. When $|\Delta\tau_1| = 0.5$ chip and $\Delta\tau_2 = 0$ chip, diversity gain compared to the first signal reach maximum 6.5321 dB; When $\Delta\tau_1 = 0$ chip and $|\Delta\tau_2| = 0.5$ chip, diversity gain compared to the first signal reach minimum 0.5115 dB. We can conclude the diversity gain figure compared to the second signal by formula 4.2. And the figure is similar to Fig. 3. Also, we can come up to the similar conclusion.

From Fig. 4, when the two channel signals do not have pseudo-code phase error, diversity gain reach maximum 3.0103 dB; When the two channel signals pseudo-code phase error are 0.5 chip, diversity gain compared to the signal without pseudo-code phase error reach minimum −3.0103 dB.

If we assume that diversity gain compared to the signal with better tracking accuracy is more than $\boldsymbol{a}$ dB (might as well set the first signal has high estimation precision), we can conclude that

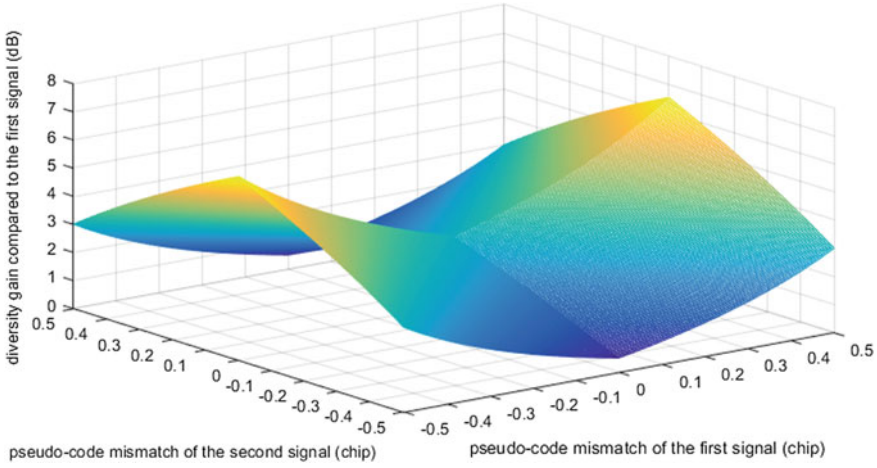


Fig. 3 The diversity gain compared to the first signal

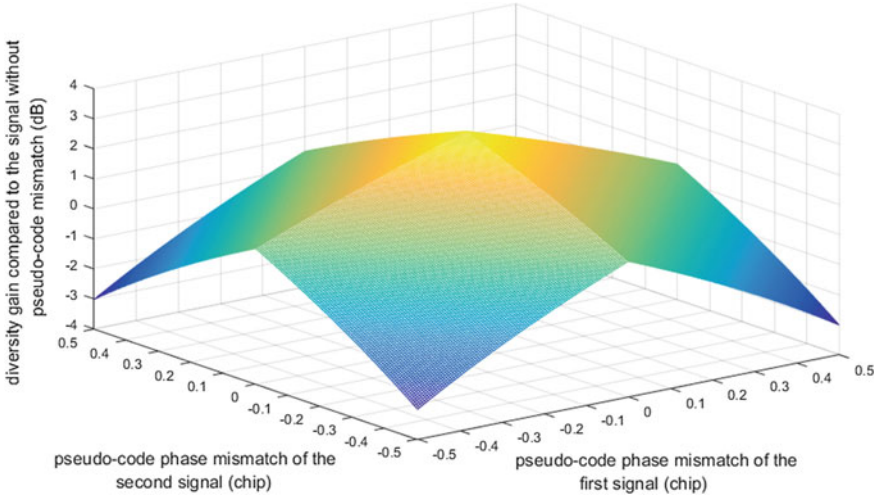


Fig. 4 The diversity gain compared to the signal without pseudo-code phase error

$$G_1 = 10\log 10\left(\frac{(\mathrm{R}(\Delta\tau_1)+\mathrm{R}(\Delta\tau_2))^2}{2\mathrm{R}(\Delta\tau_1)^2}\right) > a \tag{4.4}$$

$$\mathrm{R}(\Delta\tau_1) > \mathrm{R}(\Delta\tau_2) \tag{4.5}$$

And we can calculate the second signal pseudo-code phase error must fulfill the formula below.

$$(\sqrt{2}\times 10^{\frac{a}{20}}-1)R(\Delta\tau_1) < R(\Delta\tau_2) < R(\Delta\tau_1) \tag{4.6}$$

For example, when $\boldsymbol{a}$ = 2, it means that we have 2 dB diversity gain compared to the first signal which have better tracking accuracy and the second signal pseudo-code phase error must fulfill the relation below.

$$|\Delta\tau_1| < |\Delta\tau_2| < 0.7804|\Delta\tau_1| + 0.2196 \tag{4.7}$$

When guaranteed the diversity gain compared to the first signal, relation between the two signal pseudo-code phase is shown below (Table 1).

Table 1 Relation between the two signal pseudo-code phase

$\lvert\Delta\tau_1\rvert$	$\lvert\Delta\tau_2\rvert$
0.1	$0.1 < \lvert\Delta\tau_2\rvert < 0.29764$
0.2	$0.2 < \lvert\Delta\tau_2\rvert < 0.37568$
0.3	$0.3 < \lvert\Delta\tau_2\rvert < 0.45372$

According to the formula G_0 and Fig. 4, we can conclude that if we want to have $\boldsymbol{b}$ dB diversity gain compared to the signal without pseudo-code phase error, we get the formula below.

$$G_0 = 10\log_{10}\left(\frac{(\mathrm{R}(\Delta\tau_1)+\mathrm{R}(\Delta\tau_2))^2}{2}\right) > b \tag{4.8}$$

We conclude that the pseudo-code phase error of the two channel signals must fulfill the relation below.

$$|\Delta\tau_1|+|\Delta\tau_2|<2-\sqrt{2}\times 10^{\frac{b}{20}} \tag{4.9}$$

For example, if we want to have 2 dB diversity gain compared to the signal without pseudo-code phase error, we conclude the conclusion below.

$$|\Delta\tau_1|+|\Delta\tau_2|<0.2196\,\text{chip} \tag{4.10}$$

4.1.2 Simulation for Pseudo-Code Phase Mismatch Diversity Gain

Simulation parameter settings: the two signal IF f_i = 1.5 MHz; sampling rate f_s = 4.6875 MHz; pseudo-code rate is 4.08 MHz; data rate is 500 bps. The data length is 30 bit (60 ms) and in every experiment the Monte Carlo simulation time is 10,000. The pseudo-code phase error of the first signal is 0.5 chip and the second is 0.25 chip. Besides, the two signals do not have frequency and carrier phase error. Additive white Gaussian noise was chosen. We simulate the scenario with low SNR (−40 dB) and high SNR (−28 dB) separately. The results are shown below (Table 2).

From Table 2 we can see, basically the simulation results are equal to the theoretical value. The correctness and feasibility of the theory are verified.

Table 2 The simulation results for the pseudo-code phase mismatch

	SNR = −40 dB	SNR = −28 dB
First signal BER	31.24% (−45.94 dB)	2.64% (−33.99 dB)
Second signal BER	23.36% (−42.5 dB)	0.18% (−30.42 dB)
Diversity reception BER	19.51% (−41.05 dB)	0.03% (−29.02 dB)
Diversity gain compared to first signal	4.89 dB	4.97 dB
Theoretical gain compared to first signal	4.9485 dB	4.9485 dB
Diversity gain compared to second signal	1.45 dB	1.43 dB
Theoretical gain compared to second signal	1.4267 dB	1.4267 dB
Diversity gain compared to ideal signal	−1.05 dB	−1.02 dB
Theoretical gain compared to ideal signal	−1.0721 dB	−1.0721 dB

4.1.3 Summary

Analysis for the two channel signals, we conclude that

(1) Without taking into account of the influence of SNR on the tracking accuracy, the two signals' diversity gain have no connection with SNR but is relevant to the pseudo-code phase error of the two signals.
(2) When we want to have ***a*** dB diversity gain compared to the signal which have better pseudo-code tracking accuracy, the two signal pseudo-code phase must fulfilled formula 4.6.
(3) When we want to have 2 dB diversity gain compared to the signal without pseudo-code phase error, the summation of two signal pseudo-code phase must less than 0.2196 chip.

4.2 Performance Analysis and Simulation for Carrier Phase Mismatch Diversity Gain

In the carrier tracking phase, the frequency error can control within 1 Hz. And in the project implementation, the carrier phase is compensated in every a short time. So the carrier phase error can be controlled within 0.4 rad completely.

4.2.1 Theoretical Analysis for Carrier Phase Mismatch Diversity Gain

Assume that the first signal carrier phase mismatch is ϕ_1, the second signal is ϕ_2 and there not exist other error factors. Through Sect. 3.3, we conclude that the diversity gain compared to the first signal is

$$G_1 = 10\log 10\left(\frac{(\cos(\phi_1)+\cos(\phi_2))^2}{2\cos^2(\phi_1)}\right) \tag{4.11}$$

Compared to the second signal, the diversity gain is

$$G_2 = 10\log 10\left(\frac{(\cos(\phi_1)+\cos(\phi_2))^2}{2\cos^2(\phi_2)}\right) \tag{4.12}$$

And compared to the signal without carrier phase mismatch, the diversity gain is

$$G_0 = 10\log 10\left(\frac{(\cos(\phi_1)+\cos(\phi_2))^2}{2}\right) \tag{4.13}$$

The theoretical diversity gain is given below (Figs 5 and 6).

The conclusion of carrier phase mismatch diversity gain can be drawn from the analysis of the pseudo-code phase mismatch analogy. From Fig. 6, the curve is smooth compared to the pseudo-code phase mismatch figure. It is that the carrier phase mismatch demodulation loss is cosine-squared item and the pseudo-code phase mismatch is squared item. Therefore, the carrier phase mismatch diversity gain curve is smooth.

4.2.2 Simulation for Carrier Phase Mismatch Diversity Gain

The simulation parameter settings are the same with Sect. 4.1.2. The carrier phase error of the first signal is 0.3 rad and the second is 0.4 rad. Additive white Gaussian noise was chosen. We simulate the scenario with low SNR (−40 dB) and high SNR (−32 dB) separately. The simulation results are shown below (Table 3).

From Table 3 we can see, basically the simulation results are equal to the theoretical value. The correctness and feasibility of the theory are verified.

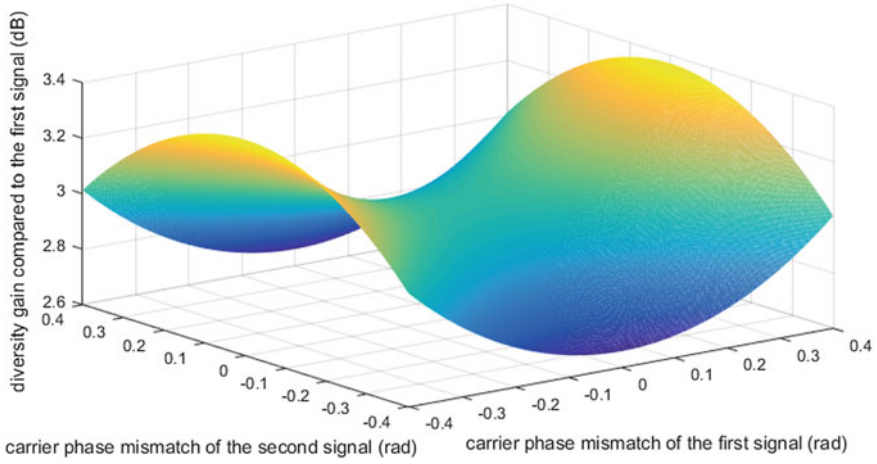


Fig. 5 The diversity gain compared to the first signal

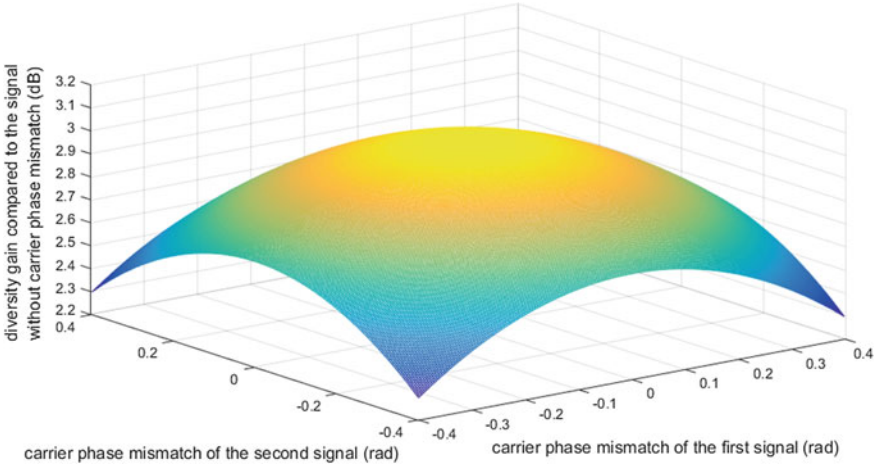


Fig. 6 The diversity gain compared to the signal without carrier phase error

Table 3 The simulation results for the carrier phase mismatch

	SNR = −40 dB	SNR = −28 dB
First signal BER	17.84% (−40.44 dB)	1.01% (−32.41 dB)
Second signal BER	18.61% (−40.72 dB)	1.24% (−32.71 dB)
Diversity reception BER	9.929% (−37.55 dB)	0.053% (−29.43 dB)
Diversity gain compared to first signal	2.89 dB	2.98 dB
Theoretical gain compared to first signal	2.8531 dB	2.8531 dB
Diversity gain compared to second signal	3.17 dB	3.28 dB
Theoretical gain compared to second signal	3.1704 dB	3.1704 dB
Diversity gain compared to ideal signal	2.45 dB	2.57 dB
Theoretical gain compared to ideal signal	2.4562 dB	2.4562 dB

4.3 Chapter Summary

(1) This chapter analyzes the diversity gain under the condition that the two channel signals have pseudo-code phase mismatch and carrier phase mismatch. We can draw the following three diversity gain formula. The diversity gain compared to the first signal is

$$G_1 = 10\log 10\left(\frac{\left(R(\Delta\tau_1)\cos(\phi_1)+R(\Delta\tau_2)\cos(\phi_2)\right)^2}{2[R(\Delta\tau_1)]^2\cos^2(\phi_1)}\right) \tag{4.14}$$

Compared to the second signal, the diversity gain is

$$G_2 = 10\log 10\left(\frac{(R(\Delta\tau_1)\cos(\phi_1) + R(\Delta\tau_2)\cos(\phi_2))^2}{2[R(\Delta\tau_2)]^2\cos^2(\phi_2)}\right) \tag{4.15}$$

And compared to the signal without carrier phase mismatch and pseudo-code phase mismatch, the diversity gain is

$$G_0 = 10\log 10\left(\frac{(R(\Delta\tau_1)\cos(\phi_1) + R(\Delta\tau_2)\cos(\phi_2))^2}{2}\right) \tag{4.16}$$

In addition, we made the simulation verification. The simulation results ensure the correctness and feasibility of the theory.

(2) Without taking into account of the influence of SNR on the tracking accuracy, the two signals' diversity gain have no connection with SNR but is relevant to the pseudo-code phase error and the carrier phase error of the two signals.

(3) In engineering practice, if we want to ensure the technical index of the diversity gain, the pseudo-code phase error and the carrier phase error of the two signals are strict limited. And the diversity gain is given by formula 4.16.

5 Conclusion

This paper started from the demodulation loss of single-channel signal, focusing on the two channel signals' diversity gain under the condition that the two signal have pseudo-code phase and carrier phase mismatch. Besides, the simulation was made to verify the correctness of the theory. We came up with the following conclusions.

(1) Single-channel signal demodulation loss is mainly caused by the pseudo-code phase mismatch and the carrier phase mismatch. When the pseudo-code phase error is 0.1, 0.3, 0.5 chip, the demodulation loss is 0.9151, 3.0980, 6.0206 dB; When the carrier phase error is 0.1, 0.2, 0.3 rad, the demodulation loss is 0.0435, 0.1749, 0.3969 dB.

(2) Based on the MRC diversity algorithm implementation, the diversity gain depends on the two signals' pseudo-code error and carrier phase error. Theoretical analysis and simulation show that if we want to have 2 dB diversity gain compared to the signal without pseudo-code phase error, the summation of two signals' pseudo-code phase error less than 0.2196 chip; When we want to get 2 dB diversity gain compared to the signal without carrier phase error, the two channel signals' carrier phase must fulfill the formula $\cos(\phi_1) + \cos(\phi_2) < 1.7804$. From formula 4.16 we can calculate that when the two signals have the same tracking performance, if the pseudo-code phase error is less than 0.1 chip and the carrier phase error is less than 0.1477 rad, the diversity gain can greater than 2 dB compared to the signal without pseudo-code phase error and carrier phase error.

References

1. Tan SS (2007) Satellite navigation and positioning engineering, 1st edn. Beijing, national defense industry press, pp 24–27
2. Zhao LW, Li GX, Dai WH, Miao ZM (2015) Opportunities and challenges of Beidou RDSS applying to navigation controls. J Glob Position Syst: 3
3. Xie G (2009) GPS principle and receiver design. The electronic industrial press, pp 244–246
4. Fan ZY (2010) The comparison and simulation of three merger way diversity reception. J Inf Sci Technol 11:50–51
5. Zhang F (1994) Polarization diversity improving telemetry receiving performance of tracking system. J Rem Sens 1:15–16
6. Rinder P, Bertelsen N (2004) Design of a single frequency GPS software receiver. Master Thesis, Aalborg University, Denmark
7. Yin Y, Zhao MS, Cai F (2006) Tracking loop design and simulation for digital DSSS receivers. Comput Simul 11:325–331
8. Benedetto JJ (2007) A software-defined GPS and GALILEO receiver. Satellite Division of the Institute of Navigation, pp 1632–1637
9. Li HJ, Liu L, Du XJ (2009) High sensitivity GPS signal capture technology research. J Projectiles Rockets Missiles Guidance 29(4)

Study on Multipath Model of BDS/GPS Signal in Urban Canyon

Yuze Wang, Xin Chen, Peilin Liu and Wenxian Yu

1 Introduction

In the field of high precision satellite navigation applications, multipath is one of the most important interferences that decrease the positioning accuracy. With the development of high-precision satellite navigation technology, various types of multipath mitigation technology has been proposed [1], but none of them can completely solve this problem. Therefore, in order to study the influence of multipath interference and improve multipath mitigation technology, characteristic model of multipath signal has attracted much attention.

Multipath model can be divided into two types according to different method, namely, accurate model and statistical model. The accurate model mainly uses the ray tracing method to accurately calculate each propagation path of the electromagnetic wave under a given 3D scene model. The advantage of this model is that the parameter characteristics are detailed and accurate. The statistical model often gives the probability density distribution model of multipath characteristic based on experiment measurement. This type of model is suitable for the large scale scene which is difficult to be calculated.

In recent years, a large number of multipath model have been proposed. Lippincott gives a multipath model for multipath delay and power fading in 1992 [2]. The DLR group impulse signal on a Zeppelin in different environment, and the impulse response data is used to model the multipath characteristics [3]. JAN P. WEISS build the 3D model for fighters and warships, and calculate the

Y. Wang · X. Chen (✉) · P. Liu · W. Yu
Shanghai Key Laboratory of Navigation and Location Based Services,
Shanghai Jiao Tong University, Shanghai, China
e-mail: xin.chen@sjtu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_8

multipath model accurately under these carriers [4]. Hannah uses a parabolic formula to simulate land-based multipath reflection models and validates it using lake-side data [5]. Yang Gao calculates the influence of satellite orbit on multipath fading theoretically [6]. Although a variety of multipath models have been proposed, there are still missing multipath statistical models based on the real navigation signal data, especially in complex environments such as urban canyons.

Because multipath signals are heavily dependent on the surrounding environment, it can be seen that the characteristics of multipath parameters are not similar in different types of environments. Therefore, it is necessary to establish corresponding statistical models in different types of scenes. Considering the multipath interference is most serious in the urban canyon environment, this paper will focus on the multipath characteristic model in this kind of scene. In order to make the model truly represent multipath characteristics, this paper extracts the characteristic parameters such as code phase delay and multipath energy attenuation in real navigation signal data, and using it to build model. Since the Beidou navigation satellite system contains GEO satellites, and it is significantly different from the motion characteristics of other types of satellites, GEO satellites and NGEO are distinguished and discussed separately.

2 Multipath Characteristic Parameters Extraction

2.1 Mathematical Model of Multipath

When the receiving antenna receives the line of sight (LOS) signal and the multipath signal at the same time, the received signal can be regarded as the combination of two signals. The formula is as follows:

$$\begin{aligned} s(t) &= AD(t-\tau_0)\cos(\omega_0 t+\theta_0) \\ &\quad + A\sum_{k=1}^{N}\alpha_k D(t-\tau_0-\tau_k)\cos[\omega_0 t+\theta_0+\Phi_k(t)] \end{aligned} \tag{1}$$

where A and τ_0 is the signal power and propagation time of LOS respectively. α_k is the power attenuation proportionality coefficient, and τ_k is the kth multipath time delay, and $\Phi_k(t)$ is the carrier phase difference.

It can be seen that α_k, τ_k and $\Phi_k(t)$ are the main parameters in multipath signal. Furthermore, $d\Phi_k(t)/dt$ is the multipath Doppler difference compared with LOS, which is also called multipath fading frequency [7]. In this paper, the statistical models of multipath signals are analyzed based on the above mentioned parameters.

2.2 Multipath Parameter Estimation

In this paper, the multipath characteristic parameters are obtained by processing the intermediate frequency navigation signal data in real environment. In order to get multipath characteristic parameters, this paper uses CADLL algorithm [8, 9, 10] in the tracking loop to track all the signals, then the parameters of all the signal can be estimated. The block structure of CADLL algorithm is as follows (Fig. 1):

In order to verify the correctness of the multipath parameters extracted by CADLL algorithm, Spirent GSS8000 navigation signal simulator is used. The configuration parameters of the simulator are shown in Table 1.

Figure 2 shows the detection result by CADLL algorithm, and each of the sub-figure represents multipath time delay τ_k, multipath power attenuation α_k, multipath carrier phase difference $\Phi_k(t)$, and spread code correlation waveform respectively. It can be seen that two multipath signals have been detected, and all the parameters are estimated in real time. In Fig. 2, the red line represents multipath 1 and the green line represents multipath 2, then the estimated error can be given as follows:

Table 2 shows that estimated error of the multipath characteristic parameters is within a reasonable range, so the statistical model of multipath based on the estimated parameter is effective. In addition, it can be seen that multipath interference distort the spread code correlation waveform. When the multipath signal is mitigated, the correlation peak returns to the triangular wave shape.

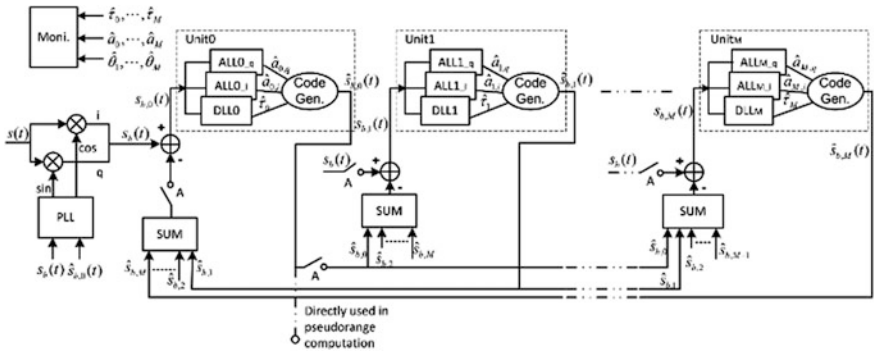


Fig. 1 Block structure of CADLL

Table 1 Configuration parameters of GNSS simulator

Constellation	GPS L1CA/PRN2	
Multipath number	Multipath 1	Multipath 2
Multipath delay (chip)	0.3	1.3
Power attenuation (dB)	5	15
Fading frequency (Hz)	0.1	0.3

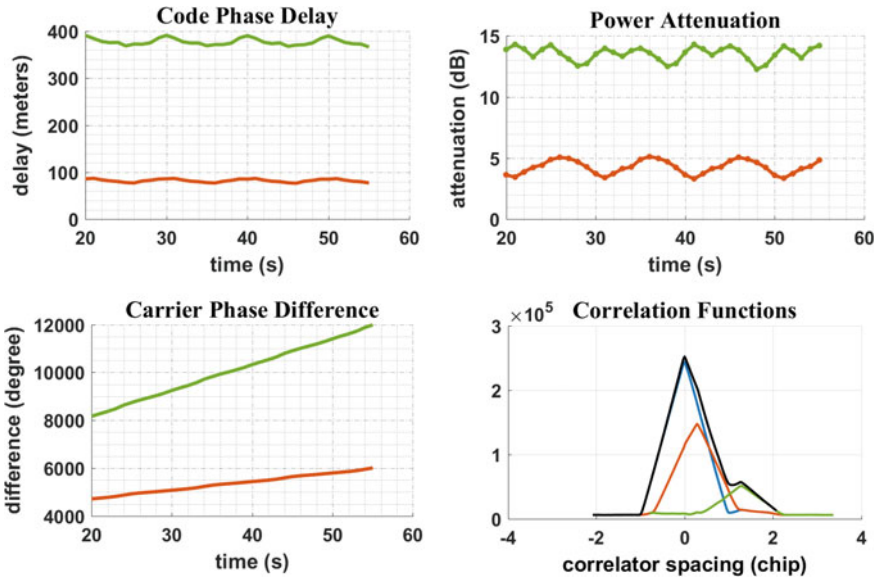


Fig. 2 Multipath characteristic parameters

Table 2 Estimated error of CADLL algorithm

Multipath number	Multipath 1	Error	Multipath 2	Error
Multipath delay (m)	82.3	5.6	377.4	3.5
Power attenuation (dB)	4.3	0.7	13.5	1.5
Fading frequency (Hz)	0.1002	0.002	0.3003	0.003

The main work of this paper is analyzing the characteristic parameters of the multipath signal in a large number of collecting data, and building the statistical model of these parameters.

3 Modeling Environment Description

3.1 Signal Collecting Environment

In order to establish the statistical characteristics model of multipath signal in urban canyon environment, this paper chooses the Lujiazui area in Shanghai as the real signal collection scene. Figure 3 shows all the collecting spots in google earth, and many skyscraper can be seen from the map. This paper randomly selects 10 collection site in the whole area with approximate uniform distribution, and each collection site is given by the yellow mark of 1–10.

Fig. 3 Site of navigation signal collecting (Color figure online)

Table 3 Configuration parameters of signal sampling machine

Location	10 spots in Lujiazui, Shanghai		
Time length	40 min per spot		
Antenna carrier	Van/static		
Constellation	BDS B1I and GPS L1CA		
Sampling frequency	62 MHz	Bit wide	8 bit
Signal bandwidth	50 MHz	Mode	Complex signal

Configuration parameters of the signal sampling machine are shown in Table 3.

Most parking place is the square before buildings for convenience, and a photo is shown in Fig. 4.

3.2 Model Description

Four kinds of statistical models are established based on the multipath characteristic parameters extracted from the real signal data. The models are described as follows:

- Multipath time delay probability density model: it is built by parameter τ_k. This model focuses on the probability density distribution of multipath time delay, and the relationship between multipath time delay distribution and satellite elevation.

Fig. 4 Photo of signal collection site

- Multipath power attenuation model: it is built by parameter α_k. This model focuses on the relationship between power attenuation and multipath time delay, and gives the probability distribution of multipath power attenuation probability in similar delay intervals.
- Multipath fading frequency probability density model: it is built by parameter $d\Phi_k(t)/dt$. When there is Doppler difference between the multipath signal and the direct path signal, the observation error of GNSS receiver caused by multipath signal is periodic, and this frequency is called multipath fading frequency.
- Multipath occurring probability model: This model gives the proportion of multipath occurring in entire experiment time.

4 Multipath Statistical Model

4.1 Multipath Time Delay Model

Multipath code phase delay is the main factor that causes the positioning error, and delay within one chip will seriously reduce the code phase measurement accuracy. Therefore, study of multipath time delay characteristic is very significant.

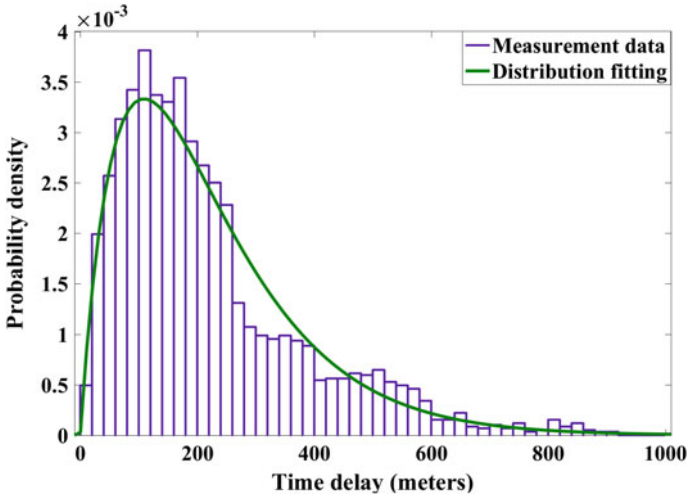


Fig. 5 Multipath time delay probability density model of NGEO satellite (Color figure online)

Figure 5 gives multipath time delay probability density model based on the experiment data. It can be seen that the general trend of the multipath delay probability distribution increases first and then gradually decreases, and the delay distribution in 100–150 m accounts for the largest proportion.

According to the statistical measurement data, Gamma function is used to fit the measurement data. The green line in Fig. 5 is the fitted curve, and its formula is as follows:

$$P_{Gamma}(\tau_n) = \frac{c}{b_\tau^a \Gamma(a)} \tau_n^{a-1} e^{-\frac{\tau_n}{b_\tau}} \tag{2}$$

Figure 6 shows the multipath time delay probability density distribution of GEO satellite. However, because the number of GEO satellites is small and the location is easy to be blocked, the measurement data is not enough to obtain the fitted distribution model. But the overall trend of GEO multipath time delay probability distribution is similar to NGEO.

4.2 Multipath Power Attenuation Model

Figure 7 gives the scatter distribution plot of the relationship between multipath power attenuation and time delay. As can be seen from the figure, with the increase of multipath delay time, the overall trend of the power attenuation is gradually increasing. Thus, liner function is used to fit the measurement, and the formula is as follows:

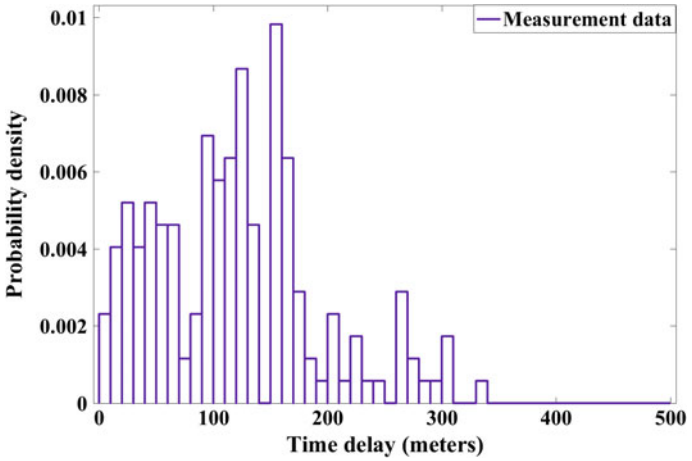


Fig. 6 Multipath time delay probability density model of GEO satellite

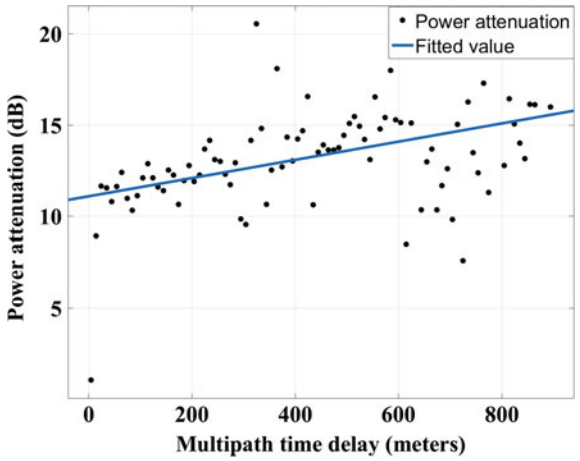


Fig. 7 Multipath power attenuation distribution

$$f(x) = p1 * x + p2 \tag{3}$$

4.3 Multipath Occurring Probability Model

In order to describe the overall effect of multipath interference, this paper further explores the occurring probability that antenna receives multipath signal. The result is shown in Fig. 8. In the figure, the occurring probability of multipath signal under different satellite elevation angles is given. The black point is the probability of

NGEO satellite, and the red fork indicates the probability of GEO satellite. In general, the multipath occurring probability is higher when the satellite elevation is lower. But Fig. 8 shows that the occurring probability decreases when satellite elevation is lower than 20°. Because multipath signal will be largely attenuated by trees or other shelter when the satellite elevation is lower than 20°, then the algorithm is not able to detect the multipath signal. The red line in Fig. 8 is the Gauss fitted curve of NGEO satellite multipath, and it can be seen that GEO satellite characteristic is similar as NGEO satellite. The formula is as follows:

$$P(\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\theta-\mu)^2}{2\sigma^2}} \tag{4}$$

4.4 Multipath Fading Frequency Model

Figures 9 and 10 show the multipath fading frequency probability density distribution of NGEO satellite and GEO satellite. It can be seen from the figure that the multipath fading frequency of NGEO satellite changes within the range of ±0.3 Hz, while the GEO satellite is within ±0.003 Hz. Although the receiver is statistic, the Doppler frequency between LOS signal and multipath signal is also different because satellite is moving. In addition, the multipath fading frequency will be large when the reflected object is moving.

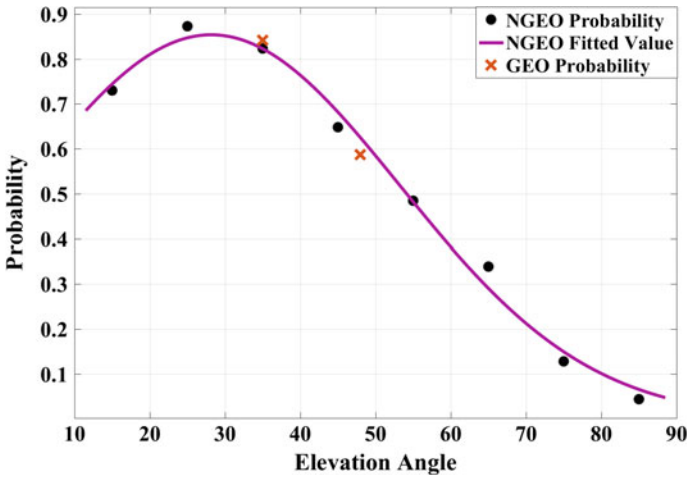


Fig. 8 Multipath occurring probability model

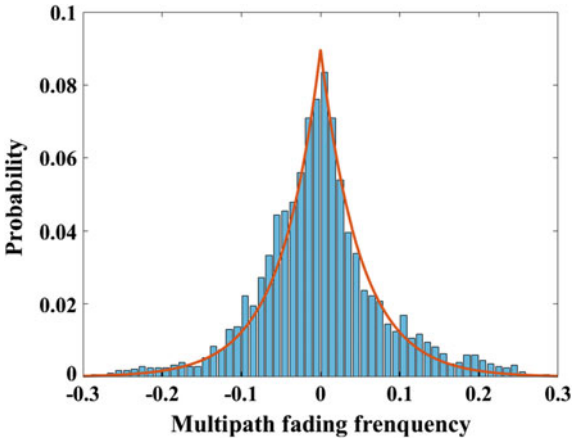


Fig. 9 Multipath fading frequency of NGEO satellite

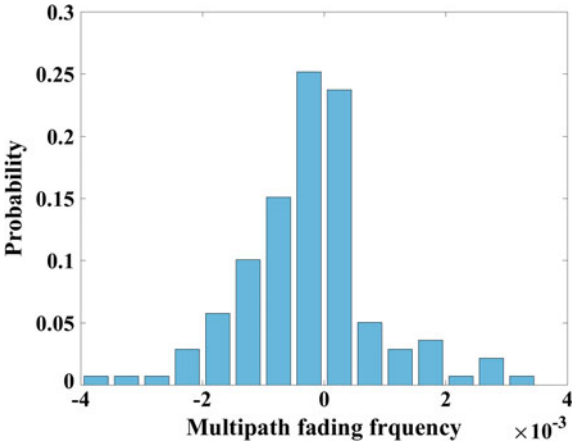


Fig. 10 Multipath fading frequency of GEO satellite

It can be seen from Fig. 9 that the probability density distribution of multipath fading is basically symmetrical, and the formula is as follows:

$$P(f) = \begin{cases} a * \exp(b * f) & f > 0 \\ a * \exp(-b * f) & f < 0 \end{cases} \tag{5}$$

Figure 10 shows that the multipath fading frequency distribution trend of GEO satellite is similar as NGEO satellite, but the value range is much smaller.

5 Conclusion

This paper extracts the characteristic parameters of multipath signal in urban canyon, and multipath delay model, multipath power attenuation model, multipath fading frequency model and multipath occurrence probability model are established by using these parameters. It reveals the feature of multipath in the urban canyon, and can help to improve multipath mitigation algorithm.

Acknowledgements This work is supported partly by grants from the National Natural Science Foundation of China (No. 61304225, No. 61401501) and Shanghai Science and Technology Committee program (No. 16DZ0504600).

References

1. Irsigler M, Eissfeller B (2003) Comparison of multipath mitigation techniques with consideration of future signal structures. In: Proceedings of ION-GPS
2. Lippincott WL, Milligan T, Igli D (1996) Method for calculating multipath environment and impact on GPS receiver solution accuracy, ION NTM 1996, Santa Monica, CA, 1996, pp 707–722
3. Model parameters for an urban environment for the physical-statistical wideband LMSS model in recommendation ITU-R P.681–7. International Telecommunication Union ITU, Geneva, Switzerland, Report P.2145, Oct 2009
4. Weiss JP, Axelrad P, Anderson S (2007) A gnss code multipath model for semi-urban, aircraft, and ship environments. Navigation 54(4):293–307
5. Hannah BM (2001) Modeling and simulation of GPS multipath propagation. Ph.D. dissertation, Queensland University of Technology, Mar 2001
6. Gao Y, Yao Z, Cui X et al (2014) Analysing the orbit influence on multipath fading in global navigation satellite systems. IET Radar Sonar Navig 8(1):65–70
7. Nee RDJ (1992) Multipath effects on GPS code phase measurements. Navigation 39(2): 177–190
8. Chen X, Dovis F, Pini M, Mulassano P (2011) Turbo architecture for multipath mitigation in global navigation satellite system receivers. J IET Radar Sonar Navig 5:517–527
9. Chen X, Dovis F, Peng S, Morton Y (2013) Comparative studies of GPS multipath mitigation methods performance. IEEE Trans Aerosp Electron Syst 49(3)
10. Chen X, Morton Y, Dovis.F (2013) A computationally efficient iterative MLE for GPS AOA estimation. IEEE Trans Aerosp Electron Syst 49(4)

Part II
Satellite Navigation Augmentation Technology

An Efficient Algorithm for Determining the Correspondence Between DFREI and σ_{DFRE} for a Dual-Frequency Multi-constellation Satellite-Based Augmentation System

Jie Chen, Zhigang Huang, Rui Li and Weiguang Gao

Abstract Global navigation satellite systems (GNSSs) will broadcast two open signals for civil aviation within the next ten years. Satellite-based augmentation system (SBAS) providers are currently developing the dual-frequency multi-constellation (DFMC) SBAS interface control document (ICD) to take advantage of the four GNSSs and two signals. In order to ensure the interoperability, a SBAS provider is allowed to determine the correspondence between DFREI and σ_{DFRE} based on the performance of its service within the frame of the ICD. Considering that the search space of the correspondence determination is too larger, an algorithm is proposed, which reduces the search space by decorrelating DFREIs. The evaluation based on real data shows that the effect of quantitation error resulted from the correspondence determined by the algorithm to SBAS service performance is neglectable, indicating that the algorithm is efficient. The algorithm could be of value for SBAS providers to design the correspondence.

Keywords Dual-frequency · Multi-constellation · Satellite-based augmentation system · DFREI · Correspondence

1 Introduction

In order to take advantage of the Galileo system and the Beidou system to be of full operational capability, and the two open signals centred around 1575.42 and 1176.45 MHz for civil aviation, the satellite-based augmentation system (SBAS) interoperability working group was formed by the SBAS providers to develop the dual-frequency multi-constellation (DFMC) SBAS interface control document (ICD) [1, 2]. The ICD defines the augmentation message types and describes the

J. Chen · Z. Huang · R. Li (✉) · W. Gao
Beihang University, Beijing, China
e-mail: lee_ruin@263.net

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_9

usage of the parameters in the augmentation messages. Because of the limited data rate of the SBAS signal and the need to augment 92 satellites simultaneously, DFMC SBAS does not broadcast equivalent terms for fast pseudorange corrections in legacy L1-only SBAS. Thus, there are only four components in the user range error (URE) variance bounding satellite clock-ephemeris (SCE) error after applying SCE corrections, denoted as σ^2_{DFC}, as follows:

$$\sigma^2_{DFC} = (\sigma_{DFRE} \cdot \delta_{DFRE})^2 + \varepsilon^2_{corr} + \varepsilon^2_{er} \tag{1}$$

where σ_{DFRE} is converted from the parameter DFREI in the message type (MT) 32, MT 34, MT 35 or MT 36 with the correspondence broadcast though the MT 37; δ_{DFRE} is a spatial parameter computed as:

$$\begin{aligned} &\delta_{DFRE} = \sqrt{I^T \cdot R^T \cdot R \cdot I} + \varepsilon_C \\ &R = 2^{ScaleExponent-5} \cdot \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{14} \\ 0 & E_{22} & E_{23} & E_{24} \\ 0 & 0 & E_{33} & E_{34} \\ 0 & 0 & 0 & E_{44} \end{bmatrix} \equiv S \cdot E \\ &\varepsilon_C = C_{covarinace} \times 2^{ScaleExponent-5} \end{aligned} \tag{2}$$

with *ScaleExponent* and E_{ij} being parameters in the MT 32, $C_{covarinace}$ being a parameter in the MT 37, I being the four-dimensional observation vector, and ε_C being used to compensating for the errors introduced by quantization; ε_{corr} is a temporal degradation parameter; ε_{er} is computed based on the approach type.

In order to simplify the description, ε_C is set to 0, which is the same value broadcast by WAAS. Thus, $\sigma_{DFRE} \cdot \delta_{DFRE}$ is computed as:

$$\sigma_{DFRE} \cdot \delta_{DFRE} = \sigma_{DFRE} \sqrt{I^T \cdot R^T \cdot R \cdot I} = \sqrt{I^T \cdot P_{brd} \cdot I} \tag{3}$$

where $P_{brd} = \sigma^2_{DFRE} \cdot R^T \cdot R$.

The master station of a SBAS decomposes the matrix P_{SIS}, which is the covariance matrix of SCE errors after applying SCE corrections, into σ_{DFRE}, *ScaleExponent* and E_{ij}. Then, σ_{DFRE} is converted to DFREI, and *ScaleExponent* and E_{ij} are quantized, so that these parameters can be broadcast to SBAS users. The MT 32, containing DFREI, *ScaleExponent* and E_{ij}, has a maximum update interval of 120 s. During the update interval of the MT 32, MT 34–36, whose maximum update interval are 6 s, broadcast DFREI to meet the alert time requirement of precision approaching operations. SBAS users receive these parameters to compute P_{brd}. The difference between P_{SIS} and P_{brd} is the quantization error.

A DFMC SBAS provider is allowed to customize the correspondence between DFREI and σ_{DFRE} based on the performance of its service within the frame of the ICD. This ensures the interoperability among SBASs with variant service performance. For instance, WAAS has a lot monitoring stations distributed in a wide area,

supporting LPV-200 operations with GPS, so that WAAS can choose a small scale factor for the correspondence. Meanwhile, a SBAS like MSAS with serval stations distributed in a limited area may choose a large scale factor. In addition, a SBAS augmenting a new GNSS whose satellite is less robust than a GPS satellite may choose a large scale factor so that the instability can be protected against. This is different from the L1-only SBAS, in which the correspondence between UDREI and σ_{UDRE} is unchangeable and is same among different SBASs, as listed in Table 1 [3]. Note that UDREI and σ_{UDRE} in L1-only SBAS are equivalent terms for DFREI and σ_{DFRE} in DFMC SBAS. Therefore, the problem that how to determine the correspondence so that the quantization error can be minimized comes up to DFMC SBAS providers, which L1-only SBAS providers do not need to face.

In the draft DFMC SBAS ICD, MT 37 is described as a critical message which is not planned to be dynamically changed [1]. Thus, the parameters in MT 37, especially the correspondence parameters, should be determined by SBAS providers in advance. According to the parameters description listed in Table 2, there are

Table 1 Correspondence between UDRE and σ_{UDRE}

UDREI	σ_{UDRE} [m]	UDREI	σ_{UDRE} [m]	UDREI	σ_{UDRE} [m]
0	0.2280	5	0.9119	10	2.2796
1	0.3040	6	1.1398	11	4.5593
2	0.3800	7	1.3678	12	15.1976
3	0.5320	8	1.5958	13	45.5927
4	0.6840	9	1.8237		

Table 2 σ_{DFRE} parameters in MT 37

Parameter name	Length [bit]	Scale factor [m]	Range [m]	
			Min	Max
σ_{DFRE}:DFREI = 0	4	0.0625	0.125	1.0625
σ_{DFRE}:DFREI = 1	4	0.125	0.25	2.125
σ_{DFRE}:DFREI = 2	4	0.125	0.375	2.25
σ_{DFRE}:DFREI = 3	4	0.125	0.5	2.375
σ_{DFRE}:DFREI = 4	4	0.125	0.625	2.5
σ_{DFRE}:DFREI = 5	4	0.25	0.75	4.5
σ_{DFRE}:DFREI = 6	4	0.25	1	4.75
σ_{DFRE}:DFREI = 7	4	0.25	1.25	5
σ_{DFRE}:DFREI = 8	4	0.25	1.5	5.25
σ_{DFRE}:DFREI = 9	4	0.25	1.75	5.5
σ_{DFRE}:DFREI = 10	4	0.5	2	9.5
σ_{DFRE}:DFREI = 11	4	0.5	2.5	10
σ_{DFRE}:DFREI = 12	4	1	3	18
σ_{DFRE}:DFREI = 13	4	3	4	49
σ_{DFRE}:DFREI = 14	2	30	10	100

$(2^4)^{14} \times 2^2 \approx 2.88 \times 10^{17}$ combinations of the correspondence if all 15 DFREIs are used. A problem like this is always solved by a brute-force search. However, the combinatorial explosion is highly likely to take place with so many candidate solutions. The key to avoid combinatorial explosion is introducing external information specific to the problem to reduce the search space, so that the brute-force search can be sped up [4–6].

Although a draft version of the DFMC SBAS ICD has been approved by SBAS providers, a public version has not been released yet. There is no public research about how to determine the correspondence. Based on the fact that the SBAS service performance is highly relied on the monitoring geometric dilution of precision (MGDOP) defined in (4), an algorithm using MGDOP to reduce the search space is proposed. This algorithm provides an efficient way for SBAS providers to determine the correspondence based on the its service performance under the constraint listed in Table 2.

2 Methodology

The proposed algorithm contains two parts. The first part decorrelates all DFREIs, so that the search space is cut into independent subspaces. The second part finds the optimum correspondence of each DFREI using a brute-force search. The criterion is that a correspondence candidate is optimum if the quantization error of it is minimum among those of other candidates.

2.1 Decorrelate DFREIs

There are 2.88×10^{17} candidates in the whole search space as mentioned above, which is infeasible to find the optimum one. After decorrelating all 15 DFREIs, the search space is cut into independent subspaces, so that there are only 16 candidates in each subspace when DFREI is smaller than 14 and only 4 candidates in the subspace when DFREI is 14. That means there are only 228 candidates totally if all 15 DFREIs are used. Thus, the optimum correspondence can be determined within a short time. The problem is how to decorrelate DFREIs.

DFREI describes the URE resulting from SCE errors after applying SCE corrections. The accuracy of SCE corrections depends highly on the configuration of SBAS monitoring stations. Given that all monitoring stations have a same tracking accuracy, MGDOP is the best indicator of tracking accuracy. The MGDOP of a satellite viewed from SBAS tracking stations is given by:

$$MGDOP = \sqrt{D_{1,1} + D_{2,2} + D_{3,3} + D_{4,4}}$$
$$D = \left(H^T H\right)^{-1} H_i = [\cos(Ele_i)\sin(Azi_i)\ \cos(Ele_i)\cos(Azi_i)\ \sin(Ele_i)\ -1] \quad (4)$$

where $D_{i,j}$ is the (i, j)th element of the matrix D; H_i is the i-th row of the matrix H and it is the observation vector from the ith monitoring station to the satellite; *Ele* and *Azi* are the elevation angle and azimuth angle of the satellite. Figure 1 shows that the MGDOP is suitable for decorrelation because its trend is similar to the trend of user differential range error indicator (UDREI) broadcast by WAAS.

The samples for determining the correspondence are classified into subsets by MGDOP. Each subset corresponds to a DFREI value. The mapping function from MGDOP to DFREI is designed and given by:

$$DFREI(MGDOP) = \left[\frac{\ln(MGDOP) - \ln(MGDOP_{\min})}{\ln(MGDOP_{\max}) - \ln(MGDOP_{\min})} \times (DFREI_{\max} - DFREI_{\min})\right] + DFREI_{\min} \quad (5)$$

where $MGDOP_{\min}$ and $MGDOP_{\max}$ are the minimum and the maximum of MGDOPs of a satellite; $DFREI_{\min}$ and $DFREI_{\max}$ are the minimum and the maximum of DFREI the SBAS plans to broadcast; $[x]$ rounds the number x to the nearest integer. Considering that the ranges of MGDOP of GNSS satellites are similar, $MGDOP_{\min}$ and $MGDOP_{\max}$ are set to be the quantiles of all satellites' MGDOPs with the probabilities 1 and 99%, respectively, for simplicity.

The problem of the mapping function above is that the computation of MGDOP needs at least 4 stations tracking the satellite. Thus, if a SBAS can provides integrity information with less than 4 stations, Eq. (5) cannot classify all samples. The MGDOP is the combination of the number of tracking stations and the geometric relation between tracking stations and the satellite. It is hard to describe the geometric relation quantitatively when there are less than 4 tracking stations, so that

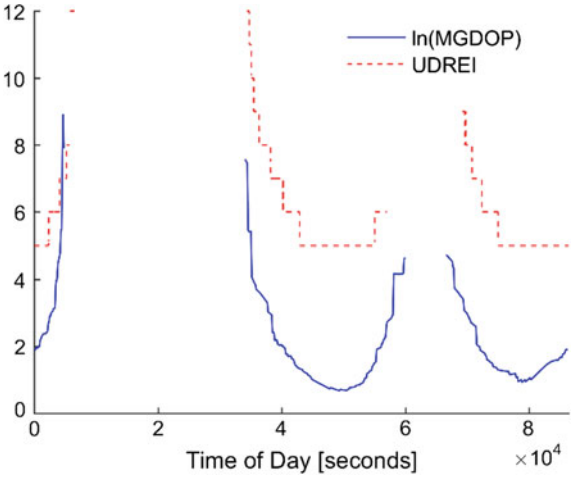


Fig. 1 MGDOP and UDREI of GPS PRN 1 satellite on May 21, 2015

the number of tracking stations is chosen to classify those samples. The amended mapping function is written as:

$$DFREI(MGDOP,n)= \begin{cases} DFREI_{\max}+n_{\min}-n, & n\le 3 \\ \frac{\ln(MGDOP)-\ln(MGDOP_{\min})}{\ln(MGDOP_{\max})-\ln(MGDOP_{\min})}\times(DFREI_{\max}-DFREI_{\min}+\min\{n_{\min}-4,0\}), & n>3 \end{cases} \tag{6}$$

where n is the number of tracking stations; $n_{\min}$ is the minimum number of tracking stations when the SBAS can still provide integrity information; $\min\{x,y\}$ returns the smaller value between x and y.

2.2 Determine the Correspondence

After all samples are classified into subsets using (6), the correspondence of a DFREI is determined with the corresponding subset. A brute-force search is used, which tries all values of σ_{DFRE} to find the one leads to the minimum quantization error. The quantization error QE is the maximum difference between $(\sigma_{DFRE}\cdot\delta_{DFRE})^2$ computed P_{brd} by and that computed by P_{SIS} within the service area, defined as follows:

$$QE=\max_{user\in servicearea}\{I^T P_{brd}(DFREI,ScaleExponent,E_{ij})I-I^T P_{SIS}I\} \tag{7}$$

where $P_{brd}(DFREI,ScaleExponent,E_{ij})$ is computed using (2) and (3).

The pseudocode of the brute-force search for a certain DFREI with a subset containing N samples is given below.

```
for k=1:N
  for n=0:n_max-1
    σ_DFRE(n) = σ_DFREI,min + n × ScaleFactor_DFREI ;
    compute R from P_SIS,k / σ²_DFRE(n) using Cholesky factorization;
    for ScaleExponent =0:7
      compute E_ij with R / 2^(ScaleExponent−5) ;
      compute QE ;
    end for;
    choose the ScaleExponent whose QE is the smallest, and add its QE to QE(n) ;
  end for;
end for;
find the n_s whose QE(n) is the smallest;
The correspondence is determined as σ_DFRE(n_s) = σ_DFREI,min + n_s × ScaleFactor_DFREI .
```

The parameter n_{max} in the pseudocode is the candidate number of the DFREI, which is 16 if the DFREI is smaller than 14 and is 4 if the DFREI is equal to 4. The parameters $\sigma_{DFREI,\min}$ and $ScaleFactor_{DFREI}$ are listed in Table 2.

3 Application and Evaluation

The data used in performance evaluation include the GPS broadcast message and precise ephemeris provided by the National Geospatial-intelligence Agency (NGA), WAAS broadcast message and WAAS stations' antenna L1 phase center positions provided by the National Satellite Test Bed, the GPS ultra-rapid ephemeris provided by the IGS, and measurements made at 1 Hz by WAAS reference stations and provided by the National Geodetic Survey. The GPS L5 service has not officially been announced to operate at full capability, so that measurements of the GPS L2 signal were used instead. The first part uses 7-day data to determine the correspondence as a SBAS provider. The second part uses anther 7-day data to find out the effect of the correspondence determine in the first part in user position domain.

3.1 Determine the Correspondence

The correspondence is determined with the SCE covariance matrix P_{SIS} computed by the AK method [7]. The mask angle, $n_{\min}$, $DFREI_{\min}$ and $DFREI_{\max}$ are set to be 5°, 1, 0 and 14, respectively. The mapping result using (6) on May 21, 2015 is showed in Fig. 2.

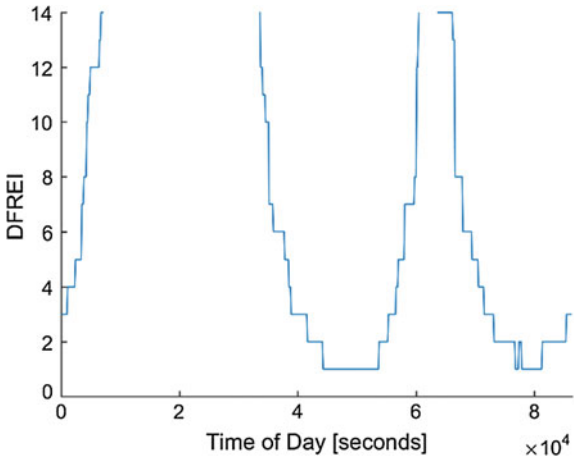


Fig. 2 Mapping result of GPS PRN 1 satellite on May 21, 2015

Table 3 Correspondence between DFREI and σDFRE

DFREI	σ_{DFRE} [m]	DFREI	σ_{DFRE} [m]	DFREI	σ_{DFRE} [m]
0	0.125	5	0.75	10	2
1	0.25	6	1	11	2.5
2	0.375	7	1.25	12	3
3	0.5	8	1.5	13	4
4	0.625	9	1.75	14	10

Based on the data from May 21 to 27, the correspondence is determined as listed in Table 3. It is noteworthy that the σ_{DFRE} of each DFREI is the minimum in its range.

3.2 Evaluate the Effect in User Position Domain

With the correspondence listed in Table 3, the SCE integrity parameters computed by the AK method are quantized. Based on the measurements of WAAS stations, navigation solutions are computed with the unquantized matrix P_{SIS} and quantized matrix P_{brd}, respectively. With the navigation solutions, the effect of quantization error introduced by the correspondence is evaluated in user position domain. The results presented below are based on data obtained from May 28 to June 3, 2015.

3.3 Position Error

Table 4 shows that there are 24 stations whose statistical results about position error before and after quantization are different. The differences in terms of the quantile with the probability 95% and the root mean square (RMS) are both less than 2 mm.

3.4 Availability of CAT-I

A DFMC SBAS is expected to support CAT-I operations. The most critical requirement is that the vertical alert limit ranges from 10 to 15 m depending on the need of an airport. The vertical alert limit is set to be 15 m in this evaluation. The availabilities of CAT-I in WAAS stations, which is the portion of time the vertical protection level is smaller than the vertical alert limit, are analyzed. There are 5 stations whose availabilities increase about 0.01%, and there are 6 stations, 1 station and 2 stations whose availabilities decrease about 0.01, 0.02 and 0.04%, respectively, while 23 stations' availabilities are unchanged.

Table 4 Position error

No.	Position error before quantization [m]				Position error after quantization [m]			
	Horizontal		Vertical		Horizontal		Vertical	
	95% quantile	RMS	95% quantile	RMS	95% quantile	RMS	95% quantile	RMS
1	1.525	0.892	2.448	1.215	1.526	0.892	2.45	1.215
2	1.153	0.654	2.711	1.424	1.153	0.654	2.712	1.424
3	1.276	0.752	1.777	0.908	1.275	0.752	1.777	0.908
4	1.651	0.93	2.367	1.174	1.651	0.930	2.366	1.173
5	1.618	0.94	2.539	1.324	1.617	0.940	2.539	1.324
6	1.621	0.923	3.746	2.046	1.621	0.923	3.745	2.046
7	1.585	0.909	2.859	1.455	1.585	0.909	2.860	1.456
8	2.027	1.129	2.856	1.467	2.027	1.129	2.855	1.466
9	2.071	1.121	3.414	1.66	2.072	1.121	3.414	1.660
10	2.296	1.303	3.155	1.571	2.296	1.303	3.155	1.571
11	1.949	1.079	2.586	1.291	1.948	1.079	2.585	1.291
12	2.057	1.134	3.34	1.662	2.057	1.135	3.341	1.662
13	1.428	0.842	2.526	1.303	1.43	0.843	2.527	1.304
14	2.372	1.272	3.555	1.802	2.371	1.272	3.556	1.802
15	1.882	1.046	2.691	1.345	1.881	1.046	2.691	1.345
16	2.167	1.176	3.166	1.648	2.167	1.176	3.166	1.647
17	2.147	1.203	3.514	1.808	2.147	1.203	3.513	1.808
18	1.941	1.094	2.760	1.425	1.941	1.094	2.761	1.425
19	1.995	1.149	2.764	1.400	1.996	1.149	2.765	1.400
20	1.968	1.093	2.946	1.508	1.968	1.093	2.947	1.508
21	2.655	1.529	3.932	1.989	2.655	1.528	3.932	1.989
22	2.028	1.132	2.914	1.471	2.029	1.132	2.914	1.472
23	1.697	0.971	2.928	1.497	1.697	0.971	2.928	1.498
24	1.878	1.023	2.526	1.282	1.878	1.023	2.525	1.282

4 Conclusion

The correspondence between DFREI and σ_{DFRE} should be determined by a DFMC SBAS provider in advance. Considering that MGDOP is an indicator of augmentation performance, an algorithm based on decorrelating DFREIs with MGDOP is proposed to decrease search space so that the combinatorial explosion in determine the correspondence concurrently can be avoided. Using 7-day real data, the proposed algorithm determined a correspondence, as listed in Table 3, within a search space of only 228 candidate solutions instead of 2.88×10^{17} candidate solutions. Another 7-day real data is used to analyze the effect of quantization error introduced by the correspondence in user position domain. The result shows that there is little influence on the service performance, proofing that the proposed algorithm is efficient. The algorithm could be of value for SBAS providers to design the correspondence.

Acknowledgements This work is supported by the National Natural Science Foundation of China (41304024).

References

1. Interoperability Working Group (2016) SBAS L5 DFMC interface control document (Issue 1, Revision 3)
2. Walter T, Blanch J, Enge P (2012) L1/l5 SBAS MOPS to support multiple constellations. In: Proceedings of the 25th international technical meeting of the satellite division of the institute of navigation, Institute of Navigation, Nashville TN, 17–21 Sept, pp 1287–1297
3. RTCA Special Committee 159 (2016) Minimal operational performance standards for global positioning system/wide area augmentation system airborne equipments (RTCA DO-229E). RTCA, Washington, DC
4. Anantharaman T, Campbell MS, Hsu FH (1990) Singular extensions: adding selectivity to brute-force searching. Artif Intell 43(1):99–109
5. Morton AB, Mareels IMY (2000) An efficient brute-force solution to the network reconfiguration problem. IEEE Trans Power Deliv 15(3):996–1000
6. Coxson G, Russo J (2005) Efficient exhaustive search for optimal-peak-sidelobe binary codes. IEEE Trans Aerosp Electron Syst 41(1):302–308
7. Chen J, Huang ZG, Li R (2017) Computation of satellite clock–ephemeris corrections using a priori knowledge for satellite-based augmentation system. GPS Solutions 21(2):663–673

Analysis on Ionospheric Delay Variogram Modeling in China

Dun Liu, Xiao Yu, Liang Chen and Weimin Zhen

Abstract Ionospheric delay variogram is essential for Kriging based Space-Based Augmentation System (SBAS) ionospheric grid model. Variogram realizations are compared for various solar activity conditions and different latitudes for China and around areas. It shows that the ionospheric anomaly in low-latitude areas impacts the variogram seriously, causing a larger intercept at zero distance and a "bump" at the distance of about 1000 km. An alternative method is proposed for variogram modeling for China areas and parameters are given for variograms under different solar conditions. A variogram with fixed constants is further proposed for practical use in SBAS, and the parameters are given also.

Keywords Ionospheric · Variogram · Kriging method · Ionospheric grid model · SBAS

1 Introduction

SBAS is established to augment GNSS performance in local region. In SBAS ionospheric grid model is used to correct the ionospheric delay which is one of the main error sources degrading system performance. Threshold to bounding the residual error effectively is also estimated in grid model. In the new development of SBAS, kriging method is adopted to estimate the delay and its bounding [1–3, 9–11].

The essence of kriging method is to estimate the grid ionospheirc vertical delay (GIVD) by making use of the spatial correlation characteristics of ionospheric delays, which is actually depicted by variogram model (or the corresponding covariance model). Variogram is also used for estimation of ionospheric delay error bound, namely the grid ionospheric vertical error (GIVE) [3, 9–11].

D. Liu (✉) · X. Yu · L. Chen · W. Zhen
No. 22nd Research Institute, CETC, Qingdao, Shandong 266107, China
e-mail: Dun.L@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_10

In previous work, characteristics of ionospheric spatial correlation structure in China area has been studied, and comparison has been made with that of North American region. Results show significant difference would be caused by the ionospheric anomaly in low latitude area of China [5, 7, 8]. For variogram realization, ionospheric anomaly effects in south of China may also exist. Study has to be carried out in variogram modeling and its further applications in SBAS for China area.

In this paper, the method of ionospheric delay variogram construction has been analyzed. Different variograms under various solar activity conditions and in different latitude areas have been studied and compared for China area. Section 2 introduces the definition of variogram and its application in ionospheric grid model. Section 3 gives the method of developing ionospheric delay variogram. Section 4 analyzes the characteristics of the variogram in detail for China area and the proposed variogram model is also presented. Conclusion is given at last.

2 Ionospheric Delay Variogram and Its Application in SBAS

2.1 Ionospheric Delay Variogram

The kriging estimated ionospheric delay can be expressed as the sum of a trend item, a random field (process noise) item r and a measurement noise item m [1–3].

$$\mathrm{I}_{\mathrm{meas}}(x_k) = a_0 + a_1 x_k^{(east)} + a_2 x_k^{(North)} + r(x_k) + m(x_k) \tag{1}$$

Where $[a_0, a_1, a_2]$ are coefficients of the planar trend item. x_k is the coordinate of ionospheric pierce point (IPP) in local coordinate system centered at ionospheric grid point (IGP) $x_k^{(North)}, x_k^{(east)}$ are respectively coordinates in the north and east direction.

For each pair of IPPs x_i, x_j, ionospheric delay variogram is defined as $\gamma(x_i, x_j)$:

$$\gamma(x_i, x_j) = \gamma(|x_i - x_j|) = \frac{1}{2} E\left(\left(r(x_i) - r(x_j) \right)^2 \right) \tag{2}$$

Covariance function $Cov(x_i, x_j)$ can then be defined as:

$$Cov(r(x_i), r(x_j)) = E(r(x_i) r(x_j)) = C_0(|x_i - x_j|) \tag{3}$$

Relation between variagram $\gamma(x_i, x_j)$ and covariance function $Cov(x_i, x_j)$ is:

$$Cov(x_i, x_j) = \sigma_\infty^2 - \gamma(x_i, x_j) \tag{4}$$

In kriging method, ionospheric delay variogram is required to be a non-zero value at distance of zero, which is called the nugget effect. The nugget-effect here is equivalent to the white noise in delay estimation. The variogram increases linearly near the origin, and at distance of infinite, it should converge to a constant value, which is called the sill effect [3].

An admissible and widespread model for the variogram is the exponential model shown as Fig. 1, where $\sigma_0 = \sigma_{decorr}^{nom}, \sigma_\infty = \sigma_{decorr}^{total}$ are respectively the nugget and sill. The corresponding ionospheric delay variogram and covariance function is shown as below [9, 10].

$$\gamma(x_k, x_l) = \left(\sigma_{decorr}^{total}\right)^2 - \left[\left(\sigma_{decorr}^{total}\right)^2 - \left(\sigma_{decorr}^{nom}\right)^2\right] \exp\left(-\frac{D_{k,l}}{d_{decorr}}\right) \tag{5}$$

$$C(x_k, x_l) = \left[\left(\sigma_{decorr}^{total}\right)^2 - \left(\sigma_{decorr}^{nom}\right)^2\right] \exp\left(-\frac{D_{k,l}}{d_{decorr}}\right) \tag{6}$$

$$D_{k,l} = \sqrt{(x_k - x_l)^T (x_k - x_l)} \tag{7}$$

Where $\sigma_{decorr}^{total}, \sigma_{decorr}^{nom}$ are coefficients that need to be estimated. Constant d_{decorr} is the de-correlated distance, which is generally selected as 8000 km.

2.2 Ionospheric Grid Model Based on Kriging Method

The kriging estimate of the grid ionospheric vertical delay (GIVD) I_{est} is determined as [3, 9]:

$$I_{est}(x) = \sum_{k=1}^{n} \lambda_k \mathrm{I}_{\mathrm{meas}}(x_k) \tag{8}$$

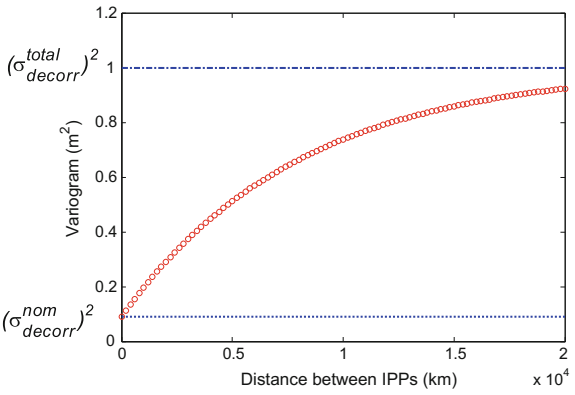


Fig. 1 Ionospheric delay variograms proposed in WAAS

where λ is weight matrix, I_{meas} are the vertical ionospheric delays at IPPs, n is the number of IPPs around the IGP for GIVD estimation.

$$\lambda = \left(W - WG\left(G^T WG\right)^{-1} G^T W\right) C(x, x_k) + WG\left(G^T WG\right)^{-1} X \tag{9}$$

$$W = \left(C(x_k, x_l) + M(x_k, x_l)\right)^{-1} \tag{10}$$

$$X = \begin{bmatrix} 1 & x^{(east)} & x^{(north)} \end{bmatrix}^T \tag{11}$$

where G is the coefficient matrix determined by the satellite and user's location. W is the weight matrix, C is the covariance matrix from ionospheric variogram model, and M is the measurement noise matrix.

GIVE, the bound for residual of delay estimation, is defined as follows [1–3]:

$$\sigma^2_{bound,IGP} = R^2_{irreg}\left[\lambda^T C(x_k, x_l)\lambda - 2\lambda^T C(x_k, x) + \left(\sigma^{total}_{decorr}\right)^2\right] + \lambda^T M \lambda \tag{12}$$

where R^2_{irreg} is the inflating factor.

As can be seen, the variogram is used to construct weighing matrix for ionospheric delay and its accuracy estimation. Moreover, it also depicts the uncertainty $\left(\sigma^{total}_{decorr}\right)$ of the background ionosphere in error bound estimation.

3 Method of Ionospheric Delay Variogram Construction

3.1 Modeling Data

GPS dual-frequency measurements are used in the study. The sites distribution covers the low and mid latitude areas around China (Fig. 2). The data period spans different solar activities (2003, 2005, 2009), and includes various status of ionosphere (calm and stormy).

3.2 Data Pre-processing

Ionospheric delays are extracted from GPS dual-frequency measurements. The following steps are taken to improve the accuracy of derived ionospheric delays [4].

(1) Carrier phase smoothed code pseudo-ranges are used to extract the ionospheric delays.
(2) Differential Code Bias (DCB) of GPS satellites released by CODE is used, so only the hardware bias for receiver is to be estimated.

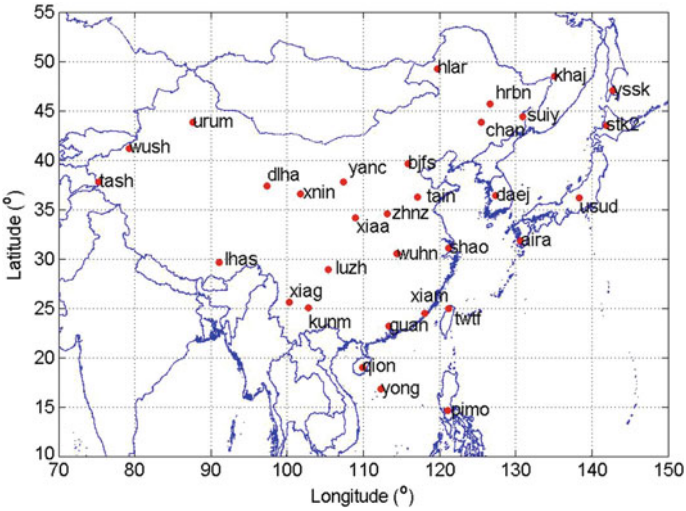


Fig. 2 GPS network around China areas

(3) Only data during nighttime is used in receiver hardware bias estimation to reduce the effects of larger variation of background ionosphere during daytime.
(4) Results from multiday are smoothed to improve the stability of receiver hardware bias estimation, especially for ionospheric disturbance period.

Comparison with results from CODE shows an accuracy of better than 0.2 ns could be achieved for the receiver hardware bias estimation in this way. As the smoothed pseudo-ranges could reach the accuracy of millimeter or even better, the resulting ionospheric delays will have the accuracy of about 0.2 ns.

3.3 Construction of Ionospheric Delay Variogram

With delays at IPPs, the variogram model can be constructed as below [3].

(1) For each pair of IPPs, we did a planar fit on the remaining measurement (up to a certain radius) with the middle point of the two IPPs as reference.
(2) At each epoch, underlying trend described by this planar fit is removed from the delays at the pair of IPPs, and then the difference of de-trended delays $I(x_i) - I(x_j)$ is calculated.
(3) Compute all the difference $I(x_i) - I(x_j)$ between pairs of de-trended measurements and the corresponding distance $|x_i - x_j|$ for all epochs.

(4) The experimental variograms defined below are estimated with the difference of de-trended delays. Because we do not have pairs of measurements separated by exactly d, we bin the pairs with the certain tolerance δ of 50 km [3].

$$\gamma_{\exp}(d) = \frac{1}{2}E\left(I(x_i) - I(x_j)\right)^2 \quad |x_i - x_j| \in \left[d - \frac{\delta}{2}, d + \frac{\delta}{2}\right] \tag{1.13}$$

(5) Fitting the experimental variograms with exponential model given in Eq. 5 and determining the σ_{decorr}^{total} and σ_{decorr}^{nom} for the variogram models.

4 Realization of Variogram for China Areas

4.1 Variogram Under Various Solar Activities

Figures 3, 4 and 5 shows the experimental variograms (blue line) for low (2009), modest (2005) and high solar activity (2003) respectively. Data from days of quiet ionospheric conditions are used in these variogram realizations. A period of three hours is used to bin the data for statistics in each day. As can be seen:

(1) At the distance of zero, experimental variograms have values of non-zero, which means the uncertainty caused by the background ionosphere. Yet compared with variograms for North America area, the intercept of experimental variograms and the range of its variation are much larger at the distance of zero [1, 3].

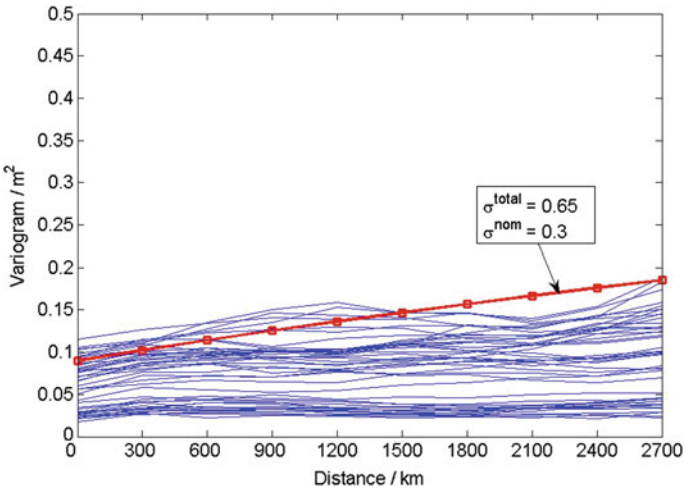


Fig. 3 Variograms for China areas (2009) (Color figure online)

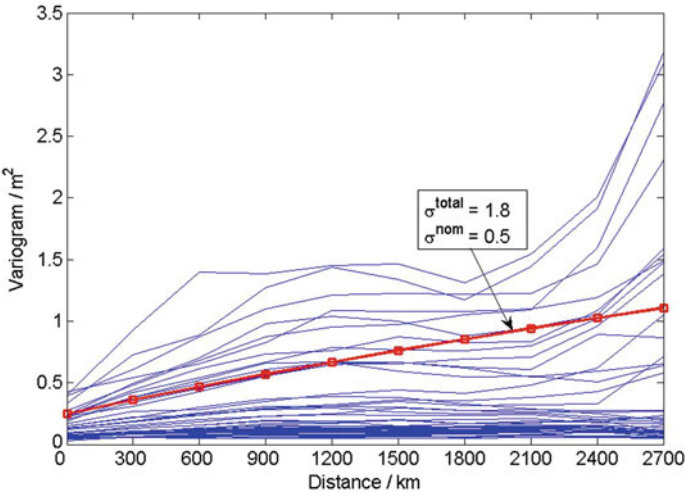


Fig. 4 Variograms for China areas (2005) (Color figure online)

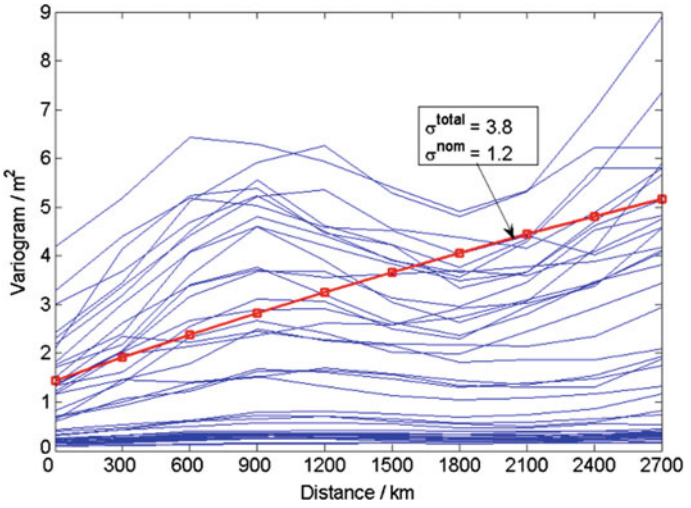


Fig. 5 Variograms for China areas (2003) (Color figure online)

Satellite signals may pass through the same IPP from different azimuths and elevations, leading to different ionospheric delays at different paths. Moreover, mapping function will also introduce error in vertical delays when converting from the slant ones. As a result, the difference of de-trended delays will not be zero even for neighboring IPPs.

This difference would be larger in areas where low latitude anomaly exists. Larger errors from excess ionospheric delays when signals pass through the anomaly region and mapping function based on the imperfect thin shell model destroyed by ionospheric anomaly contribute to this increasing difference.

(2) For years of modest and high solar activities, a "bump" at the distance of about 1000 km exists evidently in the ionospheric delay variogram. This phenomenon is not shown in the variogram for North America area.

Within approximately ±20° on both side of the magnetic equator lies the equatorial anomaly. It is the occurrence of a trough in the ionization in the F2 layer at the equator and crests at larger magnetic latitude depending solar activity. For China area, the south of Changjiang River (magnetic latitude of about 20°N) is located under the ionospheric anomaly, where the area along the River is at the northern edge of the anomaly, and the area around Guangzhou and Haikou is located under the crest of anomaly [12].

When part or all IPPs are located in ionospheric anomaly region, the anomaly structure will break down the assumption of the planar trend underlying the thin-shell ionospheric model, increasing the error in de-trended delays. The error will reach the largest especially when the two IPPs, with a distance of about 1000 km, are located respectively at the edge and the crest of the anomaly along the same longitude. In that case, most of the IPPs for grid model fitting are located in the northern area of the anomaly and the crest structure of the anomaly will have the maximum effect on delay estimation.

With the distance between two IPPs increasing, the difference of ionospheric delays will decrease even both IPPs are located in anomaly, as this two IPPs has a smaller geomagnetic latitude span.

4.2 Variogram in Different Latitude Region

To understand the anomaly effects on variogram more deeply, we divided the IPPs distribution into two parts: IPPs in low latitude area, and IPPs in middle latitude area. The latitude of 30°N is selected as the separator. Experimental variograms for these two regions are respectively estimated and results are shown in the following figure. As can be seen that,

(1) For the middle latitude area, linear variation is shown clearly in the experimental variogram. This means that in this region ionosphere has the perfect planar variation, and the fitting results of variogram is in good accordance with that of the North American [3, 9] (Fig. 6).
(2) For the low latitude area, the "bump" at distance of about 1000 km in the experimental variogram still exists. This means that the abnormal variation in experimental variogram of China area is caused by the ionospheric anomaly in the low latitude area (Fig. 7).

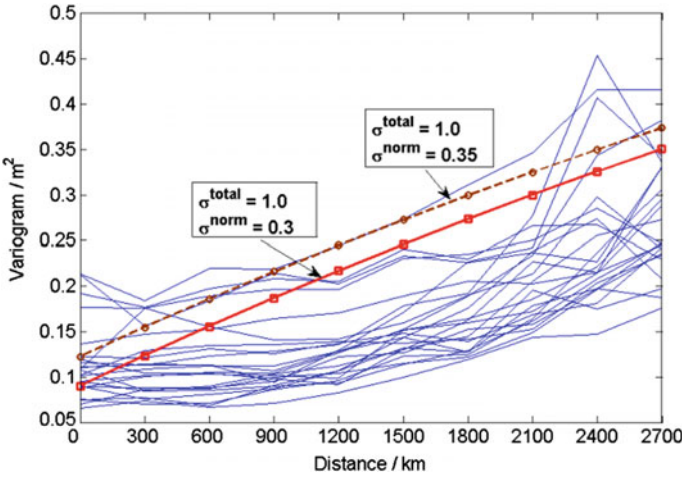


Fig. 6 Variograms for mid-latitude areas around China (North of 30°N, 2003)

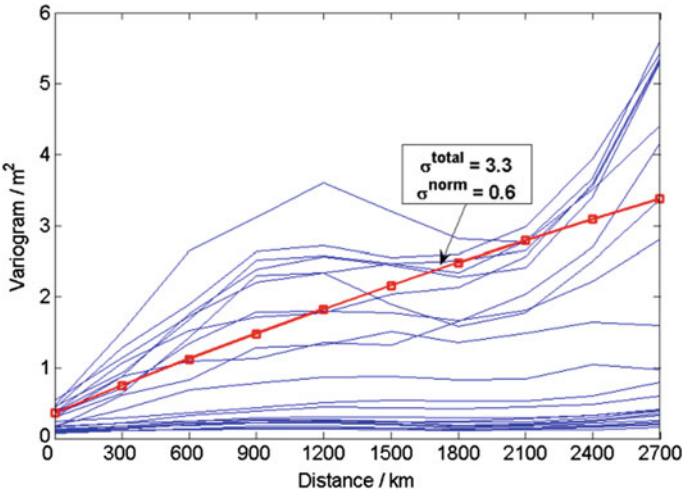


Fig. 7 Variograms for low latitude areas around China (South of 30°N, 2003)

4.3 Fitting of Experimental Variogram

With the analysis of ionopsheric anomaly effects, it can be seen that there are two features in the experimental variograms under modest and high solar activity.

(1) Larger difference between de-trending delays even for nearby IPPs.
(2) "Bump" in the experimental variogram for modest and high solar activity.

In variogram modeling, exponential model is selected to fitting the experimental variogram, in which the variogram increases linearly near the origin [1, 3]. Ionospheric anomaly effects make the modeling difficult as abnormally large values would be estimated for σ_{decorr}^{total} and σ_{decorr}^{nom} if the large error in the intercept at zero distance and the "bump" effects at 1000 km would be taken into account.

As can be seen from Sect. 2, the ionospheric delay variogram is mainly used to construct the spatial covariance function for weighting estimation in kriging method. So we could focus only on the regular change in experimental variograms, here is the linear variation with distance, and dismiss the change of "bump" variation.

At the same time, variogram fitting provides the estimation of the parameter σ_{decorr}^{total}, which describes the uncertainty of the background ionosphere. The σ_{decorr}^{total} will contribute to the estimation of threshold on grid ionospheric delay error. Since system availability is impacted adversely by an increasing error bound, smaller value for σ_{decorr}^{total} is preferred in variogram modeling.

In SBAS, spatial threat model is adopted to compensate the delay estimation error caused by disturbance. As the effects of ionoshperic anomaly on SBAS is the same as that of modest storms, it is reasonable to neglect the anomaly induced "bump" error in variogram modeling and make the errors absorbed by the spatial threat model.

Based on above consideration, we neglected "bump" part of the experimental variogram and only fit the linear trend variation in variograms in our variogram modeling efforts. Exponential function is still selected for modeling.

4.4 Realization of Variogram for China Area

Experimental variograms are fitted for various solar activities and the variogram models with different parameters are then established. In Figs. 3, 4 and 5, results of variogram models are shown as the red lines. The estimated parameters of σ_{decorr}^{total} and σ_{decorr}^{nom} are also given. In Table 1, typical values of parameter for variograms under various solar activities are given.

As can be seen, variogram models are quite distinct for different level of solar activity. Moreover, there's no multiple variation between different variogram models. In Reference [3], a method is recommended to build covariance model $C(d) = u_0^2 C_0(d)$ for different ionospheric condition, in which $C_0(d)$ is the basic covariance model for quiet conditions, and the "zooming" factor u_0^2 is derived based on the degree of disturbance in the ionosphere. Yet this way in establishing

Table 1 Typical parameters for variogram models

Solar activity	σ_{decorr}^{nom}	σ_{decorr}^{total}
Low (2009)	0.3	0.65
Modest (2005)	0.4–0.6	1.2–1.8
High (2003)	0.6–1.2	3–4

variogram models for various ionoshperic conditions is not applicable for China area. A possible resolution is to develop different ionospheric delay variogram models according to the solar activity.

However, there are still difficulties in the multi-variogram scenario mentioned above for practical use. As the spatial threat model is constructed based on variogram model, a new variogram model realization will require the corresponding spatial threat modal to be established [10, 11]. In addition, what principles are adopted to determine the solar activities and then the selection of a suitable variogram model are also the difficulty.

Sparks et al. pointed out that the kriging estimation based on varigram model with fixed constants could still has a better performance than the planar fit method [9, 10]. The same way can also be adopted in variogram modeling for SBAS in China. In practice, the varigram model with fixed constants is established based on the experimental variograms from middle latitude area for periods of various solar activities and that from low latitude area in periods of modest and low solar activities. The parameters for the final model is $\sigma_{decorr}^{nom} = 0.45$ and $\sigma_{decorr}^{total} = 1.25$. The effectiveness of this variogram modeling method has been tested in our work on ionospheric spatial threat model development which will be presented in another paper [6].

5 Conclusion

Variogram is essential for ionopsheric grid model in SBAS based on kriging method. Variogram is generally used to construct the weighing factors in the estimation of grid delay and its error bound. Different realization of variogram will result in distinct system performance.

Experimental variograms for different solar activities and different latitude areas are compared in this paper. Result shows that in the mid-latitude area of China variogram follow the linear change. But when considering the low latitude, its variation becomes more complicated with effects of the ionospheric anomaly. These effects are much more significant under modest and high solar activities. These characteristics make it difficult to modeling the ionospheric delay variogram for SBAS in China with the ways used in WAAS.

Method is recommended on ionospheric delay variogram modeling after analysis of typical features in experimental variograms caused by the ionoshperhic anomaly. In this method the anomaly caused “bump” effect is neglected in the experimental variograms and only the underlying linear variation is considered in the variogram model fitting. Any possible errors from effects of the ionospheric anomaly would be absorbed in the ionospheric spatial threat model. Ionospheric delay variograms are modeled for different solar activities and typical values of parameters in the model are given. An ionospheric delay variogram with fixed constants is recommended further for practical SBAS in China and the parameters are also given for this model realization.

References

1. Blanch J (2002) An ionosphere estimation algorithm for WAAS based on kriging. In: Proceedings of ION GPS 2002, Institute of Navigation, Portland, OR, September 24–27, pp 816–823
2. Blanch J, Walter T, Enge P (2002) Application of spatial statistics to ionosphere estimation for WAAS. In: Proceedings of ION NTM 2002, Institute of Navigation, San Diego, CA, January 28–30, pp 719–724
3. Blanch J. (2003). Using kriging to bound satellite ranging errors due to the ionosphere. Ph.D. thesis, Department of Aeronautics and Astronautics, Stanford University
4. Carrano CS, Groves KM (2006) The GPS segment of the AFRL-SCINDA global network and the challenges of real-time TEC estimation in the equatorial ionosphere. In: Proceedings of ION NTM 2006, Institute of Navigation, Monterey, CA, January 18–20, pp 1036–1047
5. Liu D, Chen L, Zhen WM (2013) Ionospheric spatial correlation analysis for China area. In: Proceedings of CSNC 2013, Wuhan China, May 15–17, 2013
6. Liu D, Feng J, Chen Li, Zhen WM (2017) A study on construction of ionospheric spatial threat model over China area. In: Proceedings of CSNC 2017, Shanghai China, May 15–17, 2017
7. Hansen A, Peteron E, Walter T, Enge P (2000) Correlation structure of ionospheric estimation and correction for WAAS. In: Proceedings of ION NTM 2000, Institute of Navigation, Anaheim, CA, January 26–28, pp 454–463
8. Hansen A, Peteron E, Walter T, Enge P (2000) Ionospheric correlation analysis for WAAS: quiet and stormy. In: Proceedings of ION GPS 2000, Institute of Navigation, Salt Lake City, UT, September 19–22, pp 634–642
9. Sparks L, Blanch J, Pandya N (2011) Estimating ionosperhic delay using kriging: 1. Methodology. Radio Sci 46, RS0D21. doi:10.1029/2011Rs004667
10. Sparks L, Blanch J, Pandya N (2011). Estimating ionosperhic delay using kriging: 2. Impact on satellite-based augmentation system availability. Radio Sci 46, RS0D22. doi:10.1029/2011Rs004781
11. Sparks L, Blanch J, Pandya N (2013) Kriging as a means of improving WAAS availability. In: Proceedings of ION GNSS 2013, Institute of Navigation, Portland, OR, September 21–24, pp 2013–2020
12. Xiong H (2000) Radio propagation. Publishing House of Electronics Industry, Beijing

Scintillation Modeling and Its Application in GNSS

Dun Liu, Xiao Yu, Jian Feng and Weimin Zhen

Abstract A method to analysis GNSS effects flexibly under various scintillation conditions is presented with a combination of irregularity strength modeling and signal propagating model. Irregularity strength model describes the spatial and temporal variation of strength of scintillation-producing irregularities. The model could be represented in terms of solar activity, magnetic latitudes, season and local time, and the coefficients are derived with measurements from China areas. Irregularity strength distribution, combined with propagating model, could produce the simulated time series of scintillation affected signals received by users. The combination also could give the scintillation index at user receivers under various geometry configurations, which could be further used to GNSS availability analysis under different scintillation conditions.

Keywords Ionospheric scintillation · Irregularity strength · Propagating model · GNSS

1 Introduction

Ionospheric scintillation impacts seriously the space based information system such as the GNSS, SAR etc. It is essential to develop models to simulate the scintillation affected signals to assess the potential degradation of system performance under scintillation [1–3].

The ionospheric scintillation model generally includes two parts: the irregularity models and the signal propagation model [4–7]. The ionospheric irregularity models, which depict the characteristics of the irregularity that caused the scintillation effects, include models on height, thickness, strength and spectral index of the irregularity [4, 6–9]. Among them, the strength model of irregularity is the most

D. Liu (✉) · X. Yu · J. Feng · W. Zhen
No. 22nd Research Institute, CETC, Qingdao, Shandong 266107, China
e-mail: Dun.L@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_11

important one. The signal propagation model is generally based on the phase screen theory, since the phase screen theory is one of the theoretical model to explain the scintillation phenomenon on one hand, and tests had also verified the good consistency between the theory analysis and the experimental results on the other hand [10, 11, 16].

There are two main ideas in the irregularity strength modeling. One is to develop model on the irregularity strength directly (actually *CsL* or *CkL*, the product of the strength of turbulence and the irregularity layer thickness), with the WBMOD (Wide-Band MODel) as the typical one [7–9]. Another is to model the local electron density perturbation. The typical models based on this method includes the GISM (Global Ionospheric Scintillation Model) in Europe, as well as the early version of the WBMOD [6].

As the irregularity strength can be retrieved from the ground observation, such as the GNSS scintillation measurements widely deployed, developing an irregularity strength model is more favorable in scintillation modeling efforts [4, 7, 8].

In the implementation of phase screen method, by using the split step algorithm and the Discrete Fourier Transformation (DFT) to solve the propagation equations, we can obtain the complex amplitude of signal reaching receiver antenna. With the assumption of "frozen field", the spatial variation of signal intensity caused by scintillation can be transformed to temporal variation of signal intensity, and then time series of signal intensity can be obtained. In addition, the scintillation index can also be obtained by theoretical calculation [11, 12]. Therefore, the phase screen method provides a flexible way in the analysis of the scintillation effect.

Efforts on scintillation modeling have been conducted in our previous work, here in this paper we focus on the analysis the scintillation effects on GNSS with these models. Section 2 introduces the irregularity strength model for low latitude areas of China. Section 3 introduces the signal propagation model based on the phase screen theory. Based on these models, the scintillation effect was analyzed for different cases in Sect. 4. Finally the conclusion is given.

2 Scintillation Related Irregularity Modeling

2.1 Irregularity Strength

The parameters of the scintillation related irregularity include the height, thickness, strength, drift velocity, axial ratio, spectral index, and so on [4, 7–9]. Among them, the strength of the irregularity is the most important one in scintillation modeling.

The strength of irregularity is generally expressed as *CsL* or *CkL*, in which the disturbance strength *Cs* is closely related to the structure constant that is used in classical turbulence studies, and *L* is the thickness of irregularity. Due to the difficulty in distinguishing the contribution of these two factors in scintillation effects, *CsL* is generally modeled in study [4, 7, 8]. *CkL* is the introduced as a new

parameter describing the irregularity strength in WBMOD [9]. The relationship between them is:

$$CkL = (1000/2\pi)^{p+1} CsL \tag{1}$$

As *CsL* or *CkL* can be retrieved from the scintillation measurement, an empirical model could be developed with the long term observation. In the estimation of *CsL* or *CkL*, auxiliary parameters needed in the modeling can be obtained from the location of the satellite and receiver, such as the geometry configuration, or from the empirical values, such as the spectral index p.

2.2 Irregularity Strength Model

The study shows that scintillation activity depends on the magnetic latitude, season, time after sunset, solar activity, and so on [4, 7, 8]. Therefore, irregularity modeling was generally been conducted with different functions fitting into these parameters. In Ref. [7], the irregularity strength is expressed as the functions of R, the sunspot number, τ, the angle between solar terminator and magnetic meridian at geomagnetic equator, t_e, the time after E-layer sunset, and λ_a, the apex latitude [11].

$$\sqrt{CsL} = C_e f_r(R) f_s(\tau) f_t(t_e) f_\lambda(\lambda_a) \tag{2}$$

The irregularity strength *CkL* is selected in our modeling effort instead, and the following model is developed with observation in China [4].

$$\sqrt{CkL} = C_E f_r(R) f_s(\tau) f_t(t_e) f_\lambda(\lambda_a) \tag{3}$$

where, the functions of $f_r(R) f_s(\tau) f_t(t_e) f_\lambda(\lambda_a)$, with the variables defined as before, are derived from observation.

(1) $f_r(R)$ **model**

The function of $A\sqrt{1 + C_r R}$ is adopted for $f_r(R)$. Coefficients of the function are estimated with a least-squares fit to the *CkL* data retrieved from observation. The result is $A = 16.48, C_r = 0.06$, [4].

(2) $f_s(\tau)$ **model**

The following function is used for $f_s(\tau)$, and the coefficients in the function were estimated with the scintillation data. The result obtained is $B = 1.11, \omega_\tau = 28.5$ [4].

$$f_s(\tau) = B \exp\left[-\left(\frac{\tau}{\omega_\tau}\right)^2\right] \tag{4}$$

(3) $f_t(t_e)$ **model**

The analysis shows that in the low latitude area of south China, ionospheric scintillations begin 1 h after sunset in the E layer, and can last up to 4 h, then its strength slowly decay [4]. Therefore, the following form is used for $f_t(t_e)$.

$$f_t(t_e) = \begin{cases} \exp\left[-\left(\frac{t_e-1}{0.5}\right)^2\right] & t_e \le 1 \\ 1 & 1 < t_e < 3 \\ \exp\left[-\left(\frac{t_e-3}{3.5}\right)^2\right] & 3 \le t_e \end{cases} \tag{5}$$

(4) $f_\lambda(\lambda_a)$ **model**

The magnetic latitude dependence of equatorial $\sqrt{CkL}$ was represented by a function that increases toward the north and south from a non-zeros value at the magnetic equator and which peaks on its way to an asymptote of zero at mid-latitudes [7, 8]. So the following form is estimated for $f_\lambda(\lambda_a)$.

$$f_\lambda(\lambda_a) = 0.5\left\{1 - erf\left[\frac{|\lambda_a| - 20}{3}\right]\right\} \tag{6}$$

where λ_a is the apex latitude in degrees, *erf* denotes the error function, $erf(x) = 2/\sqrt{\pi}\int_0^x e^{-t^2}dt$.

With the functions $f_r(R), f_s(\tau), f_t(t_e), f_\lambda(\lambda_a)$ substituted into Eq. 3, and a least-square fitting to the retrieved *CkL* data, the value for the constant C_E can then be readily estimated. The final model representation of the irregularity strength *CkL* as in Eq. 3 will be straightforward with all these functions and its coefficients available. In the study, the strength model of irregularity is established by using the scintillation observation data of Haikou area during 2003–2008 [4].

2.3 *Accuracy Assessment of Irregularity Strength Model*

The accuracy of the irregularity strength model was verified with the ionospheric scintillation data in Haikou and Guangzhou from GPS measurements. The internal and external consistency accuracy is tested for the model.

2.3.1 Internal Consistency Accuracy of Strength Model

The internal consistency accuracy of the model is analyzed with the observation data of Haikou area during 2003–2008. Figure 1 shows that the residual error of the model has a normal distribution, with a mean of 0, and a variance of 0.44.

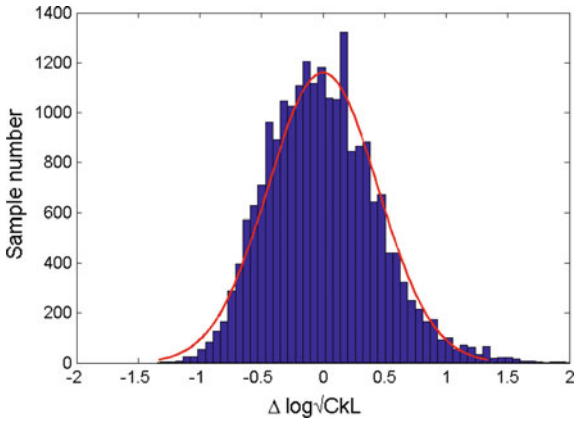


Fig. 1 Verification of the irregularity strength model with data of 2003–2008 from Haikou

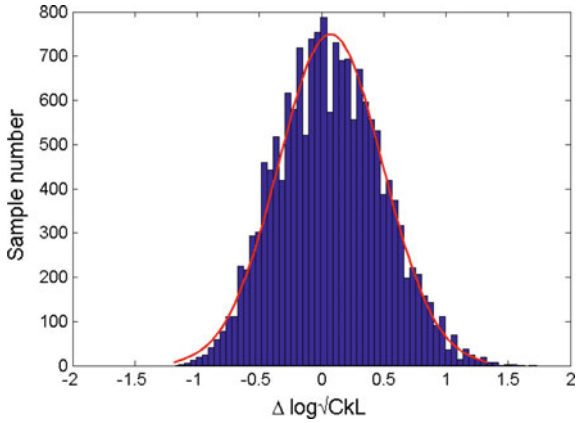


Fig. 2 Verification of the irregularity strength model with data of 2013 from Haikou

The result shows that an unbiased estimation achieved for the model, better than the performance of the model in Ref [7]. which is a biased estimation and had a larger variance [4].

2.3.2 External Consistency Accuracy of Strength Model

In order to further verify the validity of the model, the external consistency accuracy of the model is analyzed by using the GPS scintillation data in Haikou and Guangzhou in 2013. Figure 2 shows the results of the model with data in Haikou. It has a mean of 0.07, and a variance of 0.42 for the variance. The result of a mean of 0.01 and a variance of 0.42 is achieved for the test with data from Guangzhou [4].

It can be seen that the irregularity strength model has a satisfied performance for Haikou and Guangzhou areas, even in a year of higher solar activity (2013). In these two tests, the model error maintains the normal distribution with a mean close

to zero and a variance of 0.42. This result is consistent with that from the Haikou data set of 2003–2008 [4].

The consistency accuracy analysis shows that the established irregularity model has a good applicability in the low latitude area of China. Features of the zero-mean and the constant variance for the model residual make it possible to estimate the strength of irregularity with a given confidence level. This is of great significance to the studies such as the system availability assessment under scintillation.

3 Signal Propagating Model Based on Phase Screen Theory

In phase screen theory, the irregularity layer is taken as a thin screen at certain height. When wave passes the screen, only phase of the wave varies. Emerging wave from the phase screen propagates in the free space and then reaches receiver antenna. As the wave propagates towards the ground, interference across the wave-front creates complex amplitude and phase diffraction patters. Scintillations are produced when these spatial diffraction patterns are transformed into temporal ones [12, 13]. To develop the propagating model with the phase screen theory, a random phase screen should be generated and the parabolic wave equation (PWE) is resolved to get the complex amplitude of signal reaching the receiver antenna.

3.1 Generation of Random Phase Screen

According to Rino's work, when the complex amplitude signal passes the phase screen, the variation of phase caused by phase screen follows the specified power spectral density (PSD) function [12, 13]:

$$\Phi_\phi(\kappa) = \frac{\lambda^2 r_e^2 \sec^2\theta ab(2\pi/1000)^{p+1} C_k L}{\left(\kappa_0^2 + A\kappa_x^2 + B\kappa_x\kappa_y + C\kappa_y^2\right)^{(p+1)/2}} \tag{7}$$

where $\kappa_0 = 2\pi/L_0$ is the outer scale wave number, L_0 is the outer scale of ionospheric irregularity, p is phase spectral index, a, b are respectively axis ratio along and transverse to the principal axis of irregularity. κ_x, κ_y are respectively component in X axis and Y axis of the transverse wave number. $C_k L$ is strength of the irregularity. A, B, C are coefficients depend on propagation geometry angle θ, φ and geomagnetic inclination ψ. Detailed definition of these parameters can be referred to Ref. [13].

Random phase variation which follows the power spectral density function of (7) can be generated by passing Gaussian white noise with zero mean value into a filter with a transfer function of $\sqrt{\Phi_\phi}$. This can be realized in a way of digital signal processing and the algorithm has been given by Rino [12].

3.2 Propagating Model Based on Phase Screen Theory

In phase screen theory, the irregularity layer is taken as a thin screen at $z = 0$. When wave passes the screen, only the phase of the wave varies. That means, the complex amplitude of the signal just below the screen is [2, 12]:

$$U(\boldsymbol{\rho}, \mathbf{0}^+) = U(\boldsymbol{\rho}, \mathbf{0}) e^{i\phi(\boldsymbol{\rho})} \tag{8}$$

where $U(\boldsymbol{\rho}, \mathbf{0})$ is the wave signal incidence on the phase screen, $\phi(\boldsymbol{\rho})$ is the phase variation caused by electron density fluctuation ΔN_e concentrated on the phase screen.

$$\phi(\boldsymbol{\rho}) = k \int_{-L/2(\sec\theta)}^{L/2(\sec\theta)} \Delta n(\mathbf{r}) dl = \lambda r_e \sec\theta \int_{-L/2}^{L/2} \Delta N_e(\boldsymbol{\rho} + \tan\theta \widehat{a}_{k_T}(\eta - z), \eta) d\eta \tag{9}$$

Where ρ is the distance vector in the transverse (horizontal) plane, $\widehat{a}_{k_T} = (\cos\varphi, \sin\varphi)$ is a unit vector in transverse plane, r_e is the classic electric radius, and dl is the differential length element along the propagation path. The geometric parameters are as shown in Fig. 3.

Emerging wave from the phase screen propagates in the free space and then reaches receiver antenna. For signal propagation in free space (Δn is 0 then), (9) can be solved by using transverse Fourier transform of $U(\boldsymbol{\rho}, z)$ defined as:

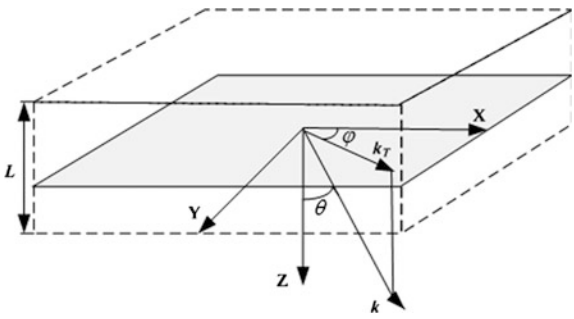


Fig. 3 Coordinate system in phase screen simulation. The origin is located at the ionospheric pierce point (IPP), axis *X*, *Y*, *Z* are respectively pointing to geomagnetic north, geomagnetic earth and vertical downward

$$U(\boldsymbol{\rho}, z) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widehat{U}(\boldsymbol{\kappa}, z) e^{i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} d\boldsymbol{\kappa} \tag{10}$$

Where, $\boldsymbol{\kappa}$ is the transverse Fourier wave number. The complex amplitude of signal reaching receiver antenna can be expressed as:

$$U(\boldsymbol{\rho}, z) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widehat{U}(\boldsymbol{\kappa}, 0+) \exp\left[-i\left(\frac{\kappa^2}{2k} z \sec\theta + (\widehat{a}_{k_T} \cdot \boldsymbol{\kappa}) z \tan\theta\right)\right] e^{i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} d\boldsymbol{\kappa} \tag{11}$$

4 Application of Scintillation Models

With the date, space environment conditions (sunspot number, magnetic field etc.), locations for satellites and users, observing geometry, and system information (operating frequency) as input, the ionosphere scintillation model can estimate the irregularity strength at the ionospheric pierce point where the signal pass through the phase screen. Then the phase screen can be produced and the signals received by the user can be resolved with the signal propagating model. Consequently, analysis could be carried out on these received signals to assess the scintillation effects on system performance. The following shows the potential applications of the established scintillation model.

4.1 Simulation of Scintillation Affected Signals

Under the "frozen field" hypothesis, the variation of the complex amplitude of signal can be transformed into the time series of the signal intensity and phase [12, 13]. These signal intensity and phase time series, modulated on the signals from a GNSS simulator, can then be used to test the tracking performance of the receiver under various scintillation conditions [15].

Figure 4 shows the simulated time series of the signal intensity with the scintillation model. The amplitude scintillation index is also shown for the signal intensity. In the simulation, the date is 12:30 (UT) 14th Sep. 2014, the satellite is in a geosynchronous orbit with the nadir point at the 85°E, 0°N, user location is set to 20°N, 110°E, and the signal frequency is 1575.42 MHz. The space environment parameters are obtained from NOAA according to the date in the analysis (the same as below). We can see that a weak ionospheric scintillation is simulated in this scenario.

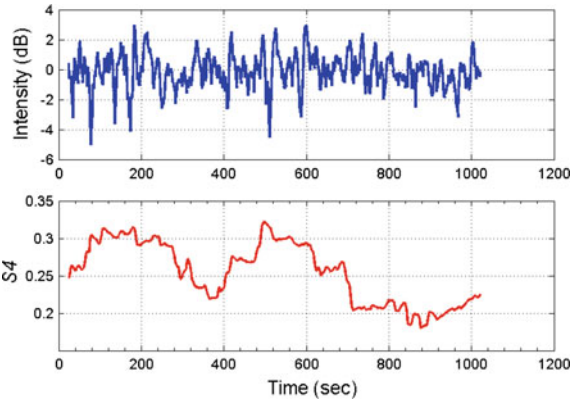


Fig. 4 Simulated signal intensity (*up*) and the corresponding S4 scintillation index (*below*)

4.2 Maps of Scintillation Effects

4.2.1 Computation of Scintillation Index

An important application of the ionospheric scintillation model is to analyze the system performance under the influence of scintillation. It can be realized with the following steps:

(1) Map of the irregularity strength distribution in the specified region is generated by using the ionospheric irregularity model.
(2) With the phase screen model, the ionospheric scintillation effects for the user within the designated area can be obtained as a map of scintillation index (such as the amplitude scintillation index).
(3) The performance degradation in the area can then be evaluated by a receiver model with the scintillation index as input.

In the above process, the scintillation index can be computed based on the phase screen theory. Rino gives the analytical resolution of the amplitude scintillation index S_4 under the phase screen theory [12].

$$S_4^2 = r_e^2 \lambda^2 (L \sec\theta) C s Z^{v-1/2} \left[\frac{\Gamma((2.5-v)/2)}{2\sqrt{\pi}\Gamma((v+0.5)/2)(v-0.5)} \right] F \tag{12}$$

$$Z = \frac{\lambda z_R \sec\theta}{4\pi} \tag{13}$$

$$v = p/2 \tag{14}$$

where r_e is the classical electron radius, λ is the wave-length of the signal, θ is the zenith angle of the incident signal at the receiver, p is the spectral index and the typical value of 2.5 is used for low latitude areas, F is the propagating geometry

factor, Γ is the gamma function [7, 8]. $z_R = zz_s/(z+z_s)$, where z, z_s is the distance from the phase screen to the ground receiver and the satellite, respectively. *CsL* is obtained with the ionospheric irregularity strength model presented before. Parameters for geometry configuration in simulation can be calculated with the positions of the satellite and user.

4.2.2 Scintillation Affects Mapping

In previous work, we have established a GNSS receiver model with the scintillation index S_4 as the input to analyze the performance degradation of GNSS [3]. Here in this paper, we only focus our efforts on the distribution of the irregularity strength and the scintillation index. With these data, the map of the degrading performance can be readily obtained.

Figure 5 shows the distribution of the irregularity strength using the model constructed in Sect. 2. The date is 16:00 (UT), 10th Sep. 2001. The ground range for the user is [−50°S, 50°N], [50°E, 180°E]. As the ionospheric scintillation index depends not only on the irregularity strength, but also on the observing geometry, two typical cases were analyzed below [12, 14].

In the first case, the user in the specified area has the same observing geometry (the same azimuth and elevation angle). This situation is specialized in the analysis of the spatial and temporal distribution of the ionospheric scintillation effects under the same observing configuration.

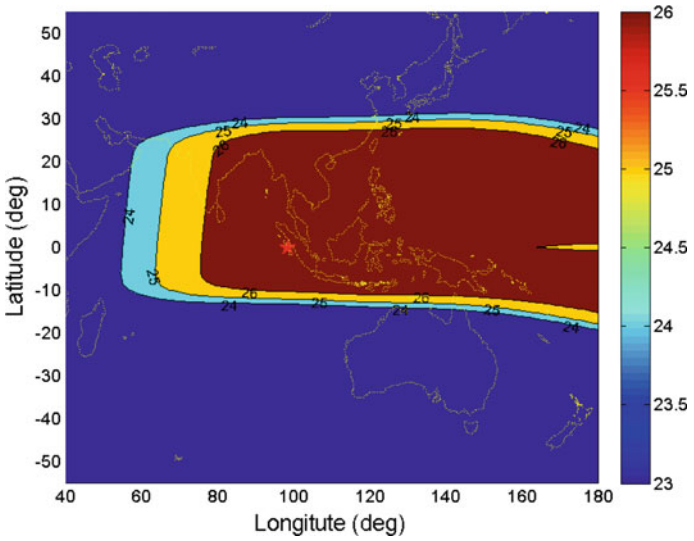


Fig. 5 *CsL* map (16:00 UT, 10th Sep. 2001)

Figure 6 shows an example of the S_4 map in this case. In simulation, the user received signals at the same azimuth angle of 0°, and elevation angle of 32°, the date is set to 16:00 (UT), 10th Sep., 2001, the user range is [−50°S, 50°N] in latitude and [50°E, 180°E] in longitude, and the signal frequency is 1575.42 MHz. The irregularity strength is from the result shown in Fig. 5.

It can be seen that there are two areas with the maximum index in the South of China and the South China Sea. The intensity of scintillation effects depend on the observing geometry and the magnetic field as well. The different geometric relation between the same propagation direction of the signals received and magnetic field results in the difference in scintillation effects distribution.

The second case is to specify the location of the satellite, and analyze the scintillation effects on users at the same time. In this scenario, the observing geometry changed with user locations. Both the changed geometry configuration and its relations with the magnetic field result in the variation in scintillation effects on the computed S_4 map.

Figure 7 shows the distribution of scintillation effects on the user in the area of [−30°S, 30°N] and [50°E, 180°E] viewing a GEO satellite at the 68.5°E. The date is 16:00 (UT), 10th Sep. 2001, and the signal frequency is 1575.42 MHz. The irregularity strength from Fig. 5 is also used here. The elevation angle user viewing the satellite is given in the figure (green line) at the same time.

It can be seen that with the lower elevation, the user is affected by a more severe scintillation effect. When the elevation is below the 10°, a modest scintillation effect happened with an amplitude scintillation index of 0.5.

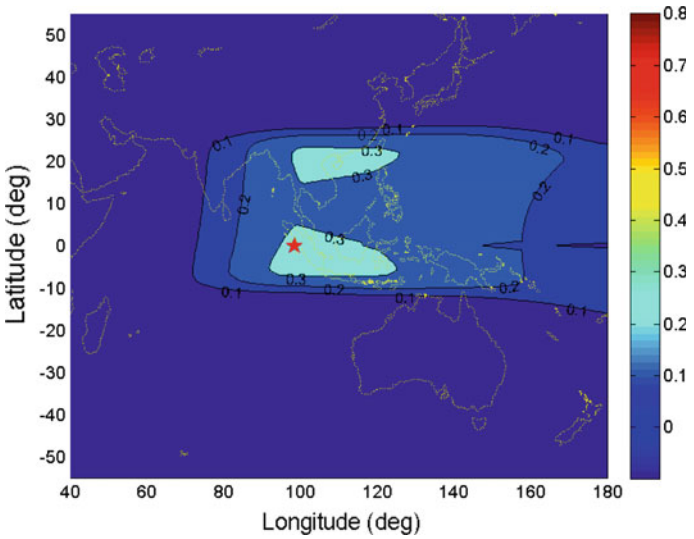


Fig. 6 S_4 map with the same geometry configuration (16:00 UT, 10th Sep. 2001)

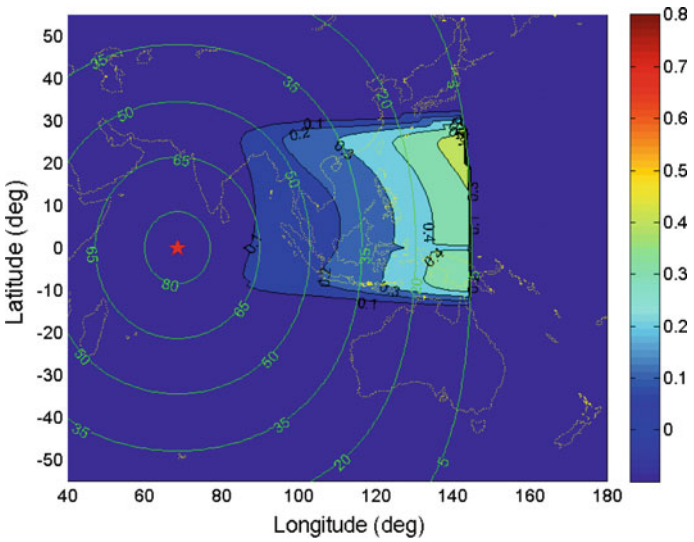


Fig. 7 S_4 map with different geometry configuration (16:00 UT, 10th Sep. 2001)

When combined the map of scintillation index with the GNSS receiver model, it can be readily to assess the users' performance distribution under scintillation. In our previous work, the GISM model is employed to produce the scintillation index map [4]. As GISM can only produce the scintillation index map for the vertical incidence, the flexibility of the analysis is limited. With the ionospheric scintillation model established in this paper, we divide the scintillation models into irregularity strength model and signal propagation model. So we can obtain the scintillation index for arbitrary geometry configuration in observation, and effectively improve the flexibility and applicability of scintillation effect analysis.

5 Conclusion

Ionospheric scintillation is an important disturbing source impacting various space application systems. The ionospheric scintillation models provide a useful tool to assess the scintillation effects on these systems.

The ionosphere scintillation models could be divided into the irregularity model and the signal propagation model. We have set up an ionospheric irregularity strength model for the low latitude area of China with GPS scintillation data from Haikou and Guangzhou. The accuracy of the model is also analyzed. The results show that error of the model observes a normal distribution with a mean of near zero and a variance of 0.42.

Based on the theory of phase screen, signal propagation model is developed by using the split step algorithm and the Discrete Fourier Transformation to solve the wave equation. With the assumption of "frozen field", the resolved complex amplitude of signal can be transformed into the spatial variation of signal intensity and then the temporal variation of signal intensity.

The ionospheric scintillation model output includes the time series of signal intensity and phase, and then the scintillation index. Examples are given for various analysis based on these different outputs. The separation between the irregularity model and the signal propagating model make it possible to assess the scintillation effects under different observation scenarios flexibly.

References

1. Hegarty C, El-Arini BM, Kim T, Ericson S (2000) Scintillation modeling for GPS/WAAS receivers. ION GPS, pp 799–807
2. Carrano CS, Groves KM, Caton RG (2010) A phase screen simulator for predicting the impact of small-scale ionospheric structure on SAR image formation and interferometry. IGARSS 2010, pp 162–165
3. Liu SH, Liu D (2015) Modeling the effects of ionospheric scintillation on BD augmentation system. Chin J Radio Sci 30(1):135–140
4. Liu D, Guo S, Yu X et al (2016) Irregularity strength modeling in ionospheric scintillation model. Chin J Radio Sci 31(3):432–437
5. Liu D, Yu X, Feng J et al (2016) Simulating the impacts of ionospheric scintillation on GNSS signals with phase screen method. Chin J Radio Sci 31(4):632–637
6. Beniguel Y (2002) GIM: a global ionospheric propagation model for scintillation of transmitted signals. Radio Sci 37(3):1032–1044
7. Fremouw EJ, Robins RE (1985) An equatorial scintillation model. Physical Dynamics, Inc., 30 September 1985
8. Secan JA, Bussey RM, Fremouw EJ, Basu S (1995) An improved model of equatorial scintillation. Radio Sci 30:607–617
9. Robins RE, Secan JA, Fremouw EJ (1986) A mid-latitude scintillation model. Northwest Research Associates Inc, October 1986
10. Fremouw EJ, Leadabrand RL, Livingston RC et al (1978) Early results from the DNA Wideband satellite experiment—complex signal scintillation. Radio Sci 13(1):167–187
11. Vanzandt TE, Clark WL, Warnock JM (1972) Magnetic apex coordinates: a magnetic coordinates system for the ionospheric F2 layer. J Geophys Res 77(13):2406–2411
12. Rino CL (2011) The theory of scintillation with applications in remote sensing. Wiley, New York
13. Rino CL, Power A (1979) Law phase screen model for ionospheric scintillation 1. Weak scatter. Radio Sci 14(6):1135–1145
14. Rino CL, Fremouw EJ (1977) The angle dependence of singly scattered wavefields. J Atmos Terr Phys 39:859–868
15. Knight MF (2000) Ionospheric scintillation effects on global positioning system receivers. The University of Adelaide, Adelaide South Austrilia
16. Yeh CK, Liu CH (1982) Radio wave scintillations in the ionosphere. Proc IEEE 70(4): 324–360

A New Algorithm for Receiver Integrity Monitoring with Receiver Clock Error Auxiliary

Ye Ren, Xiaohui Li and Longxia Xu

Abstract The receiver integrity monitoring technology is one of the important technologies to reflect the GNSS integrity. It can be divided into receiver autonomous integrity monitoring (RAIM) and auxiliary autonomous integrity monitoring. Apart from the redundant observations of the receiver are used for RAIM, there are various parameter information provided form receiver and the receiver clock error is one of them. As the stability of the receiver oscillators to 10–10/s, this makes it possible to make short-term predictions. In this paper, the receiver clock error is introduced and the receiver clock error auxiliary GNSS integrity monitoring method is proposed. In this paper, the principle of the method is introduced and the algorithm flow is established firstly. Then the key steps involved in the method are discussed, including receiver clock error prediction method, threshold value determination method, fault detection method and fault identification method. Among these key steps, threshold determination method are fully discussed and validated. Results show that the threshold determination method is proper in the method. Finally, the performance and feasibility of the proposed algorithm are evaluated. The results show that the proposed method can effectively detect and identify the faulty satellites. This method makes full use of the historical clock error information, and provides a simple and effective method for the receiver integrity monitoring.

Keywords Integrity monitoring · Receiver clock error · Fault detection · Fault identification

Y. Ren (✉) · X. Li · L. Xu
National Time Service Centre, Chinese Academy of Sciences, Xi'an, China
e-mail: renye@ntsc.ac.cn

Y. Ren · X. Li · L. Xu
Key Laboratory of Precision Navigation and Timing,
National Time Service Center, Chinese Academy of Sciences, Xi'an, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_12

1 Introduction

According to the definition of ICAO, integrity is one of GNSS four performance indicators (accuracy, integrity, availability, continuity) [1, 2]. Integrity refers to the ability of the system to issue timely and effective alerts to users when it can not be used for certain scheduled operations [3]. User integrity monitoring is one of the important methods of GNSS integrity monitoring. User integrity monitoring methods can be achieved in two ways: one is based on the redundant observations at user terminal, which is usually called as RAIM (Receiver Autonomous Integrity Monitoring) [4–6], and the other is based on the user auxiliary information, such as receiver clock error system time offset [7].

Among the auxiliary information, the receiver clock error is the deviation between the time of the atomic clock on the receiver and the GNSS constellation system. After several generations of development, the stability of the receiver oscillator has been better than 10–10/s. Theoretically, the prediction error may not exceed 1 ns for 10 s of prediction. Therefore, the high stability of the clock error information makes it possible to practice a short-term prediction of the clock error. Furthermore, the predicted clock error can be treated as a known parameter for the receiver integrity detection. However, the use of known clock error as auxiliary information is premised on: (1) any system anomalies that affect the user's position estimation will also affect the estimation of the clock error; (2) the receiver atomic clock stability maintains within 10–10/s. The receiver clock has no significant or unpredictable frequency hopping and the receiver keeps this performance long enough to accurately estimate the parameters of the receiver clock error; (3) the observations used to establish the receiver model does not contain pseudorange offset.

In this paper, the receiver clock error is introduced as auxiliary information for user integrity monitoring, which is named as CEAIM (Clock Error Auxiliary Integrity Monitoring).

2 CEAIM Algorithm

The principal of CEAIM is described as follows: (1) a current clock error t_{pre} is predicted by quadratic polynomial model with historical clock error information; (2) a current clock error t_{solve} is resolved based on the current pseudorange observation positioning solution; (3) a difference between the predicted clock error and the calculated clock error is calculated. If the difference is beyond the preset threshold value t_{TH}, the algorithm would determine that one of the current observations may contain anomalies (4) when there is an anomalies, the fault identification is lunched and the identified fault observations are removed (Fig. 1).

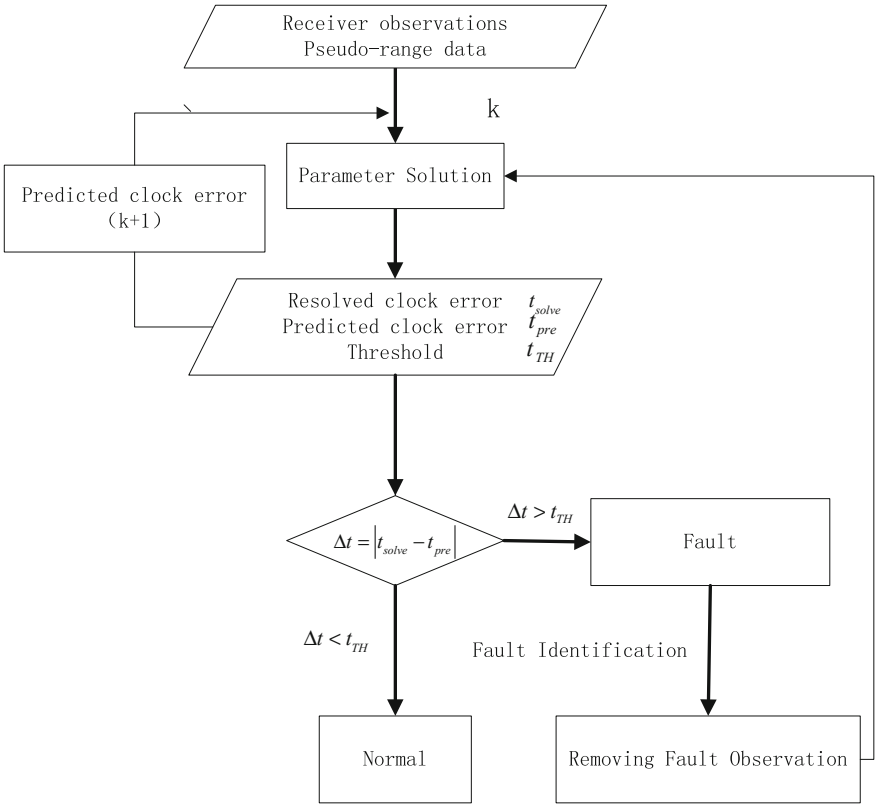


Fig. 1 CEAIM flow diagram

3 Key Steps

3.1 The Determination of Threshold

The deviant of the predicted clock error t_{pre} and the resolved clock error t_{solve} should be very close and is assumed as following Normal distribution with zero mean and σ^2 variance, under faulty-free constellation system. However, as long as there is a system failure, it may affect the resolved clock error t_{solve}, and the deviant may fall beyond the Normal distribution with 95% confidence interval. Therefore, this can be used as a basis for fault detection as well as the determination of threshold value.

A threshold should be determined prior to the integrity monitoring. The threshold is the sum of the errors introduced during the prediction and resolution of the system clock error under a faulty-free constellation system. Let the standard deviations of the clock errors introduced by t_{pre} and t_{solve} are σ_{pre} and σ_{solve}, then:

$$\mathrm{var}(t_{pre}) = \sigma_{pre}^2 \tag{1}$$

$$\mathrm{var}(t_{solve}) = \sigma_{solve}^2 \tag{2}$$

According to the error theory, the variance of $t_{pre} - t_{solve}$ is:

$$\mathrm{var}(t_{solve} - t_{pre}) = \sigma_{solve}^2 + \sigma_{pre}^2 \tag{3}$$

Under faulty-free constellation system, $t_{pre} - t_{solve}$ follows a normal distribution with a mean of zero. A faulty constellation system is defined as a situation that the variance of the deviant fall beyond 3σ. Therefore, the clock error threshold t_{TH} is determined as:

$$t_{TH} = 3\sqrt{\sigma_{solve}^2 + \sigma_{pre}^2} \tag{4}$$

3.2 Fault Detection

Let $\Delta t = \left| t_{solve} - t_{pre} \right|$, according to the hypothesis test:

$$H_0 : \Delta t < t_{TH} \tag{5}$$

$$H_1 : \Delta t > t_{TH} \tag{6}$$

when Δt does not exceed the threshold value, the algorithm decides that there is no fault in the current observations; if Δt exceeds the threshold value, the algorithm decides that there is faulty in the current observation.

3.3 Fault Identification

Based on the subset comparison method, t_{solve}^i represents as resolved clock error by removing of the *i*th observation:

$$t_{solve}^i = ((H_i^T H_i)^{-1} H_i^T)_{4,:}\, y_i, \quad i = 1, \ldots, n \tag{7}$$

The subscript "4,:" represents all the columns in the fourth row, while H_i, y_i are the measurement matrix and the pseudo-range observation vector, respectively, after removing the *i*th observation. For the current observations, n observations can deduce n groups of solution value. Then, Δt_i is defined as:

$$\Delta t_i = \left| t^i_{solve} - t_{pre} \right| \tag{8}$$

If there is a subset which corresponds to $\Delta t_{\min} = \min(\Delta t_i)$, the removed observation from this subset is considered that contains a fault.

4 The Algorithms Performance

4.1 Data Description

The underlying algorithm is verified by a experimental evaluation. The data of BeiDou B3 frequency band of Xi'an Observatory, on October 9, 2015 (0:01:06:59:59), are collected and the sampling interval is 10 s.

In order to evaluate the algorithm, two indexes are introduced: fault detection rate (FDR) and fault identification rate (FIR). During observation period, there are N observation time t_i(i = 1,…,N) and there are K observations in each observation time. At an observation time t_i, a particular observation $GNSS_{j_0}(j_0 = 1, \ldots, K)$ is artificially selected and a pseudorange bias is added to $GNSS_{j_0}$

(1) FDR: a ratio of the number of faults detected N_D to the actual number of failures N during the observation period:

$$FDR = \frac{N_D}{N} \tag{9}$$

(2) FIR: a ratio of N_I, which is the times that identified fault observation function $GNSS_j$ meets the artificial fault observation $GNSS_{j0}$, to the actual number of failures N during the observation period:

$$FIR = \frac{N_I}{N} \tag{10}$$

4.2 The Determination of Threshold

According to the key steps, the threshold is one of crucial factors that may affect the algorithm performance. Therefore, further analysis is lunched to verify the determination of threshold.

4.2.1 Threshold Under Obscure Environment

The thresholds may vary with cut-off angle. For BeiDou, under the cut-off angle of 10°, the threshold is 2.3 m, while the threshold is 4.5 m when the cut-off angle

of 30°. The possible reason is that as the cut-off angle increases, the observation quantity decreases and the parameter solution ability decreases, which followed by the increase of uncertainty of timing and positioning. Then the deviation between the predicted value and the resolved value increases. With the obscuration of the observation deteriorating, that is, the cut-off angle increases from 0° to 30°, as shown in Fig. 2, the threshold value is about 2 m when the shielding angle is less than 20°. When the cut-off angle is above 20°, the threshold value is determined to be about 4 m. Therefore, the threshold value should not be a constant value as the increase of cut-off angle.

The observation of BD 02 satellite is artificially introduced pseudo-range deviation value, which are 100, 200, 300 m, respectively. In Table 1, there is a tendency of decrease for both FDR and FIR as the cut-off angle increases. The reason contributes to this phenomenon may as the cut-off angle increases, the number of observations decreases and the uncertainty of parameter solution increases. Then the deviation of predicted clock error against the resolved clock error may increases, which may lead to the decrease of FDR. Meanwhile, with the decrease of observations, the correlation between pairs of observations my increase and the difficulty for the identification my increase which may decrease the FIR.

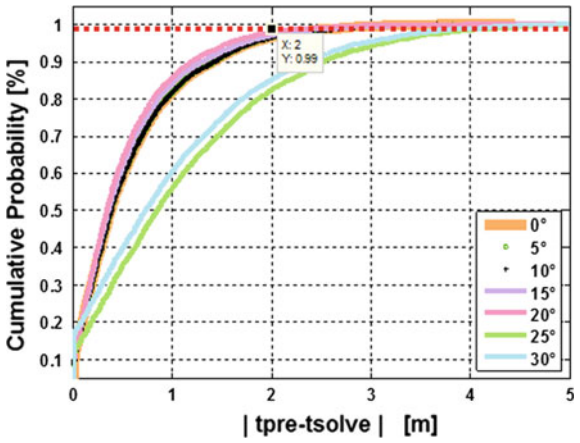


Fig. 2 FDR and FIR under pesoderange of 100/200/300 m

Table 1 FDR/FIR and threshold under different cut-off angle

Cut-off angle (°)	Threshold (m)	FDR (%)	FIR (%)
0	2	93|98|98	90|96|97
5	2	94|98|98	89|96|96
10	2	94|97|98	90|95|96
15	2	95|97|98	91|95|97
20	2	94|97|99	91|95|97
25	4	92|94|95	84|89|91
30	4	82|90|90	80|85|87

4.2.2 Threshold on the Algorithm Performance

The determination of threshold is based on a certain experience selection according to a statistical distribution. In order to verify the feasibility of the threshold determination method, the effect of the threshold uncertainty on the algorithm performance is discussed. Adding uncertainty of 1.5 and −1.5 m on the threshold 2 m, FDR and FIR are shown in Fig. 3 and Table 2. If the threshold value is set below 2 m, even if there is no fault introduced in the system, there is faulty detected. This is the false alarm. When the threshold is set at 0.5 m, the false alarm rate is as high as 40%; when the threshold is set at 3.5 m, the false alarm is reduced to 4.8%. Meanwhile, as pseudorange deviation of 300 m is introduced, when threshold is 0.5 m, the corresponding FDR and FIR is slightly higher than the value under the threshold value of 3.5 m. These indicate that the higher the pseudorange deviation, the higher FDR and FIR. The results show that the threshold determination method for this algorithm is feasible.

4.3 The Algorithm Performance

In the open area (cut-off angle of 10°), the threshold is t_{TH} = 2.8 m, Table 3 shows that with the pseudorange deviation increases, there is and increase tendency of

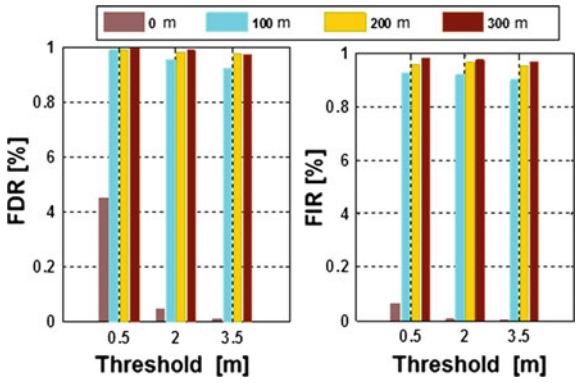


Fig. 3 FDR and FIR under different threshold

Table 2 FDR/FIR with different threshold

	Threshold (m)	Pseudo-range deviation (m)			
		0 (%)	100 (%)	200 (%)	300 (%)
FDR	0.5	45.1	98.9	99.1	99.5
	2	9.8	95.5	98.2	98.8
	3.5	4.8	92.2	97.5	97.8
FDI	0.5	6.4	92.4	95.8	98.1
	2	0.4	91.8	96.4	97.2
	3.5	0.1	89.8	95.2	96.4

Table 3 FDR/FIR under cut-off angle 10°

	Pseudo-range deviation (m)					
	50	100	150	200	250	300
FDR (%)	88.47	94.37	95.80	96.34	96.87	97.94
FIR (%)	83.02	90.71	94.10	94.73	95.62	96.51

Table 4 FDR/FIR under cut-off angle 30°

	Pseudo-range deviation (m)					
	50	100	150	200	250	300
FDR (%)	84.90	88.29	88.65	90.08	90.08	90.08
FIR (%)	70.78	81.05	82.48	85.61	87.04	87.31

FDR and FIR. Generally speaking, with the pseudorange deviation is introduced more than 100 m, the FDR can reach more than 94%.

In the case of obscure (cut-off angle of 30°), the threshold is $t_{TH} = 4.5$ m, Table 4 shows the FDR and FIR are little worse than the value under the open area. In general, with the pseudorange deviation exceeds 100 m, FDR can only reach 85%.

5 Conclusion

A user integrity monitoring algorithm is proposed by introducing receiver clock error as an auxiliary information. Firstly, the principle and flow of the algorithm are established, that is, comparing the deviant between the predicted clock error and the resolved clock error and the preset threshold value in order to determine the working status of observations. Secondly, it describes some key steps involved in the algorithm, including threshold determination method, fault detection method and fault identification method. Moreover, among these important parameters, threshold is further discussed on the aspect of the effect on the algorithm performance. Finally, the feasibility of the algorithm is verified by experiments based on real-time observation from BeiDou. The results show that: (1) the threshold value should vary with the increase of obscure environment; (2) the higher the pseudorange deviation, the weaker of algorithm performance; (3) with the deteriorate of the observation environment, the ability on both fault detection and fault identification become decrease. Generally speaking, the proposed algorithm effectively improves the utilization of the known information on the receiver, and provides a simple and effective way for user integrity monitoring.

References

1. Kovach KL, Dyke KLV (1997) GPS in ten years. In: The proceedings of the 10th international technical meeting of the satellite division of the Institute of Navigation (ION-GPS-97), Kansas City, USA, pp 1251–1260
2. Hewitson S, Wang J (2007) GNSS receiver autonomous integrity monitoring with a dynamic model. J Navig 60(2):247–263
3. Wang J, Obser P (2009) On the availability of fault detection and exclusion in GNSS reviver autonomous integrity monitoring. J Navig 62(2):1–11
4. Parkinson B, Axelrad P (1988) Autonomous GPS integrity monitoring using the pseudorange residual. Navigation 35(2):49–68
5. Sturza M, Brown AK (1990) Comparison of fixed and variable threshold RAIM algorithms. In: Proceedings of ION, pp 437–443
6. Lee YC (2004) Performance of receiver autonomous integrity monitoring (RAIM) in the presence of simultaneous multiple satellite faults. In: The 60th annual meeting of the Institute of Navigation. The Institute of Navigation, Sayton, OH, pp 687–697
7. Sun J et al (eds) (2016) A new method for multiple outliers detection in receiver autonomous integrity monitoring. In: China satellite navigation conference (CSNC) 2016 proceedings: volume II, lecture notes in electrical engineering 389, pp 151–164

FTS-Based Link Assignment and Routing in GNSS Constellation Network

Tian-yu Zhang, Gang-qiang Ye, Jing Li and Jing-wen Xu

Abstract In order to solve the link assignment problem in GNSS constellation, which is affected by the performance of dynamic topology, the method of Finite Time Slot (FTS) is proposed for GNSS constellation in this paper. Firstly, we model a GNSS constellation network as a FTS, where each time slot corresponds to an equal-length interval in system period of the GNSS network. Then, we formulate link assignment as a mixed integer linear programming problem, and proposing an iterative optimization method. Finally, the link assignment efficiency and routing performance are simulated for several kinds of time slot. The researching results will support on orbit management and maintenance for Chinese GNSS Constellation.

Keywords Finite time slot · GNSS constellation · Inter satellite links · Link assignment · Routing

1 Introduction

Inter satellite links (ISL's) is used to realize autonomous navigation and real time information transmission in GNSS constellation network. Therefore, the integrated space-ground dynamic network with function of data transmission and precision measurement is established by ISL's in GNSS. In GNSS, ISL's is a very important technique system, which is a kind of method to guarantee the security and reliability in GNSS constellation, to improve the overall performance effectively. It also can support to establish the space based information network [1–5].

T. Zhang · G. Ye · J. Li (✉) · J. Xu
The State Key Laboratory of Astronautic Dynamics, Xi'an Satellite Control Center, XSCC, Xi'an 710049, China
e-mail: 279182452@qq.com

T. Zhang
Software Engineering Institute, Xi'an Jiaotong University, Xi'an, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_13

In GNSS constellation, one satellite moves along its orbit, its relative position to other satellite and to the between ground station changes continuously and so does the visibility matrix for inter satellite and satellite-ground from it. The inter satellite links and satellite-ground station links are dynamically established [4, 5]. Therefore, we need to select the optimal interval to guarantee maximize the overall GNSS performance [6].

There are so many research about the link assignment at home and abroad. There are some quantitative analysis about Dynamic change features of LEO and Walk constellation network topology [7]. There is design dynamic switching topology of networks about navigational satellite [8]. There is communication satellite constellation reconstruction method according to lose efficacy ISL [9]. Harathi said that a fast link assignment algorithm for satellite communication networks [10]. Noakes expresses that an adaptive link assignment algorithm for dynamically changing topologies [11]. Hong expresses that FSA-Based link assignment and routing in low-earth orbit satellite networks [12]. Cain said that a distributed link assignment (reconstitution) algorithm for space-based SDI networks [13]. Xu expresses that a hybrid navigation constellation intersatellite link assignment algorithm for the integrated optimization of the inter-satellite observing and communication performance [5].

In order to solve the link assignment problem in GNSS constellation, which is affected by the performance of dynamic topology, the method of Finite Time Slot (FTS) is proposed for GNSS constellation in this paper. Firstly, we model a GNSS constellation network as a FTS, where each time slot corresponds to an equal-length interval in system period of the GNSS network. Then, we formulate link assignment as a mixed integer linear programming problem, and proposing an iterative optimization method. Finally, the link assignment efficiency and routing performance are simulated for several kinds of time slot. The researching results will support on orbit management and maintenance for Chinese GNSS Constellation.

2 Problem Analysis of ISL

ISL need to analyse the dynamic topology changes of GNSS constellation network. We traverse the whole GNSS constellation cycle and discretize it into finite time slot. The discretization whole period of GNSS constellation can be described as finite time slot, so that we can move on to the simulation in each finite time slot, which provides convenience for modelling and analysis about GNSS constellation [14]. We describe the GNSS constellation with the visible matrix in each finite time slot. In order to adapt to the antenna operation mode, we determines the satellite node in each finite time slot as communication node between ground stations and GNSS constellation. Finally we establish time slot table and routing table which can help us manage GNSS constellation more accurately and orderly. The specific model is shown in Fig. 1.

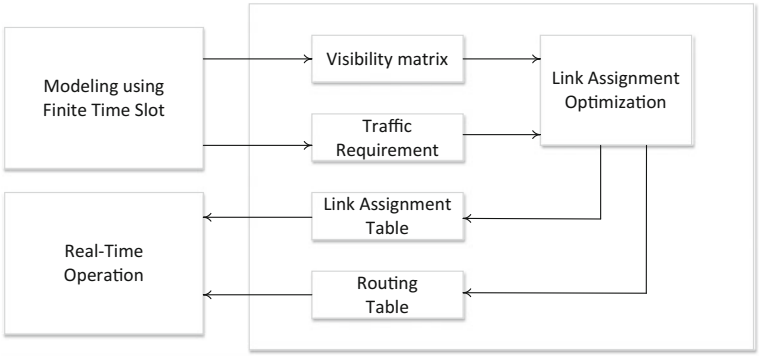


Fig. 1 FTS modeling

2.1 Calculate Visibility Matrix

The visible matrix is a record that describes some linked conditions between satellite and ground station or between satellites. Visible matrix model includes the matrix between satellites and the matrix between satellites and ground station. The visible matrix between satellites and ground stations has matured; the visible matrix between satellites need to be considered about intersatellite visible geometry, the range of signal and the range of antenna scanning [15]. The basic conditions that one satellite can link with another are shown as follows:

$$O_E E > h \tag{1}$$

$$L_{ij} \leq \max\{L_{gmax}, L_{max}\} \tag{2}$$

$$\gamma_{ij} \leq \beta_{max},\ \alpha_{ij} < \alpha_i;$$

$$\text{Or } \gamma_{ij} > \beta_{max},\ 90^\circ - \beta_{max} < \alpha_{ij} < \alpha_i \tag{3}$$

Constraint (1) specifies that the distance $O_E E$ between earth's core and satellite must be longer than the distance h between earth's core and ionized stratum. Constraint (2) expresses that the intersatellite distance L_{ij} must be longer than the upper-limit L_{gmax} of intersatellite geometrical distance and the upper-limit L_{max} of signal transmission distance. Constraint (3), stats that γ_{ij} is the geocentric angle between satellites, β_{max} is the largest coverage angle by satellite antenna beam, α_{ij} is cover angle between satellites and α_i is antenna beam half-power angle between satellites.

2.2 Confirm Node Satellite

Considering the mechanical characteristics of ground station antenna, in this paper, we choose a visible satellite as satellite node for each ground station in each finite time slot. We use the visible time between ground stations and satellites and the number of satellite that the satellite node can link as judging basis, so that we can select the greatest advantage one as a node to link with ground station.

As shown in Fig. 2, the visible satellites those the ground station can link are Satellite a, Satellite b and Satellite c. For the ground Station A, Satellite a, Satellite b and Satellite e have a higher orbit than others, so that they have larger visible radian. In this time slot, Satellite b have 4 visible satellites and Satellite a have 2 visible satellites. So that, we choose the Satellite b as the node of the Station A.

Due to the difference of visible time between satellites and ground stations in each finite time slot for GNSS constellation, which has complex topological structure, we select the optimal visible satellite as a satellite node for each ground station in each finite time slot, based on the genetic algorithm.

2.3 Program Link Assignment

In each finite time slot, ISL has the role of information interaction between satellites. Time slot table plans the linking target and the state of sending and receiving for each time slot, which can ensure the accurate of data transmission and the stable operation of entire network for GNSS constellation. The positioning will be more precise and the autonomous navigation will be more stabilized with the more

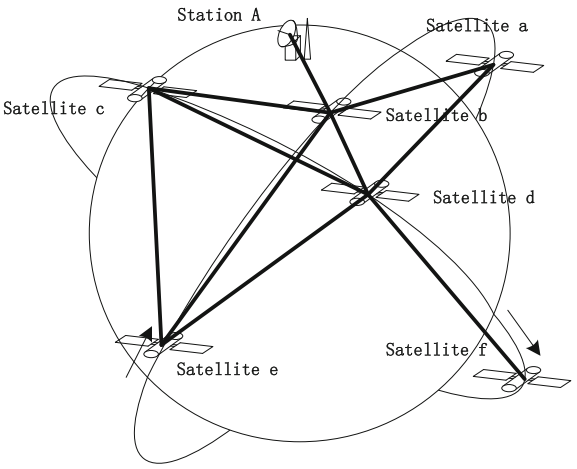


Fig. 2 Node satellite modeling

satellite links in a finite time slot [16]. Therefore we should as far as possible planning more linking target for each satellite and meet the need of traffic for all satellites in one link.

Constraint (4) specifies that the number of established links from a node cannot exceed the total number of ISL's available to the node (i.e., N_t). In this constrain, V_{ij} is the (i,j) element of the visibility matrix and is one if nodes i and j are visible from each other and zero otherwise. Constraint (2) expresses our assumption that if a link from node i to node j is established, then the backward link from node j to node i is also established. The total traffic on link (i,j) (i.e., f_{ij}) is given by Constraint (3). The link capacity constraint, Constraint (4), states that for each established link the total link traffic should not exceed the link capacity C. The traffic conservation constraint, Constraint (5), specifies that for each source–destination pair (m and n), the traffic into a node minus the traffic out of the node balances sourcing or sinking traffic at that node. Constraint (7) states that u_{ij} takes one of two integer values zero and one meaning that a link is either established (i.e., $u_{ij} = 1$) or not (i.e., $u_{ij} = 0$).

$$\sum_{\mathrm{j}} V_{ij} u_{ij} \leq N_t \quad \forall i \tag{4}$$

$$u_{ij} = u_{ji} \quad \forall i,j \tag{5}$$

$$f_{ij} = \sum_{m} \sum_{n} x_{ij}^{mn} \quad \forall i,j \tag{6}$$

$$f_{ij} \leq C * u_{ij} \quad \forall i,j \tag{7}$$

$$\sum_{j} x_{ji}^{mn} - \sum_{j} x_{ij}^{mn} = \begin{cases} -O^{mn} & \text{if } i = m \\ O^{mn} & \text{if } i = n \\ 0 & \text{otherwise} \end{cases} \quad \forall i,m,n \tag{8}$$

$$x_{ij}^{mn} \geq 0 \quad \forall i,j,m,n \tag{9}$$

$$u_{ij} \in \{0,1\} \quad \forall i, \tag{10}$$

N number of nodes (satellites);

N_t maximum number of links at each node (number of ISL's);
C link capacity;

O^{ij} traffic requirement form node i to node j;

V_{ij} visibility between node i and node j [the (i,j) element of the visibility matrix];

u_{ij} link connectivity between node i and node j;

x_{ij}^{mn} carried traffic on link (i,j) due to source-destination pair (m,n);

f_{ij} total carried traffic on link (i,j).

In this paper, we build the ISL modeling in each FTS and get the link assignment table by simulation, according to constraints (4)–(10).

2.4 Build Routing Table

Routing table is collection of paths about ground station through the node to GNSS constellation. The routing need the number of linking as much as possible and the time delay as less as possible, in order to meet the rapid information interaction between ground stations and GNSS constellation. Time delay mainly comes from the number of other satellites between ground station and the target satellite. Considering the complex intersatellite relationship, we need to continuously optimize topological structure, and find out the optimal path in each finite time slot to meet the requirements.

The uplink data need nodes as a data interchange station to GNSS constellation. So that, the linking between satellites is very important. The gather of nodes can be expressed by Eq. (11), and others are expressed by Eq. (12). n is the satellite sum of GNSS constellation, q is the number of invisible satellite.

$$S_{Ni}^{T} = \{S_{N1}, S_{N2}, \ldots, S_{Nm}\} \tag{11}$$

$$S_{oj}^{T} = \{S_{o1}, S_{o2}, \ldots, S_{o(n-m)}\} \tag{12}$$

The gather between node satellites and others can be given by Eq. (13). So that, the number of linking can be given by Eq. (14).

$$L_{No}^{T}(S_{Ni} \rightarrow S_{oj}) \tag{13}$$

$$\sum_{t,i,j} L_{No}^{T}(S_{Ni} \rightarrow S_{oj}) = n - q \tag{14}$$

3 Constraint Condition

Due to the finite time slot of the solution to the problem of dynamic of GNSS constellation network, the size of the finite time slot has a great impact on linking efficiency and routing performance for the whole of GNSS constellation.

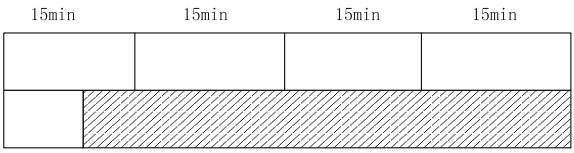


Fig. 3 Example of visible relationship

3.1 Performance Analysis

In this paper, we calculate the 7 days' visible matrix with finite time slot. If it is visible within time slot, we record as 1, else we record as 0.

The smaller intersatellite link slot is, the more accurate link is. As shown in Fig. 3, the visible time of two satellites are shown in figure in shadow, when the finite time slots choose 60 min, the intersatellite status is invisible; when the finite time slots choose 30 min, the intersatellite status is visible in latter 30 min; when the finite time slots choose 15 min, the intersatellite status is visible in latter 45 min. Therefore, choosing a small finite time slot can close to the real state of visible and improve the overall performance of the intersatellite link when we simulation the visible matrix.

3.2 Efficiency Analysis

Building more precise ISL model has not a better impact on the process of ISL. The efficiency of ISL also occupies an important position in the process. We need to improve the efficiency of ISL as possible on the basis of meeting the ISL's performance requirements.

When we choose the smaller finite time slot of ISL, we can guarantee a more accurate model, but the requirement of computation grows exponentially. As shown in Fig. 4, the smaller finite time slot we choose, the higher ISL's model cost, so that it is not conform to the requirements of the ISL when we pursuit the accuracy of GNSS constellation purely. In this paper, we will compare the corresponding performance and efficiency to conclude the optimal finite time slot for each finite time slot in the next section.

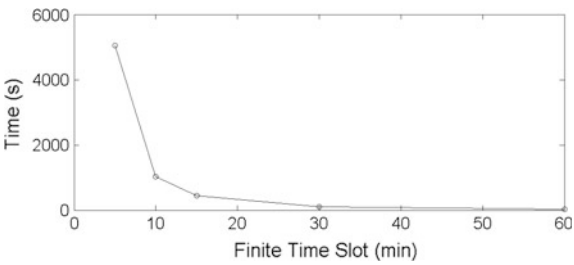


Fig. 4 Relationship between time and FTS

4 ISL Simulation Analysis of Performance

In this paper, we use the satellite simulation toolkit (Satellite Tool Kit, STK) to simulate the model of GNSS constellation. This constellation include 5 GEO satellites, 5 IGSO satellites, 24 MEO satellites, a total of 34 satellites and 5 ground stations. We calculate the intersatellite and satellite-ground visible matrixes with STK and MATLAB, divide those matrix with the finite time slot, and build the visible matrixes of 60, 30, 15, 10 and 5 min.

4.1 Choice of Node Satellite

The size of finite time slot has important influence on computing result of visible radian, which lead to the difference of visible matrix. In the aspect of choosing the nodes, the statistics of visible satellites' number can be used as the choosing the nodes for each ground station. In other word, we count the number of the least visible satellites and the average number of visible satellites for each ground station in each finite time slot.

Clearly, the bigger finite time slot we choose, the less average number of visible satellites we get, but the degree of number reduced is more and more small, and the least number of visible satellites get the same result. The result is shown as Table 1.

4.2 Comparison of Link Assignments

In the ISL, the number between satellites has an important influence on autonomous operation and navigation precision. The more links satellites can built, the better ISL's performance is. In this paper, we count the condition of linking in each finite time slot. The results are shown in Table 2.

Table 1 The number of visible satellites in different FTS

Finite time slot (min)	The least number of visible satellites	The average number of visible satellites
60	13	17.7083
30	13	17.7125
15	13	17.7133
10	13	17.7160
5	13	17.7164

Table 2 The number of linking in different FTS

Finite time slot (min)	The least number of linking	The average number of linking
60	10	11.63
30	11	12.93
15	12	13.82
10	12	13.83
5	12	13.86

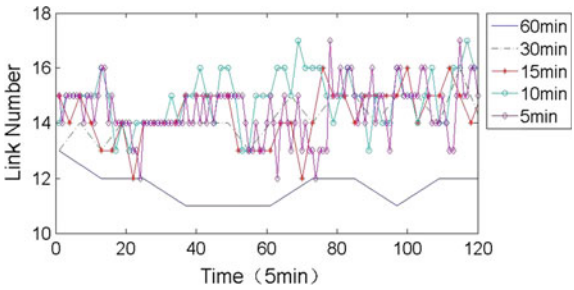


Fig. 5 Relationship between time and FTS

It is obvious from Table 2, the average number of intersatellite' links will increase slightly with the decrease of finite time slot, but when we choose 15, 10 and 5 min, the links have a little change.

As shown in Fig. 5, it is the changes of links average number in each finite time slot from 1 to 10 h. It can be concluded from Fig. 5 that links' number is significantly less than other options when we choose 60 min and the links' number is approximately equivalent when we choose 30, 15, 10 and 5 min.

4.3 Comparison of Routing Tables

The size of finite time slot we choose has an effect on visible matrix, which lead to some links can't be built. In this paper, we count the average and the maximum delay from ground to GNSS constellation in each finite time slot. The results are shown in Table 3.

We can see that there are two satellites when we inject data into GNSS constellation from ground station. The smaller time slot we choose, the smaller average delay is.

In this paper, we count the average delay in each finite time slot from 1 to 10 h. The results are shown in Fig. 6, we can get that the delay is large when we choose 60 min and the delay gap is small when we choose other slots.

Table 3 The delay in different FTS

Finite time slot (min)	Average delay	Maximum delay
60	2.4675	3
30	2.4091	3
15	2.3895	3
10	2.3721	3
5	2.3605	3

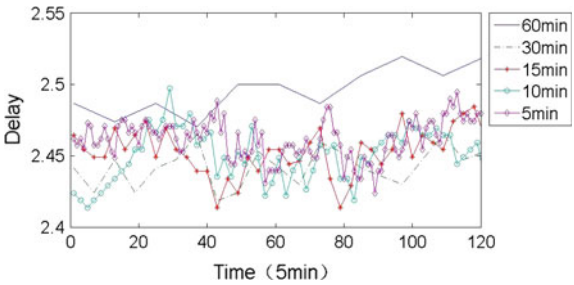


Fig. 6 Relationship between time and FTS

5 Conclusion

From the simulation experiment, we can summarize that the performance of ISL would be better when FTS become smaller, but the performance improve more and more slowly. As comparison, the computational efficiency would be exponential growth. When we choose the 60 min, the performance is significantly lower than others. When we choose 60, 30, 15 and 10 min, the consuming time can be accepted. In summary, the optimal choices of FTS are 30, 15 and 10 min. If the computing power allowed, the best choice of FTS is 10 min.

References

1. Zhang G-X (2009) Application of modern small satellite. Posts and Telecom Press, Beijing
2. Maine K, Anderson P, Bayuk F (2004) Communication architecture for GPS III. In: IEEE aerospace conference proceedings, Big Sky, USA, March 6–13, 2004
3. Zhang Y, Zhang Y (2005) The autonomous navigation constellation base on intersatellite observation. NUDT, Changsha, pp 19–23
4. Tang Y, Wang Y, Chen J (2016) The availability of space service for inter-satellite links in navigation constellations. Sensors 16(8):1327
5. Xu B, Wang D, Liu W et al (2015) A hybrid navigation constellation inter-satellite link assignment algorithm for the integrated optimization of the inter-satellite observing and communication performance. CSNC, pp 283–296
6. Yang Yu (2013) Research on ISLs allocation method in GNSS. Hunan University, Changsha
7. Zhendong Li, Shanbao He (2011) A topology design method of navigation satellite constellate link. Space Eng 27(9):32–37

8. She CD, Wang JF (2006) Topological dynamics analysis of walker-constellation satellite networks. J Commun 27(8):45–51
9. Chen X, Han C (2011) Reconfiguration of communications constellation with ISLs: fix the net. In: 2011 international conference on consumer electronics, communications and networks (CECNet), April 16, 2011
10. Harathi K, Krishna P, Newman-Wolfe RE et al (1993) A fast link assignment algorithm for satellite communication networks. In: Computers and communications, twelfth annual international phoenix conference, Tempe, USA, March 23–26, 1993
11. Noakes MD, Cain JB, Nieto JW et al (1993) An adaptive link assignment algorithm for dynamically changing topologies. IEEE Trans Commun 41(5):694–706
12. Hong SC, Byoung WK, Chang GL et al (1998) FSA-based link assignment and routing in low-earth orbit satellite networks. IEEE Trans Veh Technol 47(3):1037–1048
13. Cain JB, Adams SL, Noakes MD et al (1987) A distributed link assignment (reconstitution) algorithm for space-based SDI networks. In: Military communications conference-crisis communications: the promise and reality, MILCOM 1987, IEEE, Washington, DC, USA, October 19–22
14. Li J, Hu H, Yu PJ, Zhu J (2012) Analysis of requirement and constrained model of inter-satellite-link TT&C scheduling problem on navigation constellation. Artif Intell Comput Intell 7530:160–169
15. Wang X, Meng Y, Su Z et al (2014) A simple three-dimensional matrix method for global constellation intrasatellite link topological design. Int J Navig Obs 2014:1–17
16. Jing L, Gang-qiang YE, Tian-jiao Z, Pei-jun Y (2014) The integrated management method of slow varying ISL's for MEO satellite network. Institute of electrical and electronics engineers Inc. vol 1, pp 907–911

Construction and Performance Analysis of GPS/BeiDou/Galileo Real-Time Augmentation System

Liang Chen, Ying Liu, Changjiang Geng, Maorong Ge and Jiao Wenhai

Abstract The satellite network of Chinese Beidou system and the European Galileo System are on-contraction recently, which promotes the development of Multi-GNSS high-precision fusion applications. In this paper, the algorithm of multi-GNSS real-time clock fusion estimation is studied deeply, the real-time augmentation message processing strategy is also discussed and real-time orbit data processing and analysis are carried out based on PANDA software. Then the multi-GNSS real-time augmentation system prototype is constructed based on the GNSS real-time stream which could be obtained currently and real-time augmentation messages including GPS, BeiDou and Galileo are broadcasting on the Internet. The experiment results show that the real-time augmentation message signal-in-space ranging error(RTAM-SISRE), effecting the user's actual positioning performance, of GPS is about 4–7 cm, that of Beidou IGSO/MEO and Galileo is about 10 cm and BeiDou GEO satellites' is about 33 cm, respectively. The real-time Precise Point Positioning (PPP) results prove that 5–7 cm horizontal RMS and better-than 10 cm vertical RMS based on carrier-phase observation can be obtained.

Keywords GPS · BeiDou · Galileo · Real-time augmentation system · Signal-in-space ranging error · Real-time PPP

L. Chen (✉) · C. Geng
GNSS Engineering Center, China Academy of Aerospace Electronics Technology,
1 Fengyingdong Road, Beijing 100094, China
e-mail: sdkdchenliang@163.com

L. Chen · C. Geng
Test and Assessment Research Center, China Satellite Navigation Office,
1 Fengyingdong Road, Beijing 100094, China

M. Ge
Germany Research Center for Geosciences (GFZ), Telegrafenberg,
14473 Potsdam, Germany

Y. Liu · J. Wenhai
China Beijing Institute of Tracking and Telecommunication Technology,
Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_14

1 Introduction

Since the establishment of GPS in 1980s, the construction and application of Global Navigation Satellite System (GNSS) are developing rapidly. GPS and GLONASS are currently operating at full constellation and capability and can provide about 10 m navigation and positioning services for users of all over the world. Chinese BeiDou system and the European Galileo system are also on-construction recently. With the deepening application of satellite navigation in various industries, the demands in navigation positioning accuracy and performance become more and more high, which also promote the progress of satellite navigation and positioning technology significantly.

Up to the end of 2016, the Chinese government has launched 23 satellite BeiDou II satellites, 14 in-orbit satellites of which have been providing stable and reliable Navigation, Positioning and Time (PNT) service for the Asia-Pacific region [1]. There are 14 Galileo satellites in service including 4 In-Orbit Validation (IOV) satellites and 10 Full Operational Capability (FOC) satellites currently. The BeiDou and Galileo satellite's information are shown in Table 1. The satellite network construction of Beidou and the Galileo promotes the development of Multi-GNSS fusion applications and the revolution of Multi-GNSS high-precision fusion technology. In recent years, Multi-GNSS Precision Orbit Determination (POD) technology has been developing rapidly and the precision of POD has been also improved significantly [2–7], which provides the necessary conditions for GNSS augmentation technology based on improving GNSS Signal-In-Space performance by providing higher precision ephemeris products comparing broadcast ephemeris.

GNSS real-time augmentation relies on high-precision real-time satellite orbit and real-time clock products. At present, the International GNSS Service (IGS) has

Table 1 BeiDou and Galileo satellite's information (Up to 2016)

BeiDou	PRN	Satellite name	Galileo	PRN	Satellite name
GEO	C01	GEO-1	IOV	E11	IOV-1
	C02	GEO-6		E12	IOV-2
	C03	GEO-3		E19	IOV-3
	C04	GEO-4		E20	IOV-4
	C05	GEO-5	FOC	E01	FOC-10
IGSO	C06	IGSO-1		E02	FOC-11
	C07	IGSO-2		E08	FOC-8
	C08	IGSO-3		E09	FOC-9
	C09	IGSO-4		E14	FOC-2
	C10	IGSO-5		E18	FOC-1
	C15[a]	IGSO-6		E22	FOC-4
MEO	C11	MEO-3		E24	FOC-5
	C12	MEO-4		E26	FOC-3
	C14	MEO-6		E30	FOC-6

[a]PRN switch from C15 to C13 on 2016/10/11

provided precise satellite orbit and clock product. Even though the effectiveness, reliability and accuracy of ultra-rapid orbit provided by IGS have met the real-time requirements, its ultra-rapid clock, with a precision of about 3 ns, cannot be used in high-precise real-time applications [8]. IGS began to carry out the real-time project at 2002. Until now it has provided 0.3–0.8 ns GPS/GLONASS real-time precise clock products [9], however, which are only for scientific research and difficult to be guaranteed for users. In order to obtain the high-precise real-time clock products, the high-precise real-time orbit products are very critical. Recently, there are lots of scholars paying attention to satellite POD and acquiring many in-depth achievements all over the world. Real-time GNSS orbit estimation, the GPS 3-dimensional precision is about 3.5 cm, is realized preliminarily by Kalman filter in paper [2]; Based on PANDA software, the solar pressure model and the yaw-attitude model of Beidou satellite have been greatly studied in paper [3–6], which improve the BeiDou GEO, IGSO and MEO satellite orbit precision significantly. The precision of better than 10 cm orbit overlaps in radial-direction and better than 20 cm in cross-direction is achieved in those papers. In paper [7] the SLR and GNSS data are been used together in Galileo satellites POD in order to eliminate the systematic deviation existing in POD using GNSS observations only and the orbit precision is about ten centimeters. The traditional satellite clock estimation model is optimized and improved in order to realize the real-time clock resolved high-efficiently in paper [10–15], which can be used to updating real-time clock every second and also can realize 0.15 ns real-time clock in experiment afterwards.

This article is organized as follows. Firstly, the multi-GNSS real-time clock fusion estimation algorithm is introduced, then the data processing strategy of real-time augmentation messages is discussed and real-time orbit data processing and analysis is carried out based on PANDA software. On this basis, the multi-GNSS real-time augmentation system is constructed based on the GNSS real-time stream and real-time augmentation messages including GPS, BeiDou and Galileo are broadcasting on the Internet. Based on the augmentation system, the performance of the real-time clock estimation model and real-time augmentation message Signal-In-Space Ranging Error (SISRE) are analyzed. Finally, the real-time PPP is used to verify the GNSS augmentation precision.

2 Real-Time Augmentation System

2.1 Real-Time Orbit and Clock

In this paper, real-time orbit data processing is based on Position and Navigation Data Analyst (PANDA) software, which has achieved GPS, GLONASS, BeiDou and Galileo high-precision orbit determination [16, 17]. The POD strategy adopted in this paper is that previous 48 h observations are used for once POD processing, carried out every three hours, to obtain the real-time orbit (6 h predicted part).

Based on real-time GNSS data streams, high-precise GNSS real-time clock processing is carried out by fixing high-precise information generated in real-time POD processing like satellite real-time orbit, earth rotation parameters and station coordinates and so on. Ionosphere-free combination observations are used in dual-frequency data processing commonly to eliminate the ionosphere delay. If a unified time reference is chosen, the un-differenced (UD) observation equation of GPS/BeiDou/Galileo ionosphere-free combination is as follows:

$$\begin{cases} v^G_{j,PC} = dt + b^G_{r,PC} - dT^G_j - b^{G,s}_{j,PC} + d^G_{j,trop} + l^G_{j,PC} \\ v^G_{j,LC} = dt + b^G_{r,LC} - dT^G_j - b^{G,s}_{j,LC} + N^G_j + d^G_{j,trop} + l^G_{j,LC} \\ v^B_{k,PC} = dt + b^B_{r,PC} - dT^B_k - b^{B,s}_{k,PC} + d^B_{k,trop} + l^B_{k,PC} \\ v^B_{k,LC} = dt + b^B_{r,LC} - dT^B_k - b^{B,s}_{k,LC} + N^B_k + d^B_{k,trop} + l^B_{k,LC} \\ v^E_{m,PC} = dt + b^E_{r,PC} - dT^E_m - b^{E,s}_{m,PC} + d^E_{m,trop} + l^E_{m,PC} \\ v^E_{m,LC} = dt + b^E_{r,LC} - dT^E_m - b^{E,s}_{m,LC} + N^E_m + d^E_{m,trop} + l^E_{m,LC} \\ \cdots \end{cases} \tag{1}$$

where G, B and E represent one GPS, Beidou and Galileo satellites respectively; j, k and m are the jth GPS, the kth BeiDou, and mth Galileo satellite at the same epoch; PC, LC and N are pseudo-range, carrier-phase and ambiguity of ionosphere-free; v is the residual; dt and dT are receiver and satellite clock error; d_{trop} is the tropospheric delay; l is the difference between satellite station geometric distance and PC, LC; b_r and b^s represent the signal delay bias in receiver and satellite, respectively. The bias related with pseudo-range is called code bias and another related with carrier-phase is called un-calibrated phase delay.

In this equation, the code bias b^s_{PC} satellite-related is stable [18], and can be merged with satellite clock error, we set $d\tilde{T} = dT + b^s_{PC}$; Considering that the initial information in normal equation is provided by pseudo-range observation mostly and GPS time is chosen as the time reference, $b^G_{r,PC}$, the GPS code bias in receiver, and dt, receiver clock error, are merged like $d\tilde{t} = dt + b^G_{r,PC}$, then the equation can be simplified as follow:

$$\begin{cases} v^G_{j,PC} = d\tilde{t} - d\tilde{T}^G_j + d^G_{j,trop} + l^G_{j,PC} \\ v^G_{j,LC} = d\tilde{t} + \left(b^G_{r,LC} - b^G_{r,PC}\right) - d\tilde{T}^G_j + \left(b^{G,s}_{j,PC} - b^{G,s}_{j,LC} + N^G_j\right) + d^G_{j,trop} + l^G_{j,LC} \\ v^B_{k,PC} = d\tilde{t} + \left(b^B_{r,PC} - b^G_{r,PC}\right) - d\tilde{T}^B_k + d^B_{k,trop} + l^B_{k,PC} \\ v^B_{k,LC} = d\tilde{t} + \left(b^B_{r,LC} - b^G_{r,PC}\right) - d\tilde{T}^B_k + \left(b^{B,s}_{j,PC} - b^{B,s}_{j,LC} + N^B_k\right) + d^B_{k,trop} + l^B_{k,LC} \\ v^E_{m,PC} = d\tilde{t} + \left(b^E_{r,PC} - b^G_{r,PC}\right) - d\tilde{T}^E_m + d^E_{m,trop} + l^E_{m,PC} \\ v^E_{m,LC} = d\tilde{t} + \left(b^E_{r,LC} - b^G_{r,PC}\right) - d\tilde{T}^E_m + \left(b^{E,s}_{j,PC} - b^{E,s}_{j,LC} + N^E_m\right) + d^E_{m,trop} + l^E_{m,LC} \\ \cdots \end{cases} \tag{2}$$

If set $ISB^{B-G} = b^{B}_{r,PC} - b^{G}_{r,PC}$ and $ISB^{E-G} = b^{E}_{r,PC} - b^{G}_{r,PC}$, then the equation above can be expressed as,

$$\begin{cases} v^{G}_{j,PC} = d\tilde{t} - d\tilde{T}^{G}_{j} + d^{G}_{j,trop} + l^{G}_{j,PC} \\ v^{G}_{j,LC} = d\tilde{t} - d\tilde{T}^{G}_{j} + \left(b^{G}_{r,LC} - b^{G}_{r,PC} + b^{G,s}_{j,PC} - b^{G,s}_{j,LC} + N^{G}_{j}\right) + d^{G}_{j,trop} + l^{G}_{j,LC} \\ v^{B}_{k,PC} = d\tilde{t} + ISB^{B-G} - d\tilde{T}^{B}_{k} + d^{B}_{k,trop} + l^{B}_{k,PC} \\ v^{B}_{k,LC} = d\tilde{t} + ISB^{B-G} - d\tilde{T}^{B}_{k} + \left(b^{B}_{r,LC} - b^{B}_{r,PC} + b^{B,s}_{j,PC} - b^{B,s}_{j,LC} + N^{B}_{k}\right) + d^{B}_{k,trop} + l^{B}_{k,LC} \\ v^{E}_{m,PC} = d\tilde{t} + ISB^{E-G} - d\tilde{T}^{E}_{m} + d^{E}_{m,trop} + l^{E}_{m,PC} \\ v^{E}_{m,LC} = d\tilde{t} + ISB^{E-G} - d\tilde{T}^{E}_{m} + \left(b^{E}_{r,LC} - b^{E}_{r,PC} + b^{E,s}_{j,PC} - b^{E,s}_{j,LC} + N^{E}_{m}\right) + d^{E}_{m,trop} + l^{E}_{m,LC} \\ \ldots\ldots \end{cases} \tag{3}$$

If set $\tilde{N} = b_{r,LC} - b_{r,PC} + b^{s}_{PC} - b^{s}_{LC} + N$, the parameters estimated in the UD observation equation are as follow:

$$\left(d\tilde{t}, d\tilde{T}, ISB^{B-G}, ISB^{E-G}, \tilde{N}\right)^{T}$$

2.2 Processing Strategy

The processing strategy and parameter model of real-time orbit and clock are as follows:

2.3 System Construction

The GPS/BeiDou/Galileo real-time augmentation system is designed and the system prototype is constructed. As shown in Fig. 1, the main processes of the system prototype include obtaining the real-time observation data of global multi-GNSS tracking network, generating the real-time orbit and real-time clock augmentation messages and broadcasting GPS/BeiDou/Galileo augmentation messages.

In the system, a global multi-GNSS network needs to be built to acquire the real-time observation data and transmit to the data processing center through the Internet.

The real-time POD strategy is that once POD processing carries out every three hours using previous 48 h observations to obtain the real-time orbit (6 h predicted part). Based on the IGS core stations and M-GEX stations (http://www.igs.org/network), about 80 multi-GNSS stations shown in Fig. 2 are chosen for POD, about 40 of which can track BeiDou signal and 60 can capture Galileo signal.

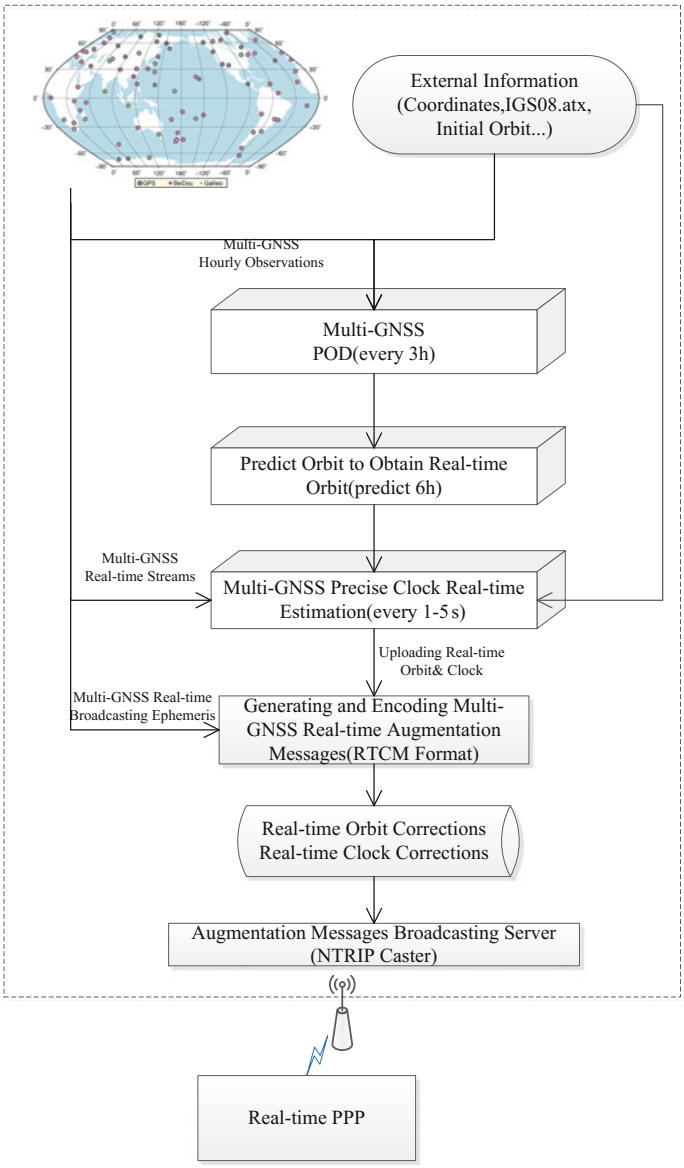


Fig. 1 The frame of GPS/BeiDou/Galileo real-time augmentation system

Using the data processing strategies in Table 2, multi-GNSS Real-time clock processing platform is constructed based on multi-GNSS real-time orbit products above and real-time data streams. As shown in Fig. 3, the total number of real-time stations, some of which come from IGS real-time streams transmitted by German

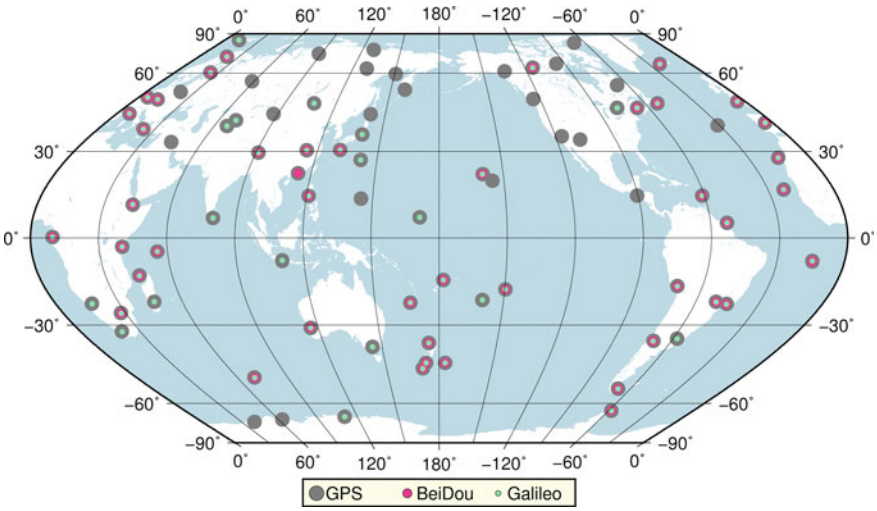


Fig. 2 The stations distribution of GPS/BeiDou/Galileo real-time augmentation system

Geoscience Research Center (GFZ) and provided by German Federal Agency for Cartography and Geodesy (BKG) (http://mgex.igs-ip.net) [23] and some are built in China, is about 60 including about 35 BeiDou stations and 50 Galileo Stations.

The messages broadcasting platform will process the multi-GNSS real-time augmentation message and GNSS broadcast ephemeris streams to generate GPS, BeiDou and Galileo real-time augmentation corrections, which are encoded in RTCM format and broadcasted in NTRIP protocol [24].

It should be noted that the messages broadcasting in the system prototype is realized based on BKG NTRIP Client (BNC) and NtripCaster software provided by BKG [25]. The following tests are carried out based on the prototype system.

3 Precision of Real-Time Augmentation Message

The users of multi-GNSS real-time augmentation need to combine real-time orbit and clock in positioning, so they take more attention to the combination precision of orbit and clock. The real-time orbit and clock in real-time augmentation messages are self-consistence, the clock products can eliminate some offsets of the orbit error in radial-direction partially. Therefore, the factor affecting the actual positioning performance is the real-time augmentation message Signal-in-Space ranging error (SISRE), which is the projected error of the difference between the satellite position obtained by the real-time augmentation messages(orbit and clock) and the truth

Table 2 The processing strategy and parameter model of real-time orbit and clock

Parameters		Models for POD	Models for real-time clock estimation
Observation information			
Observation		UD ionosphere-free observation	UD ionosphere-free observation
Prior information		P1:1.0 m; L1:0.02cycle	P1:1.0 m; L1:0.02cycle
Elevation mask		7°	7°
Observation weight		p = 1, elev > 30° p = 2sin(elev), elev $\leq$ 30°	p = 1, elev > 30° p = 2sin(elev), elev $\leq$ 30°
Simple		300 s	Real-time 5 s
Calculating arcs		48 h	–
Reference frame			
Time system		GPS Time	GPS Time
Inertial frame		ICRF J2000.0	ICRF J2000.0
Terrestrial frame		ITRF2008	ITRF2008
Precession and nutation		IAU2000	IAU2000
EOP parameters		IERS C04	IERS C04
Sun&Moon ephemeris		JPL DE405	JPL DE405
Orbit model			
Geopotential model		EGM96 model (8 × 8)	–
Ocean tides		IERS conventions 2003	IERS conventions 2003
Solid earth tides		IERS conventions 2003	IERS conventions 2003
Solid earth pole tides		IERS conventions 2003	IERS conventions 2003
Solar radiation pressure model		Reduced CODE 5-parameter with no initial value [19]	–
Correction			
Phase rotation correction		Model correct [20]	Model correct
PCO/PCV	Satellite	GPS, Galileo PCO: IGS08; BeiDou GEO PCO:IGS M-GEX; BeiDou IGSO/MEO PCO: ESA Model [21]; GPS PCV: IGS08; BeiDou, Galileo PCV: uncorrected	GPS, Galileo PCO: IGS08 BeiDou GEO PCO:IGS M-GEX BeiDou IGSO/MEO PCO:ESA model [21] GPS PCV: IGS08 BeiDou, Galileo PCV: uncorrected
	Receiver	GPS PCO: IGS08; BeiDou, Galileo PCO: same as GPS; GPS PCV:IGS08; BeiDou, Galileo PCV: same as GPS;	GPS PCO: IGS08 BeiDou, Galileo PCO:same as GPS GPS PCV:IGS08 BeiDou, Galileo PCV: same as GPS
Relativistic effects		IERS conventions 2003	IERS conventions 2003

(continued)

Table 2 (continued)

Parameters	Models for POD	Models for real-time clock estimation
Parameters estimation		
Adjustment method	Weighted least-squares algorithms	Square-root information filtering [22]
Reference clock	One satellite clock	One satellite clock
Satellite orbit	Position and velocity at given initial epoch as well as five solar radiation parameters	Fixed
Station coordinate	Estimated	Fixed
Tropospheric delay	Saastanmoine Model + GMF mapping function Piece-wise constant zenith delay of each station per 2 h	Saastanmoine Model + GMF mapping function random-walk process for each epoch
Satellite clock	Estimated as white noise	Estimated as white noise
Receiver clock	Estimated as white noise	Estimated as white noise
Ambiguity	Double-difference AR	Estimated
Inter-system bias	Estimated as constant parameters with zero-mean condition	Estimated as constant parameters with zero-mean condition

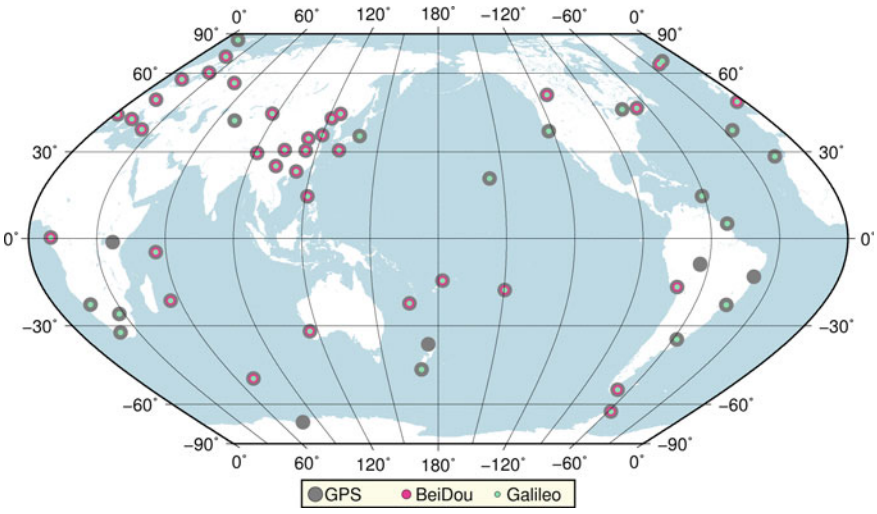


Fig. 3 The distribution of GPS/BeiDou/Galileo real-time clock estimation stations

Table 3 Coefficients of different GNSS SISRE

Satellite	BeiDou		Galileo	GPS
	GEO/IGSO	MEO		
Orbit altitude (km)	35,786	21,528	23,222	20,200
α	0.99	0.98	0.98	0.98
β	127	54	61	49

position on the line-of-sight direction from satellite to the user position. The paper [26] gives the SISRE derivation process and obtains the formula as follows:

$$SISRE = \sqrt{(\alpha \cdot dR - c \cdot dt)^2 + (dA^2 + dC^2)/\beta} \tag{4}$$

where dR, dA, dC represent the real-time orbit error in radial-direction, along-direction and cross-direction compared with the reference orbit in inertial frame, dt is the real-time clock error, and c is the speed of light. α and β are correlation coefficient determined by the satellite orbit altitude, the different GNSS coefficients are shown in Table 3. From Eq. (4) and the coefficients in the table, it can be found that main factors affecting the user positioning performance are the real-time orbit error in radial-direction and clock error.

Based on the same environment (same real-time orbit, real-time streams…) of the real-time augmentation system prototype, two servers were chosen to carry out the comparative test between UD and simplified model, and the results were analyzed from DOY 230 to 237, 2016.

3.1 Precision of Real-Time Orbit

In order to verify the precision of real-time orbit, the GBM final products provided by GFZ are selected as reference orbit. Compared to the IGS final products, the Precision of GBM products orbit is better than 1 cm and clock is better than 0.02 ns and BeiDou orbit precision compared to SLR is about 10 cm [27]. Figure 4 and Table 4 show the real-time orbit precision obtained in real-time augmentation system prototype compared to the GBM orbit. From the results, it can be found that the GEO orbit precision in radial-direction is about 30 cm and the along-direction and cross-direction are worse, the reason of which may be that there is no obvious changing of the GEO satellite geometrical distance to GNSS stations. The precision of BeiDou IGSO, MEO and Galileo satellites in radial-direction is about 10, 4.5 and 11.5 cm, respectively. Among those, the precision of GPS is the highest with about radial 2 cm, three-dimensional 7.5 cm.

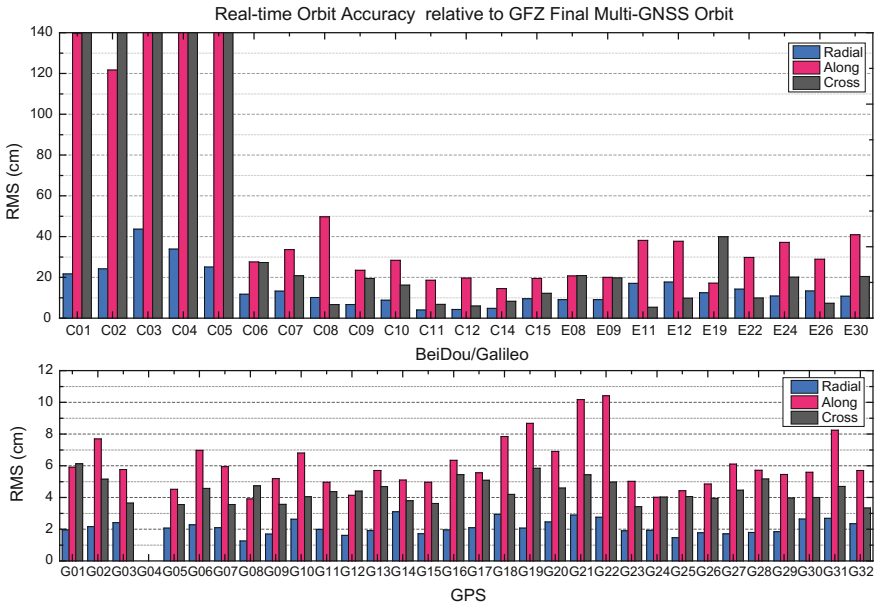


Fig. 4 The precision of real-time orbit compared to the GBM

Table 4 The RMS precision statistics of real-time orbit (cm)

Satellite type	Radial	Along	Cross
BDS-GEO	29.69	197.97	321.02
BDS-IGSO	10.01	30.35	17.09
BDS-MEO	4.34	17.60	7.01
Galileo	11.50	30.05	17.05
GPS	2.13	6.08	4.40

3.2 Precision of Real-Time Clock

The precision of real-time clock is analyzed by quadratic-difference method [12–15], namely, by one satellite clock minus itself reference satellite, we can obtain two clock difference time sequences of real-time clock and GBM clock, make second-difference between the two clock sequences then statistic the mean and STD. Figure 5 shows the precision comparison of the BeiDou, Galileo and GPS real-time clock products to the GBM clock products (C06, E08 and G01 are chosen as the reference satellites, respectively). From the results, we can see that the precision of GPS, BeiDou GEO, IGSO/MEO and Galileo estimated by UD model is about 0.18, 0.50, 0.24 and 0.30 ns, respectively. It can be also found that the

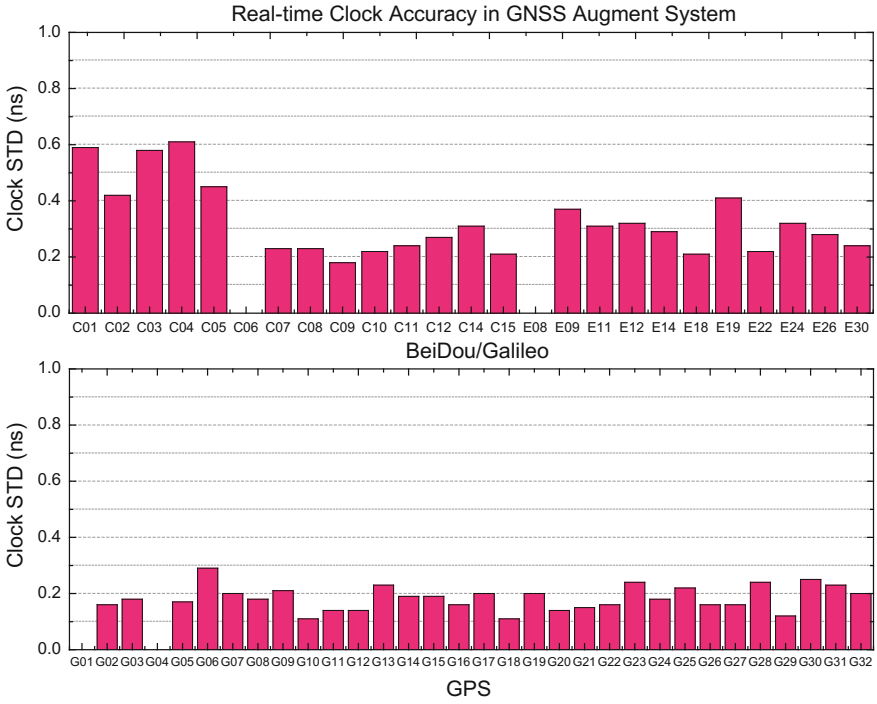


Fig. 5 Precision comparison of BeiDou, Galileo and GPS real-time clock to GBM

precision of BeiDou GEO clock is worse than the others due to the impact of orbit precision.

3.3 Real-Time Augmentation Message SISRE

By Eq. (4), the real-time augmentation message SISRE provided by the real-time augmentation system prototype can be obtained, and the results are shown in Fig. 6. From the figure, we can get that the real-time augmentation message SISRE of GPS, BeiDou GEO, IGSO/MEO and Galileo is about 4.8, 33, 10 and 8 cm, respectively.

From Eq. (4), it can be found that the orbit error in radial-direction and clock error affect SISRE mostly, and the clock error can partially eliminate some effects produced by the orbit error in radial-direction on SISRE. However, the SISRE

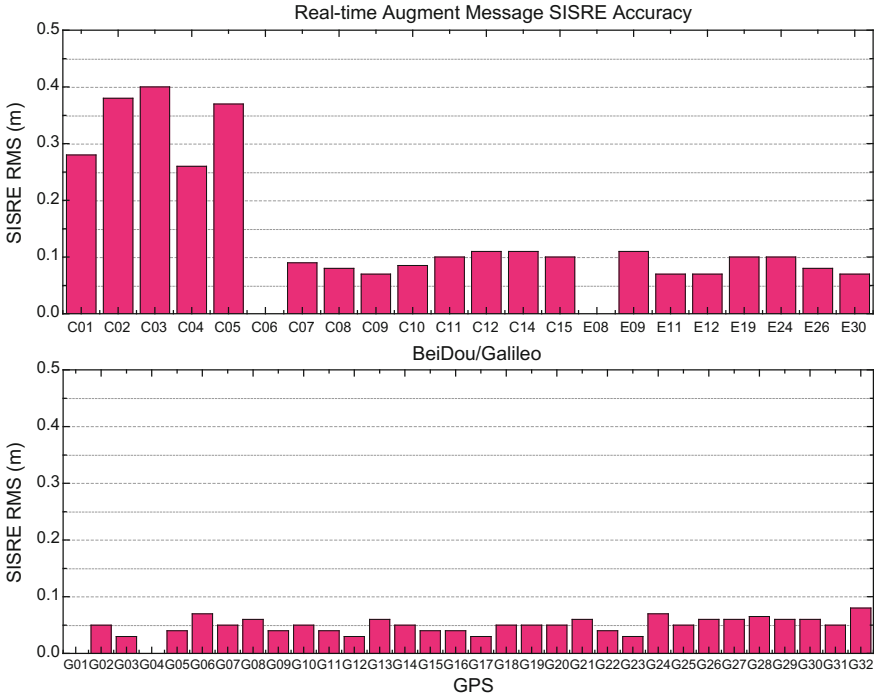


Fig. 6 Comparison of real-time augmentation message SISRE of BeiDou, Galileo and GPS

precision of BeiDou GEO real-time augmentation messages is much larger than others, so it is not suitable for real-time high-precision positioning. Therefore, BeiDou GEO satellites are not chosen in the following experiments.

4 Real-Time PPP Performance Analysis

Based on the real-time precise positioning software developed independently, the performance of GNSS real-time augmentation system is tested using five real-time stations provided by BKG, which are located in the area of Asia-Pacific and North America.

Figure 7 shows the 24-h results of GPS + BeiDou real-time PPP in UNX3 on November 10, 2016 (DOY315). From the figure, it can be found that multi-GNSS real-time PPP can converge after about 15–20 min. Table 5 gives the precision statistic results of November 4, 2016 to November 10, 2016, from which we can

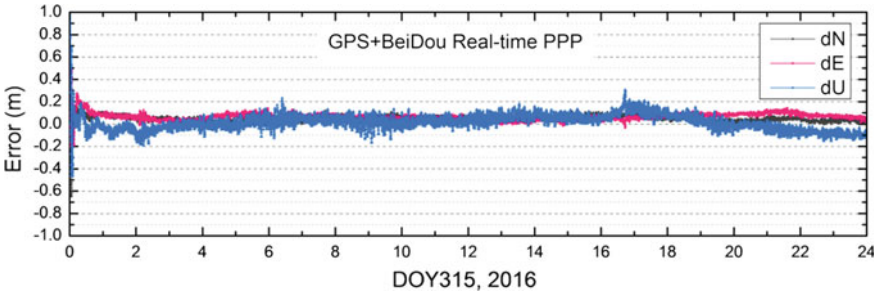


Fig. 7 Comparison between multi-GNSS PPP and GPS PPP using augmentation messages in UNX3

Table 5 Real-time PPP RMS statistics (m)

Station	Latitude	Longitude	Nation	Resolving model	Real-time PPP based on augmentation messages	
					Horizontal	Vertical
ALGO	45.96	−78.07	Canada	GPS	0.038	0.070
CUT0	−32.00	115.89	Australia	GPS + BeiDou + Galileo	0.065	0.071
GMSD	30.56	131.02	Japan	GPS + BeiDou + Galileo	0.047	0.057
JFNG	30.52	114.49	China	GPS + BeiDou + Galileo	0.050	0.064
UNX3	−33.92	151.23	Australia	GPS + BeiDou	0.048	0.054
Average					0.050	0.063

found that the GNSS global real-time augmentation system prototype can provide centimeter-level services with horizontal RMS 5–7 cm, vertical RMS better than 10 cm.

5 Conclusions

GNSS augmentation technology can improve GNSS Signal-In-Space service performance by providing higher precision ephemeris products comparing GNSS broadcast ephemeris. In this paper, the multi-GNSS real-time clock fusion estimation algorithm is studied deeply. Secondly, the data processing strategy of real-time augmentation messages is discussed and real-time orbit data processing and analysis is carried out based on PANDA software. On this basis, the multi-GNSS real-time augmentation system is constructed, and the performance and precision are analyzed. The conclusions are as follow:

1. The experiment results show that overlaps precision of GPS/Beidou MEO/Galileo satellites in radial direction is 1–5 cm, and that of Beidou

GEO/IGSO satellites is about 10 cm in the multi-GNSS real-time augmentation system prototype.
2. Based on augmentation system, real-time clock precision of GPS, BeiDou GEO, IGSO/MEO and Galileo is about 0.18, 0.50, 0.24 and 0.30 ns, respectively.
3. The SISRE of GPS real-time augmentation messages provided by the system prototype is about 4–7 cm, that of BeiDou IGSO/MEO and Galileo is about 10 cm and that of Beidou GEO is about 33 cm, which affect the user's practical application performance;
4. The GNSS global real-time augmentation system prototype can provide centimeter-level services based on the carrier-phase observation with horizontal RMS 5–7 cm, vertical RMS better than 10 cm.

From the conclusions of this paper, the GNSS real-time augmentation system can provide real-time, high-precision and all-weather precise positioning and navigation services all over the world. Combining with other GNSS technologies like RTK, it can be applied to lane meter-level navigation, traffic logistics and transportation, precision agriculture and ocean-going operations and has a wide application prospect.

Acknowledgements Thanks to the multi-GNSS high-precision products, real-time streams and data processing hardware environment provided by German Geosciences Research Center (GFZ). This paper is supported by the China Scholarship Fund ([2015]-5138) of China scholarship Council (CSC).

References

1. Yang Y (2010) Progress, contribution and challenge of compass/BeiDou satellite navigation system. Acta Geodaetica et CartograpgicaSinica 39(1):1–6
2. Laurichesse D (2013) Real time precise GPS constellation and clocks estimation by means of a Kalman filter. In: ION GNSS 2013, 16–20 Sept 2013, Nashville, Tennessee
3. Ge M, Zhang HP, Jia XL, Song SL, Wickert J (2012) What is achievable with the current COMPASS constellation? GPS World pp 29–34
4. Zhao Q, Guo J, Li M et al (2013) Initial results of precise orbit and clock determination for COMPASS navigation satellite system. J Geodesy. doi:10.1007/s00190-013-0622-7
5. Guo J, Chen G, Zhao Q et al (2016) Comparison of solar radiation pressure models for BDS IGSO and MEO satellites with emphasis on improving orbit quality. GPS Solutions. doi:10.1007/s10291-016-0540-2
6. Dai X, Ge M, Lou Y et al (2015) Estimating the yaw-attitude of BDS IGSO and MEO satellites. J Geodesy 89:1005. doi:10.1007/s00190-015-0829-x
7. Hackel S, Steigenberger P, Hugentobler U, Uhlemann M, Montenbruck O (2015) Galileo orbit determination using combined GNSS and SLR observations. GPS Solutions 19:15. doi:10.1007/s10291-013-0361-5
8. International GNSS Service(IGS).Products [EB/OL]. http://www.igs.org/products
9. Hadas T, Bosy J (2015) IGS RTS precise orbits and clocks verification and quality degradation over time. GPS Solutions 19:93–105. doi:10.1007/s10291-014-0369-5
10. Bock H, Dach R, Jaggi A et al (2009) High-rate GPS clock corrections from CODE: support of 1 Hz applications. J Geodesy 83:1083–1094. doi:10.1007/s00190-009-0326-1

11. Zhang X, Li X, Guo F (2011) Satellite clock estimation at 1 Hz for realtime kinematic PPP applications. GPS Solutions 15:315–324. doi:10.1007/s1029-010-0191-7
12. Ge M, Chen J, Dousa J et al (2012) A computationally efficient approach for estimating high-rate satellite clock corrections in realtime. GPS Solutions 16:9–17. doi:10.1007/s10291-011-0206-z
13. Chen L, Geng C, Zhou Q, Jiao W (2015) Estimation strategy and accuracy analysis of GNSS real-time precise satellite clock error. In: Proceedings of China satellite navigation conference (CSNC), vol III. Lecture notes in electrical engineering 342, pp 57–66. doi:10.1007/978-3-662-46632-2_5
14. Liang Chen, Changjiang Geng, Quan Zhou (2016) Estimation model and accuracy analysis of BeiDou/GPS real-time precise satellite clock error integrated resolving. Acta Geodaetica Cartogr Sin 45(9):1028–1034. doi:10.11947/j.AGCS.2016.20150296
15. Zhang W, Lou Y, Gu S et al (2016) Joint estimation of GPS/BDS real-time clocks and initial results. GPS Solutions 20:665–676. doi:10.1007/s10291-015-0476-y
16. Liu J, Ge M (2003) PANDA software and its preliminary result of positioning and orbit determination. Wuhan Univ J Nat Sci 8:603–609. doi:10.1007/BF02899825
17. Shi C, Zhao Q, Ge M (2010) Introduction to PANDA software and the latest development for high precision GNSS data processing and application. IGS analysis center workshop 2010, 28 June–2 July 2010, Newcastle upon Tyne, UK
18. Dach R, Schaer S, Hugentobler U (2006) Combined multi-system GNSS analysis for time and frequency transfer. In: Proceedings of the 20th European frequency and time forum EFTF06, 2006, Braunsehweig, Germany
19. Springer TA, Beutler G, Rothacher M (1999) A new solar radiation pressure model for GPS satellites. GPS Solutions 2:50–62. doi:10.1007/PL00012757
20. Wu JT, Wu SC, Hajj GA, Bertiger WI, Lichten SM (1993) Effects of antenna orientation on GPS carrier phase. Manuscr Geodesy 18:91–98
21. Dilssner F, Springer T, Schonemann E (2014) Werner: estimation of satellite antenna phase center corrections for BeiDou. IGS workshop 2014, Pasadena, USA, 23–27 June 2014
22. Zhao Q, Liu J, Ge M et al (2016) Application of square-root information filtering and smoothing on orbit determination of LEO satellites with on-board GPS Data. Geomatics and Information Science of Wuhan University, 2016(1)
23. Montenbruck O, Steigenberger P, Khachikyan R, Weber G, Langley RB, Mervart L, Hugentobler U (2014) IGS-MGEX: preparing the ground for multi-constellation GNSS science. Inside GNSS 9(1):42–49
24. RTCM Standard 10403.2 (2013) Differential GNSS (Global navigation satellite systems) Services—Version 3 (1 Feb 2013). RTCM paper 104–2013-SC104-STD
25. Stürze A, Mervart L, Weber G et al (2016) The new version 2.12 of BKG ntrip client (BNC). In: Geophysical research abstracts, vol 18, EGU2016-12012
26. 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: Proceedings of China satellite navigation conference (CSNC). Lecture notes in electrical engineering, vol 243. Springer, Berlin, pp 423–434. doi:10.1007/978-3-642-37398-5_39
27. Deng Z, Fritsche M, M Uhlemann etal. Reprocessing of GFZ Multi-GNSS product GBM. IGS Workshop, 08-12 Feb. 2016, Sydney Australia

An Improved Algorithm of Ionospheric Grid Correction Based on GPS and Compass Multi-constellation

Haipeng Li, Rui Li, Ziqi Wang and Weiguang Gao

Abstract Ionosphere delay is one of the main sources of error for single frequency user location and is the main factor influencing the integrity monitoring, which is the major problem in the building of the Beidou SBAS. At present, the Wide Area Augmentation System (WAAS) uses the Kriging method to estimate the Grid Ionospheric Vertical Delay (GIVD) and Grid Ionospheric Vertical Delay Error (GIVE). On the basis of the traditional Kriging method, some domestic scholars have proposed an improved Kriging method, taking into account the temporal correlation of IPP delay values, using the IPP delay values in the update cycle for algorithm to obtain more reasonable GIVE Value. However, because the GPS satellite constellation are all the MEO satellites, IPP observation position change over time, it is not easy to obtain accurate time variogram model. Specially, there are five GEO satellites in China's Beidou system, the IPP position formed by the reference station observing GEO satellites is almost fixed, so we introduce the Beidou GEO satellite observations in this paper, the accurate time variogram model was obtained and the improved Kriging algorithm for the Chinese region is given. In this paper, the ionospheric data of Beijing regional monitoring station are used to study. First, the spatial correlation characteristics of ionospheric delay in northern China were researched. Then, the 5 GEO satellites in Compass are selected to study the ionospheric temporal variation. Finally, we conduct a simulation on ionospheric delay estimation at IGPs in Multi-constellation through Kriging and Improved Kriging algorithm. It can be concluded that, in the premise of satisfying integrity, the GIVE is reduced more than 30% in most parts of north China; therefore, the User Ionospheric Vertical Delay Error (UIVE) can more closely envelop the correction error.

H. Li (✉) · R. Li (✉) · Z. Wang · W. Gao
Beihang University, Beijing, China
e-mail: dxhaipenglee@126.com

R. Li
e-mail: lee_ruin@263.net

R. Li
Collaborative Innovation Center of Geospatial Technology, Wuhan, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_15

Keywords Ionospheric delay · Temporal variation · Multi-constellation · China region

1 Introduction

Ionosphere delay is one of the main sources of error for single frequency user location and is the main factor influencing the integrity monitoring, which is the major problem in the building of the Beidou Satellite Based Augmentation System (SBAS). As a typical SBAS system, the single frequency user of US WAAS uses the Kriging grid correction method to estimate the ionospheric delay and confidence limits in the direction of satellite sight [1]. The fundamental of the grid ionospheric correction method is to make the ionosphere equivalent to a fixed height of the thin layer, the master station using the ground reference station dual-frequency receivers real-time collection of ionospheric delay measurements, estimated coverage area within the latitude and longitude grid point Ionospheric vertical delay correction and correction accuracy, i.e. GIVD and GIVE, moreover correctional data are transmitted on L1 through Geostationary satellites (GEO) to user.

On the basis of the traditional Kriging method, some domestic scholars have proposed an improved Kriging method, taking into account the temporal correlation of IPP delay values, using the IPP delay values in the update cycle for algorithm to obtain more reasonable GIVE Value. However, because GPS satellite constellation are all the MEO satellites, IPP observation position change over time, it is not easy to obtain accurate time variogram model. Currently, the SBASs are all designed for single GPS constellation. With the development of Compass, how to get better navigation accuracy and integrity in SBASs for GPS and Compass constellations is of significance. In this paper, in view of characteristics in Compass which contains GEO, IGSO and MEO three orbital satellites, in particular, there are five GEO satellites which will contribute to some stationary IPPs, so we introduce the Beidou GEO satellite observations in this paper, the accurate time variogram model was obtained and the improved Kriging algorithm for the Chinese region is given. In China, there is large difference between northern and southern in the ionosphere features, in addition, the ionospheric storms and disturbance are more likely to happen in the southern region. The difference is that the ionospheric changes in the northern region are more regular and the change of ionospheric characteristics is close to that of North Americathan. Therefore, the study is carried out using ionospheric observation data received on monitoring stations in Beijing region in this paper. Firstly, as the basis of ionospheric delay modeling, we make a special correlation analysis of ionospheric delay on IPPs in the northern region of China. Then, considering that some IPPs formed by Compass are motionless, the 5 GEO satellites in Compass are selected to study the ionospheric temporal variation in north China. Finally, we conduct a simulation on ionospheric delay estimation at

IGPs in dual constellations through Kriging and Improved Kriging algorithm. It can be concluded that Improved Kriging algorithm can obtain a more accurate estimate of the ionospheric delay.

2 Ionospheric Delay Estimation

The IGPs on the ionospheric thin layer model of BDS is specified in BDS Signal in Space Interface Control Document Version 2.0 (BDS-SIS-ICD-2.0), the coverage is E70°–E145°, N7.5°–55°, the IGPs are divided by latitude and longitude 5° × 2.5°, the update cycle is 5 min. In order to estimate GIVD, the ionospheric thin layer model should firstly be modeled. Thus firstly we should research on the characteristic of ionospheric delay in north China (In this paper the E100°–E135°, N30°–55°region). According the former work done for WAAS in US, we make spatial correlation analysis on the ionospheric measurements taken by the reference stations which are equipped with dual frequency receivers.

2.1 Spatial Correlation of Ionospheric Delay

We will make a statistical analysis on ionospheric delays at IPPs through dual frequency observations from reference stations, which are converted to vertical direction with obliquity factor F_{ipp} in (1):

$$F_{ipp} = \left[1 - \left(\frac{R_e \cos E}{R_e + h_I}\right)^2\right]^{-\frac{1}{2}} \tag{1}$$

Which E is the elevation angle, R_e is the approximate radius of Earth (6378.1363 km), h_I is the height of the ionospheric thin layer model (350 km). The distribution of the 13 reference stations in Beijing is shown in Fig. 1 as red triangles, where dual frequency receivers gathering observations on both L1/L2 of GPS and B1/B2 of BDS.

The ionospheric vertical delay model can be represented as the sum of the first order plane trend term and the space-dependent stochastic term.

$$I_{meas}(x_k) = a_0 + a_1 x_k^{(north)} + a_2 x_k^{(east)} + r(x_k) + \varepsilon \tag{2}$$

In the above formula, $I_{meas}(x_k)$ represents the vertical delay of ionosphere at $(x_k^{(north)}, x_k^{(east)})$. The first three terms represent the first order plane trend item, and $r(x_k)$ is the spatial correlation zero mean random term and ε is the observation noise. After the carrier phase smoothing code pseudorange, the accuracy of the measured value is close to the carrier phase accuracy, so ignore ε.

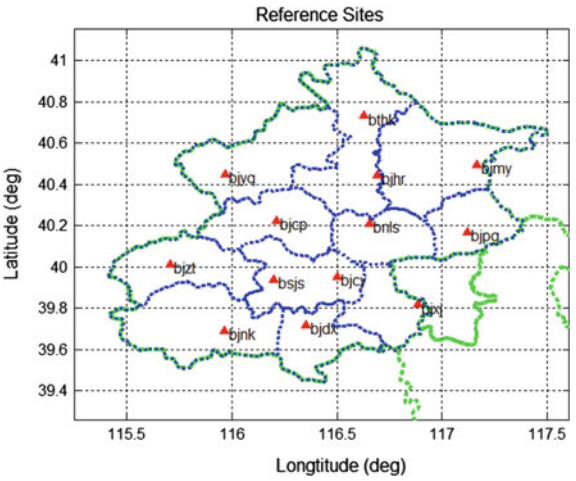


Fig. 1 13 reference station receivers (color figure online)

Using the experimental variation function, the spatial correlation characteristics of the vertical delay model of the ionosphere in the stationary period are analyzed, the result is showed as Fig. 2, the blue lines representative the space experiment variation functions of different time and the red reflects the fitting space variation function model, then we can get the following Gaussian variation function model:

$$\gamma(h) = c\left(1 - e^{-\left(\frac{h}{a}\right)^2}\right) + v, \gamma(0) = 0 \tag{3}$$

The parameters are:$c = 1.55\,\mathrm{m}^2, a = 9000\,\mathrm{km}, v = 0.05\,\mathrm{m}^2$. The covariance function of $r(x_k)$ that can be obtained is as follows:

$$C(h) = c + v - \gamma(h) = ce^{-\left(\frac{h}{a}\right)^2}, C(0) = c + v \tag{4}$$

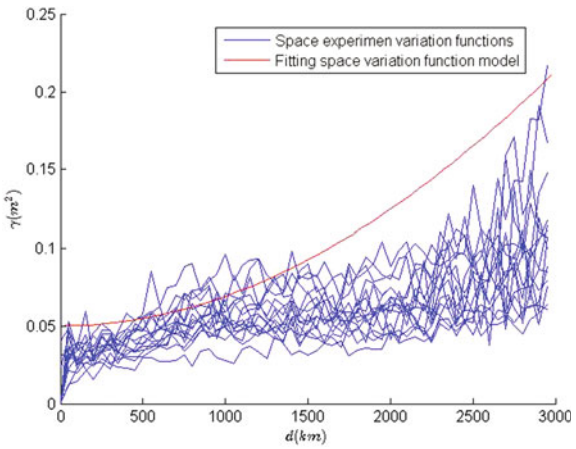


Fig. 2 Spatial experiment variation function results (color figure online)

2.2 Kriging Algorithm

2.2.1 Ionospheric Vertical Delay Model

The Kriging algorithm ionospheric vertical delay model is given by (2).

2.2.2 Grid Ionospheric Vertical Delay Estimation

Suppose that there are n sampling points $x_1, x_2, \ldots, x_n$ around the ionospheric penetrating point x, the corresponding vertical delay observations for the ionosphere are $I(x_1), I(x_2), \ldots, I(x_n)$, then the vertical delay estimate at the x-point is as follows:

$$\begin{aligned} GIVD = \hat{I}(x) &= \sum_{k=1}^{n} \lambda_k I(x_k) = \Lambda^T I \\ I &= [I_1, I_2, \ldots, I_n]^T, \quad \Lambda^T = [\lambda_1, \lambda_2, \ldots, \lambda_n] \end{aligned} \tag{5}$$

where, Λ is the coefficient vector. The problem of estimating is to find the optimal value under certain constraint conditions, that is, the Kriging coefficient Λ satisfies the following formula:

$$\begin{cases} Min & (\hat{\sigma}^2 = \Lambda^T C(\mathbf{h}) \Lambda - 2\Lambda^T C(\mathbf{h_0}) + C(0)) \\ s.t. & G^T \Lambda = X \end{cases} \tag{6}$$

where, $\hat{\sigma}^2$ is the estimated error variance, $C(\mathbf{h})$ is a matrix of n × n, representing the spatial covariance matrix between observations. $C(\mathbf{h_0})$ is a n-dimensional vector, represents the spatial covariance vector between the estimator and the observed quantity. $C(0)$ is the spatial covariance between two observations with a distance of zero, i.e. the prior variance, independent of position x. The G^T is the observation matrix and X is the position vector:

$$G^T = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ Lat_1 & Lat_2 & \ldots & Lat_n \\ Long_1 & Long_2 & \ldots & Long_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 \\ Lat \\ Long \end{bmatrix}$$

Using the Lagrangian method to solve the above equation, we can get the final Kriging coefficient Λ:

$$\begin{aligned} \Lambda &= (W - HG^T W) C(\mathbf{h_0}) + HX \\ H &= WG(G^T WG)^{-1}, \; W = C^{-1}(\mathbf{h}) \end{aligned} \tag{7}$$

So the GIVD value at the grid point can be obtained.

2.2.3 Grid Ionospheric Vertical Delay Error Estimation

In the ionospheric model, the GIVE value at the grid points represents the accuracy of the GIVD estimate. However, anomalies such as ionospheric disturbances and under sampling reduce the accuracy of GIVD estimates. This paper does not consider the IPP under-sampling conditions of the ionospheric anomaly, only the introduction of expansion model, a reasonable expansion of GIVE value to ensure the integrity of the client. Therefore, the GIVE value can be obtained from the following equation:

$$GIVE = \kappa_{99.9\%}\sigma_{GIVE}, \quad \sigma^2_{GIVE} = R^2_{irreg}\hat{\sigma}^2_{IGP} \tag{8}$$

where, $\kappa_{99.9\%}$ is the quantile of the standard normal distribution 99.9% confidence intervals. $\hat{\sigma}^2_{IGP}$ is the variance of the delay value at IGP obtained by using the covariance function during the calm period. R^2_{irreg} is the expansion factor to overcome the inaccurate model, the specific expression is as follows:

$$R^2_{irreg} = \alpha_n I^T_{meas}(W - WG(G^T WG)^{-1}G^T W)I_{meas} \tag{9}$$

$$P = \frac{2}{\pi}\int_{\phi=0}^{\frac{\pi}{2}} \left(K^2\frac{\alpha_n}{sin^2(\phi)} + 1\right)^{-\frac{n-3}{2}} d\phi \tag{10}$$

where I_{meas} is the IPP ionospheric vertical delay observation sequence, α_n can be obtained using the integrity probability assigned to the ionosphere portion, which can be computed from Eq. (10). In Eq. (10), $P = 2.25 \times 10^{-8}$ is the probability of goodness assigned to the ionospheric error part, and $K = 5.592$ is the bilateral quantile of the standard Gaussian distribution corresponding to 1 − P [2].

2.3 Improved Kriging Method

Just as we have introduced above, the traditional Kriging method utilizes only sample space-related features and only GPS constellation is used in conventional Kriging, and GIVD is estimated using only the IPP delay value at GIVD update. On the basis of the traditional Kriging method, some domestic scholars have proposed an improved Kriging method, taking into account the temporal correlation of IPP delay values, using the IPP delay values in the update cycle for algorithm to obtain more reasonable GIVE Value. However, because the GPS satellite constellation are all the MEO satellites, IPP observation position change over time, it is not easy to obtain accurate time variogram model. Specially, there are five GEO satellites in

China's Beidou system, the IPP position formed by the reference station observing GEO satellites is almost fixed, so we introduce the Beidou GEO satellite observations in this paper, it provides a convenient way to eliminate the influence of spatial location and obtain more accurate temporal variogram model.

In this paper, the IPP delay observations in the grid point update cycle are used in the estimation algorithm [3]. Since the time dimension of the IPP delay value is extended, the ionospheric vertical delay observation model of Eq. (2) changes to:

$$I_{meas}(x_k, t) = a_0 + a_1 x_k^{(north)} + a_2 x_k^{(east)} + a_3 \triangle t + r(x_k, t) \tag{11}$$

where, $\triangle t = t_0 - t$, t and t_0 is the observation time and the reference time, respectively. That is to say t_0 is the current update time. Note that the influence of Earth's rotation during update cycle requires to transform the location of IPP in the update cycle to the location in coordinate system at the time of the present update. At this point, G^T and X have turned into:

$$G^T = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ Lat_1 & Lat_2 & \ldots & Lat_n \\ Long_1 & Long_2 & \ldots & Long_n \\ \triangle t_1 & \triangle t_2 & \ldots & \triangle t_n \end{bmatrix}, X = \begin{bmatrix} 1 \\ Lat \\ Long \\ 0 \end{bmatrix}$$

Before solving the spatiotemporal correlation, the temporal correlation of the ionosphere in the northern part of China is studied. Taking into account the fixed position of the IPP formed by the Beidou GEO satellites, five GEO satellites of BDS were selected to study the temporal variation characteristics of the ionosphere in the northern part of China. The spatial variational function model and the covariance function of the random term have been obtained. The same method is adopted to research its time-dependent properties as in Sect. 2.1, the result is showed as Fig. 3, the blue lines representative the temporal experiment variation

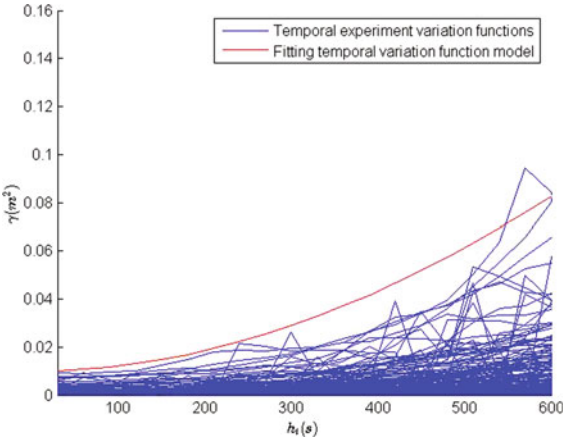


Fig. 3 Temporal experiment variation function results (color figure online)

functions of different locations and the red reflects the fitting temporal variation function model and we can obtain the following Gauss variogram model:

$$\begin{aligned} \gamma_t(h_t) &= c_t\left(1 - e^{-\left(\frac{h_t}{b}\right)^2}\right) + v_t, \gamma_t(0) = 0 \\ C_t(h_t) &= c_t + v_t - \gamma_t(h_t) = c_t e^{-\left(\frac{h_t}{b}\right)^2}, C_t(0) = c_t + v_t \end{aligned} \tag{12}$$

The parameters are: $c_t = 0.7\,\mathrm{m}^2, b = 1800\,\mathrm{s}, v_t = 0.01\,\mathrm{m}^2$

Finally, the temporal and spatial correlation characteristics of the ionospheric calm period are analyzed and the covariance function is expressed by the following model [4]:

$$C_{st}(h_s, h_t) = k_1 C_s(h_s) C_t(h_t) + k_2 C_s(h_s) + k_3 C_t(h_t) \tag{13}$$

or equivalent to its spatial and temporal variation function model:

$$\begin{aligned} \gamma_{st}(h_s, h_t) = &[k_2 + k_1 C_t(0)]\gamma_s(h_s) \\ &+ [k_3 + k_1 C_s(0)]\gamma_t(h_t) - k_1 \gamma_s(h_s)\gamma_t(h_t) \end{aligned} \tag{14}$$

Among them, C_{st}, C_t, C_s are space-time covariance function, time covariance function and space covariance function, γ_{st}, γ_t, γ_s are corresponding variation function respectively,and $C_{st}(0,0)$,$C_t(0)$, $C_s(0)$ are 'sills' of γ_{st}, γ_t, γ_s, respectively. The coefficients k_1, k_2, k_3 in the model are determined by the following equation [5].

$$\begin{aligned} k_1 &= [C_s(0) + C_t(0) - C_{st}(0,0)]/C_s(0)C_t(0) \\ k_2 &= [C_{st}(0,0) - C_t(0)]/C_s(0) \\ k_3 &= [C_{st}(0,0) - C_s(0)]/C_t(0) \end{aligned} \tag{15}$$

Using the method of experimental variation function to calculate the convergence value of the spatiotemporal variation function tends to infinity, it is easy to get $C_{st}(0,0) = 2.315\,\mathrm{m}^2$. And further get the $k_1 = 0.0044$, $k_2 = 1.0031$, $k_3 = 1.007$. At this point, the covariance function of $r(x_k, t)$ has been obtained and used in the improved Kriging estimation algorithm.

In addition, the α'_n of Improved Kriging is equivalent to α_{n-1} because the degree of freedom in Eq. (11) is 4, and the Traditional Kriging method is 3. The rest of the calculation steps are the same as those of the traditional Kriging method.

3 Algorithm Implementation and Results

In this paper, we select the 13 reference station receivers as shown in Fig. 1 of the dual-band hybrid constellation observations on the algorithm proposed in the simulation verification. This paper sets the GPS observation satellite elevation

cutoff angle to be 15°, BDS observation satellite elevation cutoff angle to be 10°. For the traditional Kriging method, the IPP search parameters: $R_{min} = 800\,\text{km}$, $R_{max} = 2000km$, $\triangle R = 50\,\text{km}$, $N_{t\arg et} = 30$ and $N_{min} = 10$. Since the Improved Kriging method uses IPP delay values in the update period of the dual constellations grid point information, So the IPP search parameters is adjusted to: $N_{t\arg et} = 200$, $N_{min} = 50$, the remaining parameters remain unchanged [3].

3.1 Algorithm Accuracy Analysis

The method of data culling is used to analyse the estimation accuracy of the algorithm. The specific steps are as follow: For each IPP ionospheric delay measurement, it is temporarily removed, the value is estimated with the remaining IPP measurements, and then the temporarily removed IPP measurements are returned. Repeat the above steps for all IPP delays Value, the estimated error of all IPP delay values is obtained.

In this section, the data of different ionospheric disturbances are selected (the three-day ionospheric disturbance index K_p is 2, 4 and 7 respectively), the error maximum value and error mean square value are calculated to compared the fitting accuracy of traditional Kriging method and improved Kriging method. The results are shown in Table 1; the results show that the maximum error of IPP delay estimation obtained by improved Kriging method is significantly smaller than that of traditional Kriging method under different ionospheric conditions. However, the root mean square of the estimated error obtained by the two Kriging methods is very close, that is the estimation accuracy of the two is generally equivalent.

3.2 Algorithm Performance Analysis

So as to assess the performance of the improved Kriging algorithms, we compare the GIVE In the service area. This section selects March 1, 2016 (K_p is 3) when the GPS is 21600s, carries on the contrast analysis. Figure 4 shows the GIVE contrast

Table 1 The maximum (Max) and the root mean square (RMS) of the estimation error calculated by Kriging and improved Kriging methods under different ionospheric conditions

Date	Kriging		Improved Kriging	
	Max(m)	RMS (m)	Max (m)	RMS (m)
2016.3.13	2.4591	0.2155	1.7909	0.2179
2016.3.19	1.9951	0.2240	1.9581	0.2190
2016.3.6	2.8157	0.2207	2.7647	0.2149

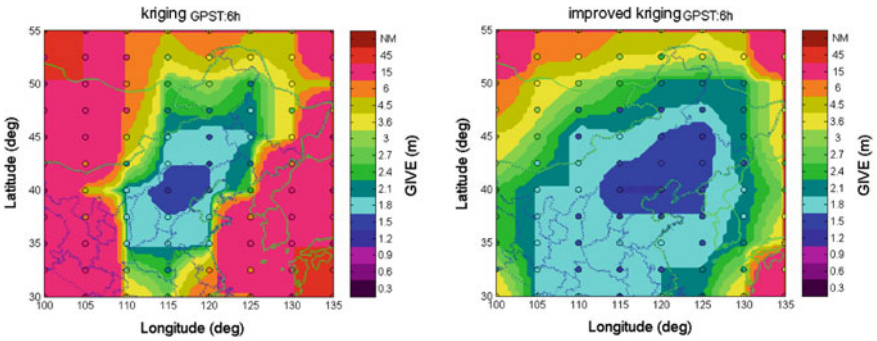


Fig. 4 The GIVE calculated by the two Kriging algorithms at GPS time of 6 h

of the two algorithms. The results show that the GIVE value computed by Improved Kriging method is smaller than traditional Kriging method in the 80% of selected service region, especially in the sparsely region of IPP point. At different times in the day, compared to the traditional Kriging method the improved Kriging method calculates the GIVE value, which mostly reduces more than 30%.

3.3 User Integrity Analysis

This article selects BDSQ and BJLX receivers as user stations and compare the estimated error of the user delay value and GIVE calculated by the two methods. Figure 5 shows the results when March 1, 2016 BDSQ and March 19, 2016 BJLX stations continuously track the GPS PRN 1and PRN 22, respectively. It indicates that Improved Kriging's UIVE can tighten the user IPP delay estimate error.

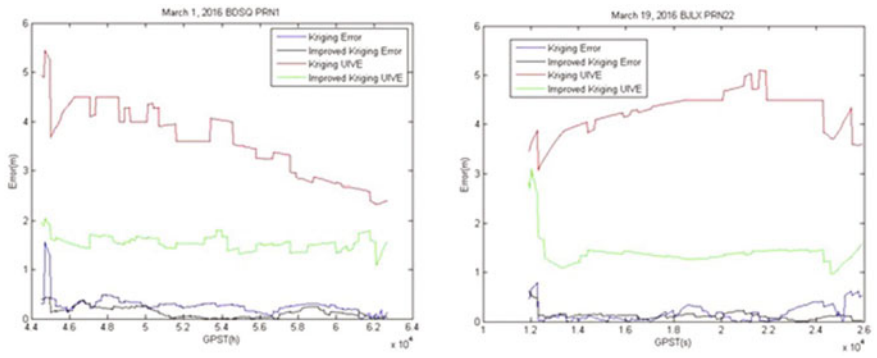


Fig. 5 Results of user ionospheric delay correction error and error bound from two algorithms

4 Conclusion

In order to solve this difficulty which the GPS satellite constellation are all the MEO satellites, IPP observation position change over time, it is not easy to obtain accurate time variogram model. In this paper, in view of characteristics in Compass which contains GEO, IGSO and MEO three orbital satellites, in particular, there are five GEO satellites which will contribute to some stationary IPPs, so we introduce the Beidou GEO satellite dual frequencies observations in this paper, the accurate time variogram model was obtained and the improved Kriging algorithm for the Chinese region is given. By which the User Ionospheric Vertical Delay Error (UIVE) can more closely envelop the correction error. First, the spatial correlation characteristics of ionospheric delay in northern China were researched. Then, the 5 GEO satellites in Compass are selected to study the ionospheric temporal variation. Finally, we conduct a simulation on ionospheric delay estimation at IGPs in Multi-constellation through Kriging and Improved Kriging algorithm. It can be concluded that, in the premise of satisfying integrity, the GIVE is reduced more than 30% in most parts of north China.

Therefore, under the premise of satisfying the integrity, the improved Kriging method improves the usability of the system service and is of great significance to the engineering application of the grid correction method.

Acknowledgements Thanks to the Collaborative Innovation Center of Geospatial Technology. This paper has been supported by the National Natural Science Foundation of China (41304024).

References

1. Sparks L, Blanch J, Pandya N (2011) Estimating ionospheric delay using kriging: 1. Methodology. Radio Sci 46(6)
2. Blanch J (2013) Using Kriging to bound satellite ranging errors due to the ionosphere. Stanford University, Stanford
3. Zhang Q, Li R (2015) A new method of ionospheric grid correction based on improved Kriging. In: Sun J, Liu J, Fan S, Lu X (eds) China satellite navigation conference (CSNC) 2015 proceedings: volume II. Lecture notes in electrical engineering), vol 341. Springer, Berlin, Heidelberg
4. Cesare LD, Myers DE, Posa D (2001) Estimating and modeling space-time correction structures. Stat Probab Lett 51(1):9–14
5. Wang S, Meng F, Zhu B (2015) The ionospheric delay correction methods in SBAS of BeiDou. 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 1596–1602

A Study on Construction of Ionospheric Spatial Threat Model for China SBAS

Dun Liu, Jian Feng, Li Chen and Weimin Zhen

Abstract Finite sampling of the ionosphere from the limited number of observations will threat the integrity of SBAS (Space-Based Augmentation System). Ionospheric spatial threat model is introduced to over-bound residual errors in delay estimation for this under-sampled scenario. Ionospheric anomaly existing in low latitude areas exerts considerable influence on threat model realization. Methods are presented to construct spatial threat model for China area with data during periods of severe storms, including the modeling of variogram, selection of disturbance detection threshold and design of data-deprivation strategy. Verification is also conducted to show its applicability in SBAS with GNSS data from China and around areas from typical storm events.

Keywords Ionospheric · Spatial threat model · Kriging method · Ionospheric grid model · SBAS

1 Introduction

SBAS is a regional system developed to augment the performance (accuracy, integrity availability etc.) of GNSS. Currently the largest error source of positioning error in the system is signal delay caused by the ionosphere. To allow users to take account of such error, SBAS computes and broadcasts ionospheric delays, called the Grid Ionospheric Vertical Delay (GIVD), and integrity bound, called the Grid Ionospheric Vertical Error (GIVE), at a set of regularly-spaced ionospheric grid points (IGP). A variety of measures are designed in this bound estimation to protect the user against positioning error due to the ionosphere and the possible ionospheric irregularity [1–3].

D. Liu (✉) · J. Feng · L. Chen · W. Zhen
No.22nd Research Institute, CETC, Qingdao 266107, Shandong, China
e-mail: dun.l@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_16

Limited observations from the SBAS could lead to a situation in which an ionospheric area, where great gradient exists, is under-sampled. When the ionospheric delay at an ionospheric pierce point (IPP) located in this under-sampled area is estimated with delays at IPPs from the surrounding quiet ionosphere, a risk arise as the estimated GIVE cannot bound the actual error effectively. This threat is especially serious when an ionospheric storm occurs. The ionospheric spatial threat model is an effective measure to increase the error bound to protect user for this kind of influence [4].

Ionospheric anomaly is a phenomenon of the increasing of electron density along the latitude in the low latitude area. The impact of ionospheric anomaly on SBAS is similar to that of a modest storm. For China, the south of the Changjiang River is generally considered to be in the ionospheric anomaly zone, and regions near Haikou and Guangzhou are under the crest of the anomaly [5]. It is expected that the SBAS integrity risk due to the under-sampling measurement of ionosphere will be higher in south of China than in mid-latitude areas, such as north of China.

The ionospheric anomaly will impact the SBAS on various aspects, including the construction of ionospheric correlation model, selection of the irregularities detection threshold, and the implementation of spatial threat model etc. [2, 3, 6, 7]. In SBAS, it is necessary to make tradeoff among various parameters in these models to maximize the performance of system integrity and availability. Consequently, constructing the ionospheric threat model for SBAS in China area will have to take into account the ways the ionospheric variogram model and the disturbance detection threshold are realized.

We have constructed the ionospheric variogram model for China area [8]. Here in this paper, we will focus on the ionospheric spatial threat model for SBAS in China. The second section introduces the implementation of SBAS ionospheric grid model based on kriging method; the third section introduces the strategies in variogram model realization, disturbance detection threshold selection and data depriving method in developing the spatial threat model; analysis is made on the spatial threat model in the fourth section, and conclusion is given in the end.

2 Description of Ionospheric Spatial Threat Model

2.1 Ionospheric Grid Model Based on Kriging Method

In ionospheric grid model, delay at the grid (GIVD) can be estimated with kriging method as [2, 9–11]:

$$I_{est}(x) = \sum_{k=1}^{n} \lambda_k \mathrm{I}_{\mathrm{meas}}(x_k) \tag{1}$$

where λ_k is the weighting matrix, I_{meas} is the vertical ionospheric delays at the IPPs located in x_k, n is the number of IPPs available for GIVD estimation.

$$\lambda = \left(W - WG\left(G^T WG\right)^{-1} G^T W\right) C(x, x_k) + WG\left(G^T WG\right)^{-1} X \tag{2}$$

$$W = \left(C(x_k, x_l) + M(x_k, x_l)\right)^{-1} \tag{3}$$

$$X = \begin{bmatrix} 1 & x^{(east)} & x^{(north)} \end{bmatrix}^T \tag{4}$$

where G is the coefficient matrix determined by satellite and user's locations, W is the weighting matrix, C is the covariance matrix derived from the ionospheric delay variogram model, M is the measurement noise matrix, $x^{(North)}, x^{(east)}$ are coordinates in the north and east direction respectively in a local frame with IGP as the origin.

Formal uncertainty in ionospheric delay estimation on the grid point is:

$$\sigma^2_{IGP} = \lambda^T C(x_k, x_l)\lambda - 2\lambda^T C(x_k, x) + \left(\sigma^{total}_{decorr}\right)^2 + \lambda^T M\lambda \tag{5}$$

SBAS sets up mechanism to detect possible ionospheric disturbance. When the statistic χ^2_{irreg} exceeds the threshold, irregularity detector is tripped and the grid is set as "unavailable". When χ^2_{irreg} below the threshold, then "inflates" the estimated error variance to improve the effectiveness of error bounding. The formal error used to evaluate the GIVE considering the "inflating" factor is [3, 6, 7]:

$$\tilde{\sigma}^2_{IGP} = R^2_{irreg}\left[\lambda^T C(x_k, x_l)\lambda - 2\lambda^T C(x_k, x) + \left(\sigma^{total}_{decorr}\right)^2\right] + \lambda^T M\lambda \tag{6}$$

where R^2_{irreg} is the "inflating" factor used to account for statistical uncertainty in the level of ionospheric disturbance [1–3].

When there are not enough observations sampling the disturbed ionosphere, the undetected spatial variation of ionospheric delays will threat the effectiveness of error bounding at the grid. The ionospheric threat model is designed as counter-measures against this under-sampling scenario. The ionospheric threat model consists of tabulated corrections that augment the confidence bound of the delay estimates [3, 4, 6, 7]. The correction $\sigma^2_{undersampled}$ could be retrieved with two metrics describing the skewed IPPs distribution. The first metric is the radius of planar fit R_{fit}, i.e. the distance from the IGP to the most distance fit IPP. The second metric, the relative centroid of IPPs distribution $RCM = R_{Cent}/R_{fit}$, is a measure of how uniform the distribution is. Definition of the two metrics is shown is Fig. 1. GIVE augmented by threat model is given by:

$$\sigma^2_{GIVE} = \tilde{\sigma}^2_{IGP} + \sigma^2_{undersampled} \tag{7}$$

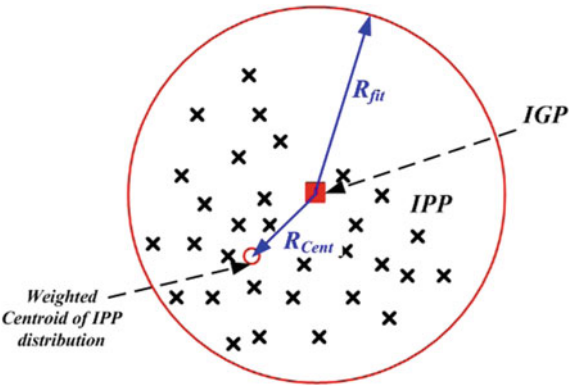


Fig. 1 Fit radius and relative centroid metric (RCM) plot

where σ^2_{GIVE} is the over-bounding error variance used to define the GIVE at an IGP. It can be seen that effects from uncertainty in ionospheric grid model, ionospheric spatial de-correlation, measurements error and spatial under-sampling measuring are included in GIVE estimation.

2.2 Construction of Ionospheric Spatial Threat Model

According to the definition of ionospheric delay error bound in SBAS, the GIVE should meet the following requirement:

$$\left|\bar{I}_k - \tilde{I}_k\right|^2 \le K^2_{undersampled}\left[\tilde{\sigma}^2_k + \sigma^2_{undersampled}\right] \tag{8}$$

where $\bar{I}_k$, $\tilde{I}_k$ are respectively the measurement and estimated ionospheric delays at IPPs. $K_{undersampled}$ is a scalar that controls how far out on the tail of the residual distribution we wish to be (in practice, $K_{undersampled}$ is set to a value of 5.33). $\tilde{\sigma}^2_k$ is the inflated variance of the delay estimate at the IPP. $\sigma^2_{undersampled}$ is the augmentation term from ionospheric spatial threat model.

According to Eq. 8, $\sigma^2_{undersampled}$ can be derived as:

$$\sigma^2_{undersampled} = \frac{\left|\bar{I}_k - \tilde{I}_k\right|^2}{K^2_{undersampled}} - \tilde{\sigma}^2_k \tag{9}$$

Actually in ionospheric threat modeling, the tabulation of the raw data is performed using the following equation. It indicates the maximization is performed over measurements k and over the time interval T following each fit epoch.

$$\sigma^{raw}_{undersampler}\left(R_{fit}, RCM\right) = \max_{over\ k,T}\left(\sigma_{undersampled,k}\right) \tag{10}$$

In our initial effort in establishing the threat model in this paper, $\sigma_{undersampled,k}$ was evaluated only for delays at the same epoch as the fit, and no update time interval T is considered.

3 Realization of Ionospheric Threat Model for China Area

3.1 Modeling Data

GPS measurements from stations in China and around areas are used for modeling (Fig. 2). Data are selected from periods when severe ionospheric storms occurred during the last solar cycle (Table 1).

3.2 Modeling of Ionospheric Variogram

Ionospheric delay variogram affects the kriging-based grid model estimation in different ways, including the formal error estimation at IGPs, inflating factor calculation, and realization of ionospheric spatial threat model [9–11].

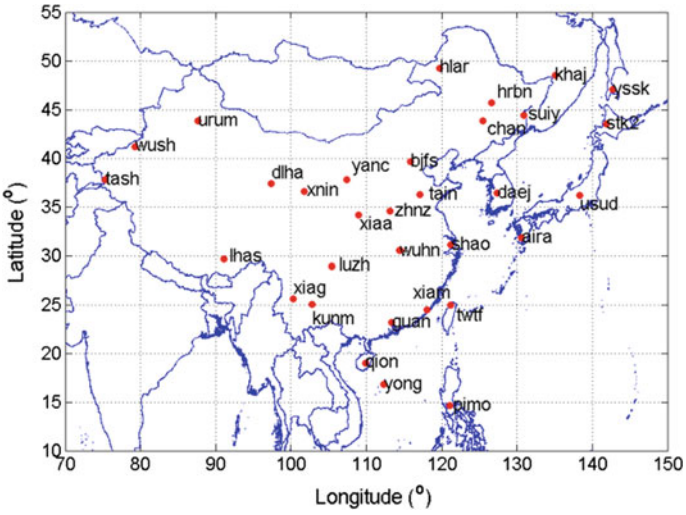


Fig. 2 GPS network in China and around areas

Table 1 List of threat model storm days

Storm day	K_p	*Dst*	Storm class
4/6/2000	8	−287	Severe
7/6/2000	9	−288	Extreme
7/15/2000	9	−289	Extreme
7/16/2000	8	−301	Strong
3/31/2001	9	−387	Extreme
10/29/2003	9	−350	Extreme
10/30/2003	9	−383	Extreme
10/31/2003	8	−307	Severe
11/20/2003	9	−422	Extreme
11/21/2003	7	−309	Extreme
9/11/2005	9	−147	Severe

Ionospheric anomaly existing in low latitude of China results in larger intercept at zero distance and "bump" effect at the distance of about 1000 km in ionospheric experimental variogram [8]. To construct a reasonable variogram model, only the linear variation in the experimental variogram was fitted, while the "bump" effect was neglected. The realized ionospheric variogram model and the corresponding covariance model are shown below.

$$\gamma(x_k, x_l) = \left(\sigma_{decorr}^{total}\right)^2 - \left[\left(\sigma_{decorr}^{total}\right)^2 - \left(\sigma_{decorr}^{nom}\right)^2\right] \exp\left(-\frac{D_{k,l}}{d_{decorr}}\right) \tag{11}$$

$$C(x_k, x_l) = \left[\left(\sigma_{decorr}^{total}\right)^2 - \left(\sigma_{decorr}^{nom}\right)^2\right] \exp\left(-\frac{D_{k,l}}{d_{decorr}}\right) \tag{12}$$

$$D_{k,l} = \sqrt{(x_k - x_l)^T (x_k - x_l)} \tag{13}$$

where $\sigma_{decorr}^{total} = 1.25$, $\sigma_{decorr}^{nom} = 0.45$, d_{decorr} is de-correlation distance and a constant value of 8000 km is adopted. x_k, x_l is coordinates of IPPs in the local frame centered at the IGP.

Modeling the variogram in this way, the residual in delay estimates caused by the "bumps" is actually accounted for in the ionospheric threat model. The following part will show the feasibility of this variogram modeling effort.

3.3 Selection of Threshold in Ionospheric Disturbance Detection

In SBAS an irregularity detector is designed for a possible ionospheric disturbance. When the statistic χ_{irreg}^2 exceeds the specified threshold value, the GIVE at the IGP is set to 45 m, indicating the grid 'unavailable' for use. Only when the value of

χ^2_{irreg} is smaller than the threshold, an effective error bound will be estimated. As the ionospheric spatial threat model acts to further improve the effectiveness of delay error bound, only those data which have passed the disturbance detection could be included in modeling [1, 3].

The irregularity metric used for ionospheric disturbance detection is defined as below [3]:

$$\chi^2_{irreg} = \frac{R_{noise}\chi^2}{\chi^2_{threshold}} \tag{14}$$

where χ^2 is the goodness-of-fit statistic associated with the delay estimate at a given IGP. $\chi^2_{threshold}$ is a threshold that can be calculated using the inverse χ^2 cumulative probability distribution with the number of IPPs and the allowable false alarm rate. R_{noise} is the constant determined by observation noise and ionospheric delay covariance model.

As can be seen from Eq. 9, the selection of disturbance detection threshold is a trade-off between the two components. On one hand, larger disturbance detection threshold means more irregularity affected delays will be incorporated in grid model estimation, increasing the grid error variance $\tilde{\sigma}^2_k$ and leading to a smaller value of spatial threat model. On the other hand, degradation of grid model will increase the error of $\left|\bar{I}_k - \tilde{I}_k\right|$, and consequently increase the value of spatial threat model. Additionally, the value of the detection threshold can directly affect system availability since the grid status will be set to "unavailable" when the threshold is passed to trip the irregularity detector.

Comprehensive analysis has been conducted among various threshold values, and the resulting system integrity and availability performance. In our study, the final detection threshold was selected as 4 (in dimensionless units).

3.4 Depriving Strategy of Ionospheric Delay Data

Due to the finite sampling of the ionosphere from the limited number of observations, a user within the SBAS service volume may have a user IPP that experiences an ionospheric irregularity that the system did not sufficiently sample. Such irregularities occur often during severe ionospheric storms. In order to protect the user from such threat, an under-sampled ionospheric threat model is applied to the GIVE. Two distinct types of threat posed by the under-sampled irregularities have been identified [4, 12]:

(1) "*Blob*" *spatial threat model*

Subject to limited measurements, the SBAS would sample a regional disturbed ionosphere with few observations. When delays for IPPs located in the disturbed area are estimated with delays from the surrounding IPPs in quiet ionosphere, a risk

would arise in bounding the residual errors sufficiently. This is also called the "*blob*" spatial threat model.

(2) "*Wall*" *spatial threat model*

In the "*wall*" spatial threat model, ionospheric pierce points locate on the same side of the grid points, large residual in delay estimate at the grid point will be caused with measurements at these pierce points. It describes an extreme scenario which occurs often near the border of SBAS when a severe storm happens.

In construction of the ionospheric spatial threat model, the data-deprivation scheme is adopted to simulate multiple examples of threats using the same set of observation data. For each IGP, various masks are employed to select IPPs to be removed from fits and subsequently treated as possible user measurements. Residual error is then tabulated by comparing predicted and actual delays for the excluded observations that lie near the IGP.

"*Blob*" threat is simulated by single station deprivation, i.e. excluding IPPs from a single reference station. Single station deprivation is performed sequentially by removing one station at a time until all stations in SBAS have been considered.

"*Wall*" threat is simulated by directional station deprivation, in which multiple sites located at the edge of service volume in a particular direction is removed sequentially.

Effects of ionospheric anomaly to SBAS can be treated as that from a modest ionospheric storm to some extent. As the ionospheric anomaly affects most areas to the south of Changjiang River in China, a reasonable station deprivation strategy should be considered.

In this study, directional station deprivation is conducted to eliminate sites of different numbers respectively from 8 directions: east, west, south, north, northeast, northwest, southeast and southwest. For station deprivation from the directions of south, southeast and southwest, multi-site, up to all sites in the south of Changjiang River, will be removed considering the possible effects caused by ionospheric anomaly.

3.5 Realization of Ionospheric Threat Model

Based on previous data processing strategy, ionospheric spatial threat model for China area has been constructed. Figure 3 shows the tabulation of $\sigma^{raw}_{undersampler}$ for the ionospheric threat model.

The actual threat model used by SBAS is defined as the two-dimensional over-bound of the raw data [3, 4]. This ensures that $\sigma_{undersampler}$ is monotonically increase with respect to the IPPs distribution metrics R_{fit} and RCM. Figure 4 shows the final under-sampled ionospheric spatial threat model based on kriging.

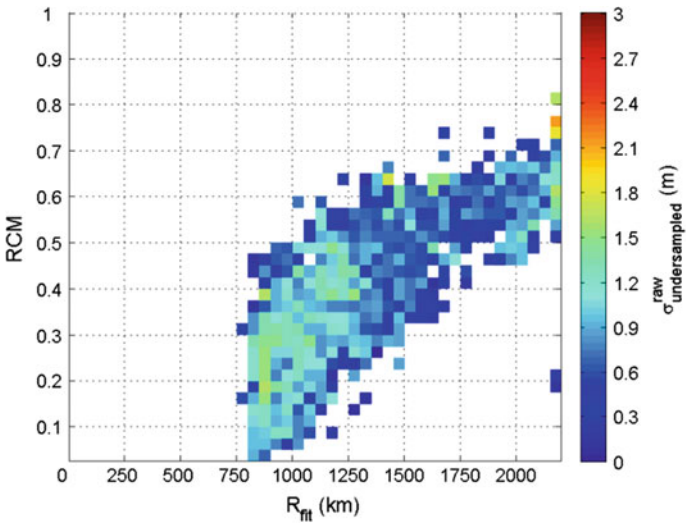


Fig. 3 Raw data for ionospheric spatial threat model for China areas

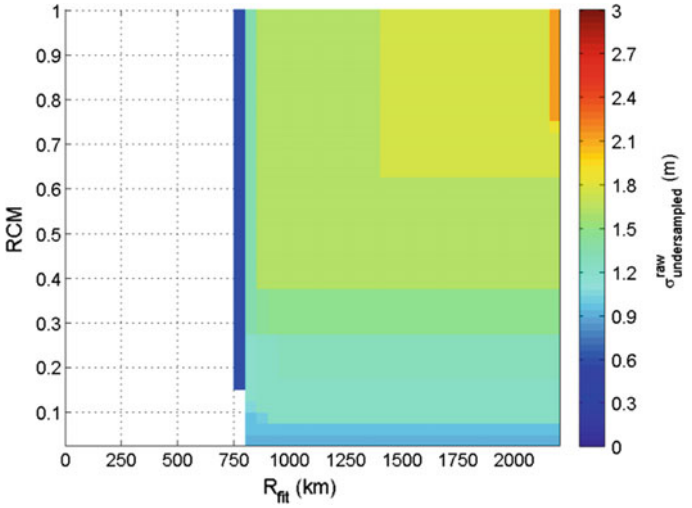


Fig. 4 Ionospheric spatial threat model for China areas

4 Validation of Ionospheric Threat Model

4.1 Cross Validation of Ionospheric Threat Model

The ionospheric spatial threat model was analyzed with cross validation method, in which each IPP was taken as a virtual ionospheric grid point, and ionospheric delay and its variance were estimated with the ionospheric grid model. The normalized residual was then given by the following equation with different over-bounding error variance $\sigma_{bound,IPP}$ [9–11]:

$$\frac{I_{v,IPP} - \widehat{I}_{v,IPP}}{\sigma_{bound,IPP}} \tag{15}$$

where $I_{v,IPP}$, $\widehat{I}_{v,IPP}$ are respectively measurement and estimated vertical ionospheric delays. Error bound $\sigma_{bound,IPP}$ will be calculated from Eqs. 5, 6 and 7 respectively. In Figs. 5, 6 and 7 these different generalized distributions of ionospheric delay errors are given. The standard normal distribution is also given in each figure (red line).

In SBAS, it is required that the ionospheric delay error should be effectively bounded by the over-bounding error variance. For the distribution of generalized delay error, this requires that the distribution of the normalized residuals should be always over-bound by the standard normal distribution.

It can be seen that, as the formal error variance is inflated and augmented by the ionospheric threat model, the normalized residuals are well overbounded by the zero mean unit-variance Gaussian distribution (red line). There is some exception in

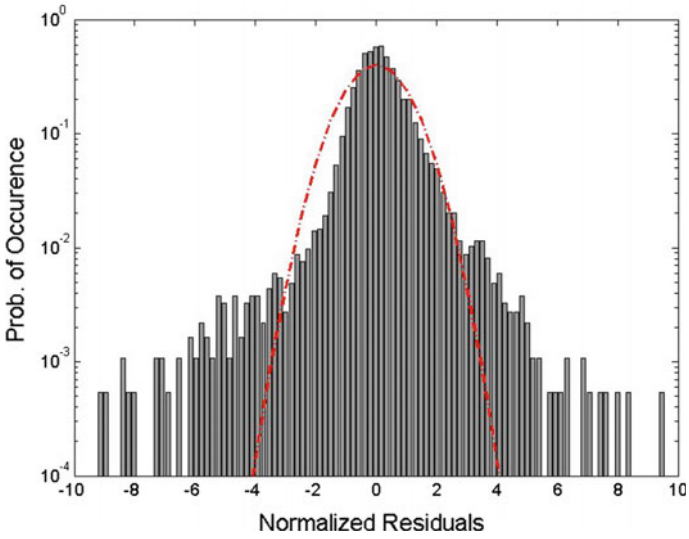


Fig. 5 Histogram of normalized residuals with σ_{IGP} (Color figure online)

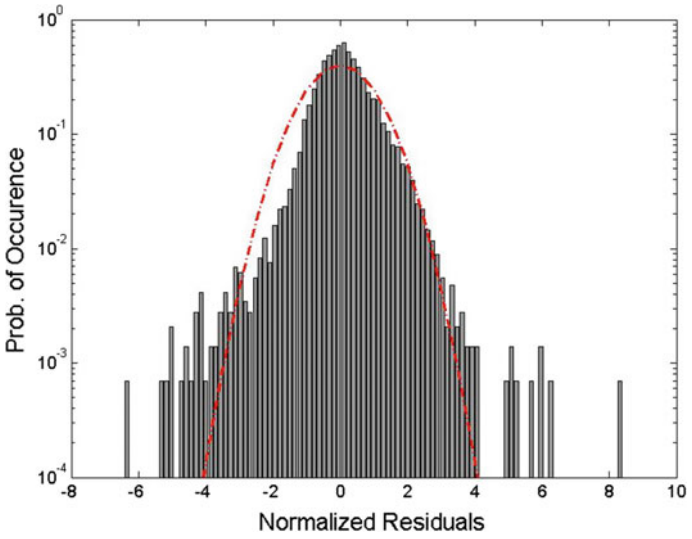


Fig. 6 Histogram of normalized residuals with $\tilde{\sigma}_{IGP}$ (Color figure online)

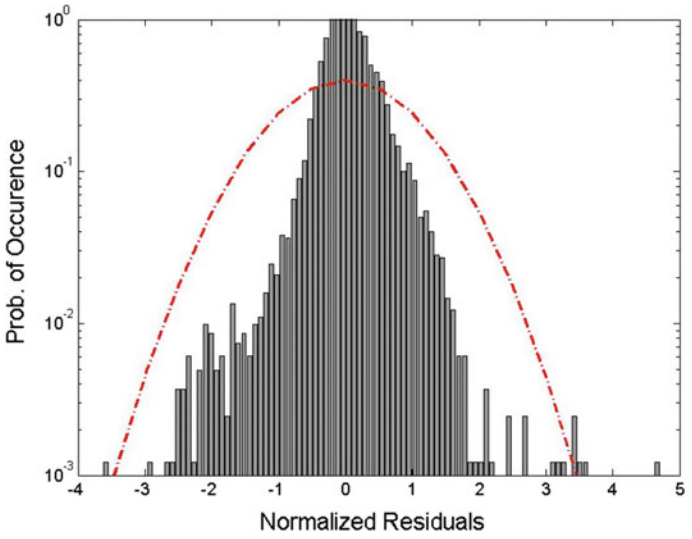


Fig. 7 Histogram of normalized residuals with σ_{GIVE} (Color figure online)

Fig. 7 at the end of the tail of residual distribution. Upon closer examination of these abnormal values, we found it caused by the observations at the outermost south edge where the test measurement was poorly surrounded by measurement. As the results show, with the inflating factor and threat model incorporated, the confidence bound is more conservative to over-bound the delay residual.

4.2 Availability of SBAS in China

Figure 8 shows the availability of APV-I service of SBAS in China areas with the ionospheric spatial threat model. The data are from a stormy day under modest solar activity (11th Sep. 2005). The analysis is conducted with a modified MAAST software [13]. As can be seen, in most region of the mainland in China the availability can reach 99%. For the region south to Guangzhou, Kunming, the availability reduces to about 95% because of its location under the ionospheric anomaly crest, the most affected area, and few ground stations. For the west and north-east region, the availability also decreases because of less ground stations available in these areas.

The analysis shows the established ionospheric spatial threat model can augment the delay error bound effectively and the availability of 95% of could be available for SBAS in China during this typical ionospheric storm period. More efforts will be made to analyze the performance of the model under even more stormy days in the future.

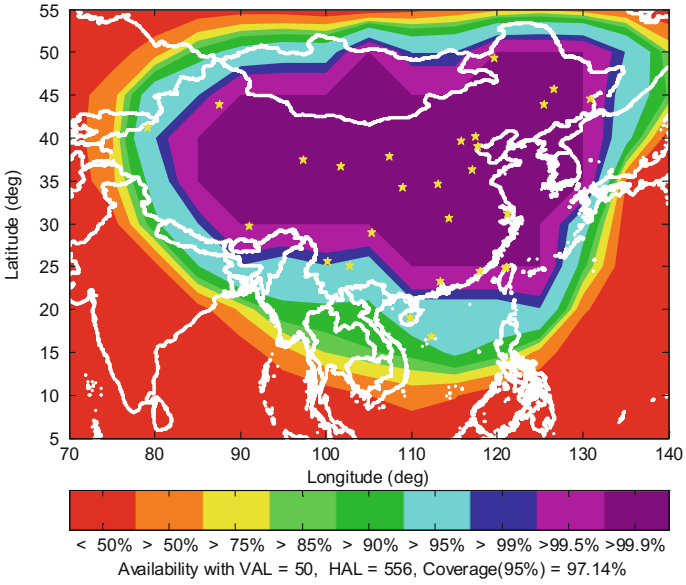


Fig. 8 SBAS availability assessment for APV-I service in China and around areas

5 Conclusion

The ionospheric spatial threat model is designed to provide protection from delay estimation error due to the under-sampling of irregularities in the ionosphere. It will be used to increase GIVEs for a more effective error bounding for SBAS.

Ionospheric threat model is subject to the factors such as the ionospheric delay variograms, threshold of disturbance detection etc. Ionospheric anomaly in low latitude area of China will affects the variogram models, the disturbance detection method etc. and consequently the realization of ionospheric threat model.

Ionospheric delay spatial threat model was developed with consideration on reasonable ionospheric variogram model construction, appropriate detection threshold selection, and station deprivation strategy accounting for the ionospheric anomaly effects. Data from typical severe ionospheric storms are also used to assess the performance of the threat model with cross validation and system availability analysis. Result shows the effectiveness of the constructed ionospheric spatial threat model.

It should be pointed out that, as the ionospheric threat model is depend on variogram models and detection threshold, a new threat model should be realized when these underlying models are improved. This will be our future work. Also, system performance under different models and parameters will be studies for an optimal ionospheric grid model for SBAS in China.

References

1. Walter T, Hansen A, Blanch J et al (2000) Robust detection of ionospheric irregularities. In: Proceedings of ION GPS 2000, Institute of Navigation, Salt Lake City, Utah, pp 209–218
2. Sparks L, Blanch J, Pandya N (2011) Estimating ionospheric delay using kriging: 1. Methodology. Radio Sci 46 RS0D21. doi:10.1029/2011RS004667
3. Sparks L, Blanch J, Pandya N (2011) Estimating ionospheric delay using kriging: 2. Impact on satellite-based augmentation system availability. Radio Sci 46, RS0D22. doi:10.1029/2011RS004781
4. Altshuler ES et al (2001) The WAAS ionospheric spatial threat model. In: Proceedings of ION GPS 2000, Institute of Navigation, Salt Lake City, UT, pp 2463–2467
5. Xiong H et al (2000) Radio propagation. Publishing House of Electronics Industry, Beijing
6. Sparks L, Blanch J, Pandya N (2013) Kriging as a means of improving WAAS availability. In: Proceedings of ION GNSS 2013, Institute of Navigation, Portland, OR, pp 2013–2020
7. Pandya N, Sheng F, Castaneda O et al (2012) Using Kriging to optimize WAAS performance over the entire solar cycle. In: Proceedings of ION GNSS 2012, Institute of Navigation, Portland, OR, pp 1310–1333
8. Liu D, Yu X, Chen L, Zhen WM (2017) Analysis on ionospheric delay variogram realization in China area. In: Proceedings of CSNC 2017, Shanghai China, 15–17 May 2017
9. Blanch J (2002) an ionosphere estimation algorithm for WAAS based on Kriging. In: Proceedings of ION GPS 2002, Institute of Navigation, Portland, OR, pp 816–823
10. Blanch J, Walter T, Enge P (2002) Application of spatial statistics to ionosphere estimation for WAAS. In: Proceedings of ION NTM 2002, Institute of Navigation, San Diego, CA, pp 719–724

11. Blanch J (2003) Using Kriging to bound satellite ranging errors due to the ionosphere. Ph.D. thesis, Department of Aeronautics and Astronautics, Stanford University
12. Sparks L, Komjathy A, Manucci A (2005) Extreme ionospheric storms and their impact on WAAS. In: Proceedings of the ionospheric effect symposium 2005, Alexandria, VA, May 2005
13. Jan S-S et al (2001) Matlab simulation toolset for SBAS availability analysis. In: Proceedings of ION GPS 2001, Institute of Navigation, Salt Lake City, UT, pp 2366–2375

RAIM Algorithm Based on Robust Extended Kalman Particle Filter and Smoothed Residual

Zhen Li, Dan Song, Fei Niu and Chengdong Xu

Abstract With the rapid development of Global Navigation Satellite System (GNSS), receiver autonomous integrity monitoring (RAIM) has become an essential part of integrity monitoring of navigation satellite system. Conventional RAIM algorithm requires the observation noise obeying Gaussian distribution, but in some conditions it obeys non-Gaussian distribution. Particle filter is applicable to nonlinear and non-Gaussian system, thus RAIM algorithm performance is improved under non-Gaussian noise with particle filter method. However, particle degeneration interferes the performance of particle filter. In this paper, robust extended Kalman particle filter is proposed and applied to receiver autonomous integrity monitoring. The importance density function of particle filter is calculated by robust extended Kalman filter in order to improve the accuracy of state estimation when pseudo-range bias exists, and the particle degeneration is restrained. On this basis, the smoothed residual test statistics is set up for satellite fault detection and isolation. The simulation results show that RAIM algorithm based on robust extended Kalman particle filter and smoothed residual can well detect and isolate the faulty satellite under the condition of non-Gaussian observation noise. Compared to the RAIM algorithm based on particle filter, the new RAIM algorithm has a better performance on fault detection, and its position accuracy is improved.

Keywords Receiver autonomous integrity monitoring · Robust extended Kalman particle filter · Importance density function · Smoothed residual

Z. Li (✉) · D. Song · C. Xu
School of Aerospace Engineering, Beijing Institude of Technology, Beijing, China
e-mail: 18811528909@163.com

D. Song
e-mail: songdan0207@163.com

C. Xu
e-mail: xucd@bit.edu.cn

F. Niu
Beijing Satellite Navigation Center, Beijing, China
e-mail: niufei009@sina.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_17

1 Introduction

With the rapid development of Global Navigation Satellite System (GNSS), integrity has become an important service performance in navigation satellite system. Receiver autonomous integrity monitoring is an essential part of navigation satellite system integrity monitoring [1], which uses the redundant observations of receiver to achieve the fault detection and isolation of satellite. RAIM algorithm can be divided into Snapshot method and filtering method. Snapshot method includes parity method, least square method et al. which uses the redundant observations to detect and isolate fault by collection consistency test [2, 3]. It's quick and easy to implement without external devices. Filtering method mainly uses Kalman filter algorithm which has a better robustness than Snapshot method [4]. It makes full use of prior information to increase the observation redundancy. The performance of RAIM algorithm depends on the observation noise distribution which should be Gaussian distribution, but usually the observation noise is non-Gaussian in practical application which causes the performance degraded [5]. Particle filter is applicable to nonlinear and non-Gaussian system, thus RAIM algorithm performance is improved under non-Gaussian noise with particle filter method [6, 7]. However, particle degeneration interferes the performance of conventional particle filter, and the calculation burden is heavy because of multiple particle filter which monitor different failure modes in parallel.

Aiming at these problems, this paper proposes a new RAIM algorithm based on robust extended Kalman particle filter (REK-PF) and smoothed residual. REK-PF algorithm is an optimal importance density function method, which chooses an appropriate importance density function to limit particle degeneration [8]. Robust extended Kalman filter (REKF) is used for calculating the importance density function of particle filter, and the accuracy of importance density function under pseudo-range bias is improved by robust estimation, which improves the state estimation accuracy of particle filter. Meanwhile, the fault detection statistics is conducted by smoothed residual method, which achieves the fault detection and isolation of satellite without multiple particle filter. The simulation results verified the effectiveness of the new RAIM algorithm.

2 Robust Extended Kalman Particle Filter Algorithm

2.1 Particle Filter Algorithm

Particle filter is a non-parameterized and sequential Monte-Carlo simulation which based on bayes estimation. It uses a few random sample, also called particle, to represent the posteriori probability density of random variable of the system. It prefers to get the approximate optimal numerical solution which based on the physical model, rather than the optimal filtering of approximate model. So particle filter is applicable to strongly nonlinear, non-Gaussian system.

The state equation and observation equation of dynamic system can be expressed as follows [9].

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{w}_{k-1}) \tag{1}$$

$$\mathbf{z}_k = h(\mathbf{x}_k, \mathbf{v}_k) \tag{2}$$

where $\mathbf{x}_k$ and $\mathbf{z}_k$ is the state vector and observation vector of the system, f is the state transition function, h is the function relationship between state vector and observation vector, $\mathbf{w}_k$ is the process noise and $\mathbf{v}_k$ is the observation noise.

The algorithm process of conventional particle filter based on sequential importance resample is as follows.

1. Particle initialization. $k = 0$.

The initial particle swarm $\{\mathbf{x}_0(i); i = 1, 2, \ldots N\} \sim p(\mathbf{x}_0)$ is selected from the prior probability density $p(\mathbf{x}_0)$, where N is the number of particle. The weight of every single particle is initialized as $1/N$, which can be express as $\{\omega_0(i) = 1/N; i = 1, 2, \ldots N\}$.

2. $k = 1, 2\ldots$

I. Sequential importance sampling.

The prior particle at epoch k is selected from importance density function $q(\mathbf{x}_k|\mathbf{x}_{k-1}(i), \mathbf{z}_k)$.

$$\left\{\mathbf{x}_{k/k-1}(i); i = 1, 2, \ldots N\right\} \sim q(\mathbf{x}_k|\mathbf{x}_{k-1}(i), \mathbf{z}_k) \tag{3}$$

II. Weight updates.

The weight of every particle is updated by the observations and the observation equation.

$$\omega_k(i) = \omega_{k-1}(i)p(\mathbf{z}_k|\mathbf{x}_k(i)) \tag{4}$$

Then the weight is normalized by

$$\tilde{\omega}_k(i) = \frac{\omega_k(i)}{\sum_{i=1}^{N} \omega_k(i)} \tag{5}$$

III. Resample.

Calculating the effective number of particle $N_{eff} = \frac{1}{\sum_{i=1}^{N} (\tilde{\omega}_k(i))^2}$, and compared with the threshold N_{th}. The prior particle swarm $\left\{\mathbf{x}_{k/k-1}(i), \tilde{\omega}_k(i); i = 1, 2, \ldots N\right\}$ need to resample to get a new particle swarm $\left\{\tilde{\mathbf{x}}_{k/k-1}(i), \tilde{\omega}_k(i) = \frac{1}{N}; i = 1, 2, \ldots N\right\}$ if $N_{eff} < N_{th}$. Usually $N_{th} = \frac{2}{3}N$.

IV. State estimation.

The system state estimation at current epoch is as follows.

$$\hat{\mathbf{x}}_k = \sum_{i=1}^{N} \tilde{\mathbf{x}}_{k/k-1}(i)\tilde{\omega}_k(i) \tag{6}$$

3. $k = k+1$, turn to the calculation of next epoch.

2.2 Robust Extended Kalman Particle Filter

Particle degeneration is the main problem of conventional particle filter. As time goes on, the weights are concentrated in a small amount of particles while most of the weights are almost zero. Particles are sampled from importance density function which has a great influence on filter performance, so an appropriate importance density function can repress particle degeneration. According to the minimum variance principle, the optimal importance density function is $q(\mathbf{x}_k|\mathbf{x}_{k-1}(i), \mathbf{z}_k)_{\text{opt}} = p(\mathbf{x}_k|\mathbf{x}_{k-1}(i), \mathbf{z}_k)$. But it is unpractical to use the optimal importance density function because it is difficult to sample particle from $p(\mathbf{x}_k|\mathbf{x}_{k-1}(i), \mathbf{z}_k)$ and calculate the integration $p(\mathbf{z}_k|\mathbf{x}_{k-1}(i)) = \int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1}(i))d\mathbf{x}_k$ [10]. Conventional particle filter chooses the state transition probability as importance density function which is easy to implement, but particles will transfer to the state space without observations. Some researchers proposed extended Kalman particle filter which uses local linearization method. It uses extended Kalman filter to combine the observations with the Gaussian approximation of state to produce a Gaussian distribution as the importance density function. The particle distribution is more approximated to posterior probability distribution and the state estimation accuracy is improved [11].

However, when the observations contain bias, the particle distribution by extended Kalman filter will be influenced and the accuracy of importance density function decreases. In this paper, robust extended Kalman particle filter is proposed. The importance density function is produced by robust extended Kalman filter, thus the observations are considered and the particle sampling is more accurate under pseudo-range bias by robust estimation. The algorithm process of robust extended Kalman particle filter is as follows.

1. Particle initialization, the same as particle filter.
2. $k = 1, 2\ldots$

I. State estimation for every single particle by REKF.

$$\hat{\mathbf{x}}_{k/k-1}(i) = \boldsymbol{\Phi}_{k/k-1}\hat{\mathbf{x}}_{k-1}(i) \tag{7}$$

$$\mathbf{P}_{k/k-1}(i) = \boldsymbol{\Phi}_{k/k-1}\mathbf{P}_{k-1}(i)\boldsymbol{\Phi}_{k/k-1}^{\mathrm{T}} + \mathbf{Q}_k \tag{8}$$

$$\bar{\mathbf{K}}_k(i) = \mathbf{P}_{k/k-1}(i)\mathbf{H}_k^{\mathrm{T}}\left(\mathbf{H}_k\mathbf{P}_{k/k-1}(i)\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k\boldsymbol{\alpha}_k^{-1}\right)^{-1} \tag{9}$$

$$\hat{\mathbf{x}}_k(i) = \hat{\mathbf{x}}_{k/k-1}(i) + \bar{\mathbf{K}}_k(i)\left(\mathbf{z}_k - f\left(\hat{\mathbf{x}}_{k/k-1}\right)\right) \tag{10}$$

$$\mathbf{P}_k(i) = (\mathbf{I} - \bar{\mathbf{K}}_k(i)\mathbf{H}_k)\mathbf{P}_{k/k-1}(i) \tag{11}$$

where $\boldsymbol{\Phi}_{k/k-1}$ is the state transition matrix, $\mathbf{P}_{k/k-1}(i)$ is the prediction state covariance matrix, $\mathbf{P}_k(i)$ is the estimation state covariance matrix, $\mathbf{Q}_k$ and $\mathbf{R}_k$ are covariance matrixes of $\mathbf{w}_k$ and $\mathbf{v}_k$ respectively, $\mathbf{H}_k$ is the observation matrix and $\bar{\mathbf{K}}_k$ is the robust filter gain matrix [12]. The equivalent weight matrix of observation data is $\boldsymbol{\alpha}_k = \mathrm{diag}[\bar{p}_1, \bar{p}_2, \ldots \bar{p}_m]$ and it can be obtained by Huber function or IGG series solution. This paper chooses IGGIII solution [13], and the calculation of equivalent weight is as follows.

$$\bar{p}_i = \begin{cases} 1, & |s_i| \le k_0 \\ \frac{k_0}{|s_i|}\left\{\frac{k_1 - |s_i|}{k_1 - k_0}\right\}^2, & k_0 < |s_i| \le k_1 \\ 0 & |s_i| > k_1 \end{cases} \tag{12}$$

where $\mathbf{s}_i$ is the standardized residual [14], k_0 and k_1 are robust parameters. In this paper $k_0 = 1.5$ and $k_1 = 3.0$. When pseudo-range observations contain bias, the weight of abnormal pseudo-range observations is descended and the corresponding observation noise variance is ascended in REKF. Therefore the gain of abnormal pseudo range observations in $\bar{\mathbf{K}}_k$ is descended to improve the state estimation accuracy of particle $\hat{\mathbf{x}}_k(i)$. Then the particle estimation sample which is less effected by pseudo-range bias is $\{\hat{\mathbf{x}}_k(i); i = 1, 2, \ldots N\}$.

II. Sequential importance sampling.

After calculating $\bar{\mathbf{x}}_k = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{x}}_k(i)$ and $\mathbf{s}_k^2 = \frac{\sum_{i=1}^{N}(\hat{\mathbf{x}}_k(i) - \bar{\mathbf{x}}_k)^2}{N-1}$, which is the mean and the variance of particles respectively, the particle $\mathbf{x}_k(i)$ is sampled from importance density function $q(\mathbf{x}_k|\mathbf{x}_{k-1}(i), \mathbf{Z}_k) = N\left(\bar{\mathbf{x}}_k, \mathbf{s}_k^2\right)$.

III. Weight updates and resample, the same as particle filter.

IV. State estimation, the same as particle filter.

3. $k = k + 1$, turn to the calculation of next epoch.

3 RAIM Algorithm Based on REK-PF and Smoothed Residual

The state equation of navigation satellite system is as follows.

$$\mathbf{x}_k = \mathbf{\Phi}_{k/k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{13}$$

where $\mathbf{x} = \left[r_x, r_y, r_z, \Delta t_u\right]^T$ and r_x, r_y and r_z is the three-dimensional position coordinates of receiver in ECEF coordinates. The receiver clock bias is Δt_u, the state transition matrix is $\mathbf{\Phi}$ and the process noise is $\mathbf{w}$.

The observation equation which uses pseudo-range observations is as follows.

$$\rho_k^i = R_k^i + c(\Delta t_u - \Delta t_s) + T_k^i + I_k^i + v_k^i \tag{14}$$

where ρ_k^i is the pseudo-range observations between receiver position and the position $\left(s_x^i, s_y^i, s_z^i\right)$ of satellite i, c is the speed of light, Δt_u is the receiver clock bias, Δt_s is the satellite clock bias, T_k^i is the troposphere delay, I_k^i is the ionospheric delay, v_k^i is the observation noise of pseudo-range and R_k^i is the real distance between receiver and satellite i which can be expressed as follows.

$$R_k^i = \sqrt{\left(r_x - s_x^i\right)^2 + \left(r_y - s_y^i\right)^2 + \left(r_z - s_z^i\right)^2} \tag{15}$$

This paper combines REK-PF with smoothed residual method to achieve the fault detection and isolation of satellite. Smoothed residual method is one of the fault diagnosis method using particle filter. Its basic idea is, when the system is fault-free, then the predicted observations which is calculated by state estimation of particle filter should be close to the true observations. When system exists fault, there should be a large difference value between the predicted observations and the true observations. Using absolute value of the difference between them as residual, the average value of residual in M epochs is compared with the threshold to determine whether fault exists [15]. RAIM algorithm based on REK-PF and smoothed residual can calculate the receiver state estimation accurately, and the smoothed residual test is tested in every single observations to achieve fault detection and isolation of satellite without multiple particle filter.

The process of RAIM algorithm based on REK-PF and smoothed residual is as follows.

1. Particle initialization.

The initial particle swarm $\left\{\mathbf{x}_0(i), \omega_0(i) = \frac{1}{N}; i = 1, 2, \ldots N\right\}$ of REK-PF is calculated according to receiver position coordinates $\left(r_x, r_y, r_z\right)$ and process noise $\mathbf{w}$.

2. $k = 1, 2, \ldots$

I. State estimation of every single particle is calculated by REKF, and the priori particle swarm $\{\mathbf{x}_{k/k-1}(i); i = 1, 2, \ldots N\}$ is calculated by Sequential importance sampling.

II. Weight updates.

Substituting the priori particle $\left\{\mathbf{x}_{k/k-1}(i); i = 1, 2, \ldots N\right\}$ and the position coordinates $\left(s_x^i, s_y^i, s_z^i\right)$ of satellite i into observation Eq. (14) to calculate the predicted pseudo-range $\tilde{\rho}_k^i$ of satellite i. Substituting $\tilde{\rho}_k^i$ and the true pseudo-range ρ_k^i into (4) to calculate the weight and normalized it.

III. Resample.

The particle swarm after resampling is $\{\tilde{\mathbf{x}}_{k/k-1}(i), \tilde{\omega}_k(i) = \frac{1}{N}; i = 1, 2, \ldots N\}$.

IV. State estimation at current epoch.

$$\hat{\mathbf{x}}_k = \sum_{i=1}^{N} \tilde{\mathbf{x}}_k^M(i)\tilde{\omega}_k^M(i) \tag{16}$$

V. Fault detection.

Substituting $\hat{\mathbf{x}}_k$ into the observation Eq. (14) to calculate the predicted observations $\hat{y}_k$. Calculating the corresponding residual $d_k^i = \left|\hat{y}_k - \rho_k^i\right|$ of satellite i, and the averaged value of residual $T_k^i = \frac{1}{M}\sum_{j=k-M+1}^{k} d_k^i$ in latest M epochs. Fault detection test statistics is $Tr_k = \max\left(T_k^i\right), (i = 1, 2, \ldots m)$. The detection threshold is τ. If $Tr_k > \tau$, then the system exists fault, or the system is fault-free.

VI. Fault isolation.

If $T_k^i > \tau$, then satellite i has fault which should be isolated.

3. $k = k + 1$, turn to the calculation of next epoch.

4 Simulation

The GPS observation data was collected from JFNG tracking station which located in Jiufeng Township, Wuhan City, China. The observation time was from 27,800 to 28,400 s in the 342 th day of the year 2014. The simulation step is 1 s and the

satellite cutoff angles is 10°. During this time, there are seven GPS satellite that can be observed, and their prn number is 5,15,18,21,24,26,29 respectively. The pseudo-range observations is $\mathbf{Y} = (y_1, y_2, y_3, y_4, y_5, y_6, y_7)$ which obeys Gaussian core-Laplacian distribution [16]. The particle number $N = 100$, the window length $M = 30$ which is used to calculate the residual, and the detection threshold $\tau = 10$, which is chosen as a balance between false alarm rates and miss alarm rates. Add 30 m pseudo-range bias to the satellite 18 from 200 to 500 s. The position accuracy is measured by Root Mean Squared Error (RMSE). When the fault has not been isolate, the RMSE value of particle filter (PF) and REK-PF is shown in Fig. 1.

As shown in Fig. 1, when existing pseudo-range bias, the RMSE value of PF has a significant increase while that of REK-PF has little change. It means that the receiver state estimation can be calculated accurately by REK-PF even if the fault has not been isolate.

The performance of RAIM algorithm based on REK-PF (hereinafter short for REK-PF method) and RAIM algorithm based on PF (hereinafter short for PF method) is compared in the simulation. The fault detection test statistics and the test statistics of every single satellite of two methods is shown in Figs. 2 and 3 respectively.

As shown in Fig. 2, when existing pseudo-range bias, the fault detection test statistics of REK-PF method is higher than that of PF method, which means REK-PF method is more sensitive to failure. The fault detection test statistics of REK-PF method exceeds the detection threshold and give an alarm at 208 s, while the PF method give an alarm at 216 s, which means REK-PF method can detect the

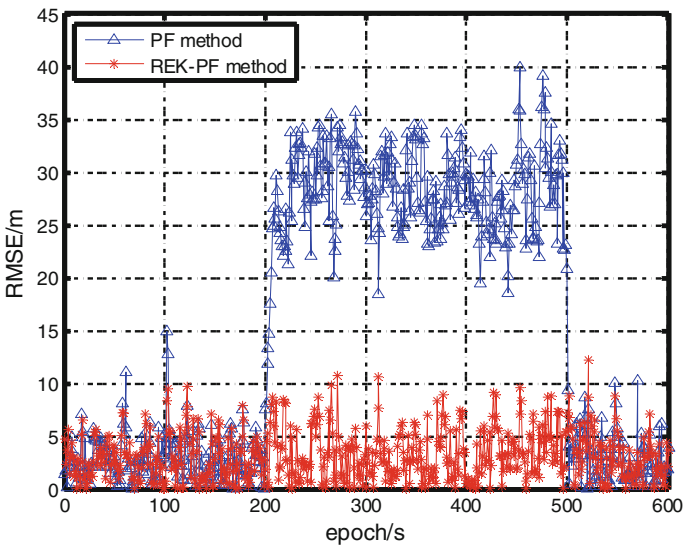


Fig. 1 RMSE value of two methods

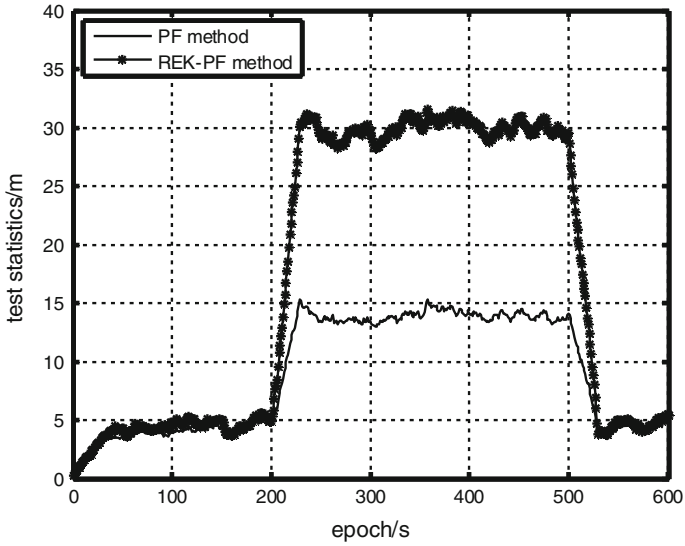


Fig. 2 Fault detection test statistics of two methods

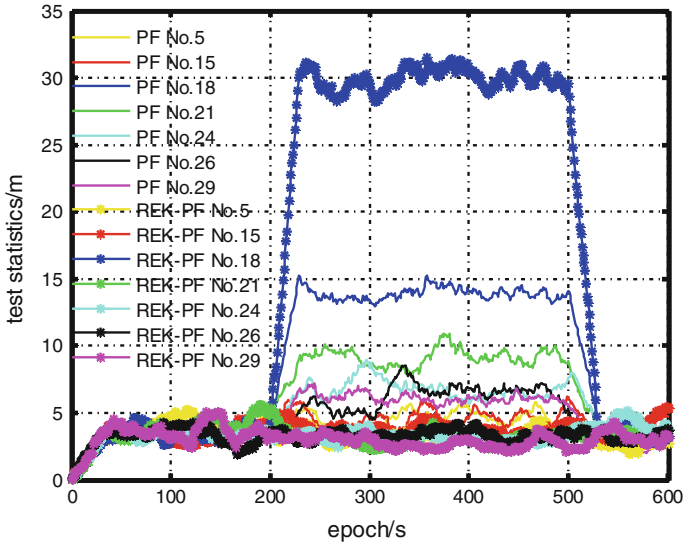


Fig. 3 Test statistics of every satellite of two methods

fault quicker. As shown in Fig. 3, the test statistics of satellite 18 in REK-PF method exceeds the detection threshold while the test statistics of the other satellite stay steady, so that satellite 18 can be judged to be the faulty satellite which should

be isolated directly. On the other hand, the test statistics of every satellite in PF method are all changed, which increase the risk of false alarm. The basic reason of this phenomenon is that the particle state estimation of PF has bias when existing fault, so that the residual is effected by the bias. But REK-PF method has an accurate state estimation because of its robustness, so that the residual of fault-free satellite is less effected by pseudo-range bias while the residual of faulty satellite is more sensitive to the bias.

To further verify the performance of REK-PF method, different pseudo-range bias are added in satellite 18. The average effective number of particles $\bar{N}_{eff}$ and the alarm time of two methods is shown in Table 1. The average effective number of particles $\bar{N}_{eff}$ is calculated as follows.

$$\bar{N}_{eff} = \frac{1}{T}\sum_{i=1}^{T} N_{eff}^{i} \tag{17}$$

where T is the total time in simulation and N_{eff}^{i} is the effective number of particles at epoch i. As shown in Table 1, whether fault-free or existing different pseudo-range bias, the average effective number of particle of REK-PF method are more than that of PF method, and the alarm time of REK-PF method is shorter.

The RMSE value of two methods under different pseudo-range bias are shown in Table 2. As shown in Table 2, whether fault-free or existing different pseudo-range bias, the RMSE value of REK-PF method are smaller than that of PF method. When existing pseudo-range bias, the RMSE value of PF has an obvious increase while that of REK-PF stay small, which means the position accuracy is improved because of the robustness of REK-PF.

The effectiveness of REK-PF method can be testified in Tables 1 and 2.

Table 1 The parameters of two methods in fault detection and isolation

Pseudo-range bias (m)	PF method		REK-PF method	
	Average effective number of particles	Alarm time (s)	Average effective number of particles	Alarm time (s)
0 (fault-free)	29.35		89.22	
20	28.61	228	90.01	213
25	29.06	223	90.04	210
30	28.75	216	89.65	208
35	28.67	213	89.44	207
40	28.57	210	89.19	205

Table 2 RMSE value of two methods under different pseudo-range bias

Pseudo-range bias (m)	RMSE (m)	
	PF method	REK-PF method
0 (fault-free)	4.10	3.64
20	10.24	3.89
25	12.31	3.90
30	14.25	3.93
35	16.95	3.95
40	19.55	3.97

5 Conclusion

RAIM algorithm based on robust extended Kalman particle filter and smoothed residual is proposed in this paper. The importance density function of particle filter is calculated by robust extended Kalman filter, and the fault detection test statistics is conducted by smoothed residual method to achieve the fault detection and isolation of satellite. The simulation results show that, RAIM algorithm based on robust extended Kalman particle filter and smoothed residual can well detect and isolate the faulty satellite under the condition of non-gaussian observation noise. Compared to the RAIM algorithm based on particle filter, the proposed RAIM algorithm increases the average effective number of particles, shortens the alarm time and improves the position accuracy, thus this new RAIM algorithm has a better performance.

Acknowledgements This work was supported by the National Natural Science Foundation, China (No. 61502257 and No. 41304031).

References

1. Elliott DK, Christopher JH (2006) Understanding GPS: principles and applications, 2nd edn. Artech House, MA
2. Ran JH, Jin MZ, Gan XL (2016) A new RAIM algorithm based on multivariate cumulative sum and its improvement. Neurocomputing 171:113–117
3. Alban R, Bexhet K, Shkelzen C et al (2015) Integrity monitoring in navigation systems: fault detection and exclusion RAIM algorithm implementation. J Comput Commun 3:25–33
4. Yun Wu (2010) GNSS fault detection and identification performance comparison of snapshot and filtering. Geomatics Inf Sci Wuhan Univ 35(6):649–652
5. Xiaohao X, Chuansen Y, Ruihua L (2013) Review and prospect of GNSS receiver autonomous integrity monitoring. Acta Aeronaut et Astronaut Sin 34(3):451–463
6. Ershen W, Tao P, Pingping Q, et al (2015) RAIM Algorithm based on particle filter and likelihood ratio method. J Nanjing Univ Aeronaut Astronaut 47(1):46–51
7. Alrowaie F, Gopaluni RB, Kwok KE (2012) Fault detection and isolation in stochastic non-linear state-space models using particle filters. Control Eng Pract 20:1016–1032
8. Arulampalam MS, Maskell S, Gorden N et al (2002) A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. IEEE Trans Signal Process 50(2):174–188

9. Stelio PM, Chris R (2016) Real-time failure detection in the carrier phase measurements of GPS by robust and conventional Kalman filter. Mar Geodesy 21(1):41–65
10. Anton JH (2012) Bayesian estimation and tracking: a practical guide. Wiley, New York
11. Jing H, Zhanrong J, Yan Y (2013) Extended Kalman particle filter algorithm for target tracking in stand-off jammer. J Electron Inf Technol 35(7):1587–1592
12. Yuanxi Y, Xia R, Yan X (2013) Main progress of adaptively robust filter with applications in navigation. J Navig Positioning, 1(1):10–12
13. Jian W, Yang G, Zengke L et al (2016) A tightly-coupled GPS/INS/UWB cooperative positioning sensors system supported by V2I communication. Sensors 16(7):1–16
14. Yuanxi Y, Junyi X (2016) GNSS receiver autonomous integrity monitoring (RAIM) algorithm based on robust estimation. Geodesy Geodyn 7(2):117–123
15. J L, Liyan Q, Xiyuan P (2007) Fault detection based on sir state estimation and smoothed residual. Acta Electron Sin 35(12):32–36
16. Sayim I, Pervan B, Pullen S et al (2002) Experimental and theoretical results on the LAAS sigma overbound. In: Proceeding of 2002 ION GPS, Portland, pp 29–38

Research on RAIM Algorithm Based on GPS/BDS Integrated Navigation

Fuxia Yang, Ershen Wang, Tao Pang, Pingping Qu and Zhixian Zhang

Abstract In multi-constellation integrated navigation system, the multiplication of visible satellites makes the receiver autonomous integrity algorithm (RAIM) applied in approach flight phase of civil aviation become possible. For the GPS/BDS integrated navigation RAIM algorithm applied to approach procedure with vertical guidance (APV) phase of flight, an algorithm based on BDS and GPS positioning solution weighted average solution is proposed. Firstly, the algorithm obtains the weighted average solution satisfying the vertical-guided approach by searching for the optimal weight ratio and uses it as the final solution. Then, the distance between the final solution and BDS, GPS positioning solution is used as the test statistic, and the detection threshold is calculated through the maximum allowable false alarm probability. In the end, the fault is detected by comparing the detection statistic and the detection threshold. By collecting the raw GPS/BDS data, the proposed algorithm was verified. The simulation results demonstrate that the proposed algorithm can effectively detect fault satellites.

Keywords Global positioning system (GPS) · BeiDou Navigation Satellite System (BDS) · Receiver autonomous integrity monitoring (RAIM) · Fault detection

1 Introduction

The receiver autonomous integrity monitoring (RAIM) algorithm is derived from the high reliability requirements of the navigation system of the civil aviation users, which is the ability of the navigation system positioning error to exceed the

F. Yang · E. Wang (✉) · T. Pang · P. Qu · Z. Zhang
School of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang, China
e-mail: wanges_2016@126.com

E. Wang
Liaoning General Aviation Key Laboratory, Shenyang Aerospace University, Shenyang, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_18

allowable error and timely alarm. At present, there are three main ways to realize the integrity monitoring of the navigation system. One is the satellite autonomous integrity monitoring (SAIM); the other is the receiver autonomous integrity monitoring (RAIM), which uses the receiver's internal redundant information; third is the augmentation system. Compared with the other two integrity monitoring methods, the receiver autonomous integrity monitoring based on the redundant information inside the receiver to achieve satellite fault detection and isolation can satisfy the user's requirements for the alarm time [1]. Early receivers' autonomous integrity algorithms were based on GPS single-constellation, and the design goal was to ensure the integrity from the route to the non-precision in the civil aviation flight. However, with the development of global navigation satellite system (GNSS), future civil GNSS receivers can use multi-constellation and multi-frequency satellite signals to position and navigate. The number of participating satellites increase, which can effectively reduce RAIM hole in the civil aviation flight. The RAIM technology can also be applied to some more strict application scenarios, such as vertical guidance approach (APV-I, APV-II) and precision approach (CAT-I).

Multi-constellation RAIM algorithm refers to check the consistency of two or more satellite navigation system redundancy measurement information. The development of GNSS technology brings a lot of opportunities to RAIM. The research on RAIM based on GPS, BDS, GLONASS and Galileo has been done by domestic and foreign scholars. T. Walter proposed the weighted RAIM method [2]. PB Ober and S. Hewitson proposed a method for monitoring integrity in multi-system multi-fault scenarios [3–5]. RAIM of APV-I and APV-II is analyzed for the combination of GPS, Galileo and GLONASS. ENE analyzes the RAIM availability of GPS and Galileo integrated navigation during the vertical-guided approach [6, 7]. Martineau et al. conducted a hypothesis analysis of GPS and Galileo integrated RAIM in cutoff height, user equivalent range error and surface grid resolution [8], and Xu et al. conducted a global survey of the vertical protection level in the vertical guidance approach [9]. Yong C et al. proposed the optimal weighted average solution (OWAS) algorithm [10], and domestic researchers also extended this algorithm. Guo Jing et al. expanded its use of new grouping method to meet the requirement of the double constellation double fault application scenario [11]. The algorithm of optimal weighted average solution can easily and quickly be compatible with different constellations, which avoids the error caused by the conversion of different time and coordinate systems. It has the characteristics of fast processing speed and high availability. Therefore, the optimal weighted average solution method is used to study the BDS and GPS combined navigation in the vertical guidance approach.

This paper studies the OWAS method. In the position domain, the distance between the grouping position solution and the weighted average solution is taken as the test statistic. The detection thresholds for fault detection are constructed by using the maximum false alarm probability that OWAS double constellation RAIM allow. However, in the single constellation fault identification, this method is affected by the fault constellation positioning error, and the fault-free constellation

will generate false alarm. Therefore, the paper completes fault satellite detection and identification by the weighted least squares residual method. Simulation results show that the algorithm availability supports the APV1 and APV2 vertical guidance approach phase and the hypothesis that the GPS/BDS dual system simultaneously fault.

2 The Optimal Weighted Average Solution (OWAS) RAIM Algorithm

The optimal weighted average solution (OWAS) RAIM algorithm is a typical representative of the multi-constellation receiver integrity algorithm in recent years. It has been developed for application to a combined GPS/Galileo position solution that would be used for vertical guidance approach. Base on two existing RAIM methods: group separation (GS) algorithm [12] and weighted NIORAIM (Novel Integrity Optimized RAIM) algorithm [13]. Their principle is that algorithm separates visible satellite into two groups by different navigation system to respectively calculate the position results.

2.1 The Derivation of Covariance Matrix

The main words in all headings (even run-in headings) begin with a capital letter. Articles, conjunctions and prepositions are the only words which should begin with a lower case letter. The final solution of the algorithm is obtained by weighting average of GPS and BDS position solution, as follows:

$$x_A = \frac{w_1 x_1 + w_2 x_2}{w_1 + w_2} = \frac{w_1}{w_1 + w_2} x_1 + \frac{w_2}{w_1 + w_2} x_2 = r x_1 + (1 - r) x_2 \tag{1}$$

where $r \equiv w_1/(w_1 + w_2)$ and $0 \leq r \leq 1$.

The relationship between the position error of the optimal solution and the weighting factor is as follows:

$$\Delta x_A = r \Delta x_1 + (1 - r) \Delta x_2 \tag{2}$$

The general expression of the position error of the weighted least squares solution is as follows:

$$\Delta x_n = (G_n^T W_n G_n^T)^{-1} G_n^T W_n \Delta R_n \tag{3}$$

In Eq. (3), ΔR_n is the pseudo-range error vector; W_n is the inverse matrix of pseudo-range error covariance matrix; G_n is the correlation matrix of pseudo-range

error and the receiver position, clock; n is the solving equation of nth navigation system. So the covariance equation that the separation distance error between the optimal weighted mean solution and each navigation system position solution is as follows:

$$dP_n \equiv E[(\Delta x_n - \Delta x_A)(\Delta x_n - \Delta x_A)^T] = E[\Delta x_n \Delta x_n^T - \Delta x_A \Delta x_n^T - \Delta x_n \Delta x_A^T + \Delta x_A \Delta x_A^T] \quad (4)$$

Equation (4) can be deduced by the following:

$$P_n \equiv E[\Delta x_n \Delta x_n^T] = (G_n^T W_n G_n^T)^{-1} \quad (5)$$

$$E[\Delta x_A \Delta x_n^T] = rP_1 + (1 - r)E[\Delta x_2 \Delta x_1^T] \quad (6)$$

Assuming GPS and BDS pseudo-range error is not relevant, then there is $E[\Delta R_2 \Delta R_1^T] = 0$. Thus, Eq. (7) can be derived.

$$E[\Delta x_2 \Delta x_1^T] = (G_2^T W_2 G_2^T)^{-1} G_2^T W_2 E[\Delta R_2 \Delta R_1^T] W_1 G_1^T (G_1^T W_1 G_1^T) = 0 \quad (7)$$

Therefore,

$$E[\Delta x_A \Delta x_1^T] = E[\Delta x_1 \Delta x_A^T] = rP_1 \quad (8)$$

Similarly,

$$P_A = E[\Delta x_A \Delta x_A^T] = r^2 P_1 + (1 - r)^2 P_2 \quad (9)$$

Through Eq. (5), (8) and (9), it can be gotten as follows:

$$dP_1 = r^2 P_1 + (1 - r)^2 (P_1 + P_2) \quad (10)$$

$$dP_2 = E[(\Delta x_2 - \Delta x_A)(\Delta x_2 - \Delta x_A)^T] = r^2 (P_1 + P_2) \quad (11)$$

2.1.1 The Weighting Ratio Determines the Protection Level

First, two parameters are defined as

$$K_{\text{fd}} \equiv Q^{-1}(P_{fd}/4) \quad (12)$$

$$K_{md} \equiv Q^{-1}(P_{md}) \quad (13)$$

where, Q^{-1} is the inverse of the complement of the one-sided standard normal cumulative distribution function (CDF). P_{fd} is a maximum allowable false alert

probability(Dual constellation including horizontal and vertical distribution need to divide by 4). P_{md} is a maximum missed detection probability.

Further define,

$$\sigma_{V,1} \equiv \sqrt{P_1(3,3)} \tag{14}$$

$$\sigma_{V,2} \equiv \sqrt{P_2(3,3)} \tag{15}$$

$$\sigma_{V,S} \equiv \sqrt{P_S(3,3)} = \sqrt{P_1(3,3) + P_2(3,3)} \tag{16}$$

Then, based on Eq. (10) and (11), it can be gotten:

$$\sigma_{dV,1} \equiv \sqrt{dP_1(3,3)} = (1-r)\sqrt{P_S(3,3)} = (1-r)\sigma_{V,S} \tag{17}$$

$$\sigma_{dV,2} \equiv \sqrt{dP_2(3,3)} = r\sqrt{P_S(3,3)} = r\sigma_{V,S} \tag{18}$$

The vertical protection level (VPL_1 and VPL_2) is the sum of the vertical protection thresholds and the vertical positioning error upper limit under the assumption of no fault detection, so VPL_1 and VPL_2 are written as:

$$VPL_1 = Td_{V,1} + a_{V,1} = K_{\mathrm{fd}}\sigma_{dV,1} + K_{\mathrm{md}}\sigma_{V,1} = K_{\mathrm{fd}}(1-r)\sigma_{V,S} + K_{\mathrm{md}}\sigma_{V,1} \tag{19}$$

$$VPL_2 = Td_{V,2} + a_{V,2} = K_{\mathrm{fd}}\sigma_{dV,2} + K_{\mathrm{md}}\sigma_{V,2} = K_{\mathrm{fd}}r\sigma_{V,S} + K_{\mathrm{md}}\sigma_{V,2} \tag{20}$$

Set $VPL_1 = VPL_2$, we can get,

$$r_{\min\_VPL} = \frac{1}{2} + \frac{1}{2}\left(\frac{K_{md}}{K_{fd}}\right)\left(\frac{\sigma_{V,1} - \sigma_{V,2}}{\sigma_{V,S}}\right) \tag{21}$$

$$VPL = \frac{1}{2}K_{fd}\sigma_{V,S} + \frac{1}{2}K_{md}(\sigma_{V,1} + \sigma_{V,2}) \tag{22}$$

The weight ratio in Eq. (21) gives the smallest VPL, and the weight ratio that minimizes HPL can also be derived in a similar manner.

2.1.2 The Weighting Ratio Determines the Error Precision σ_V and d_{major}

The covariance matrix of the weighted average solution position error, P_A, was showed in Eq. 23

$$P_A = E[\Delta x_A \Delta x_A^T] = r^2 P_1 + (1-r)^2 P_2 \tag{23}$$

σ_V and d_{major} are given, respectively, by:

$$\sigma_V = \sqrt{P_A(3,3)} \tag{24}$$

$$d_{major} = \sqrt{\frac{P_A(1,1)+P_A(2,2)}{2} + (\sqrt{\frac{P_A(1,1)+P_A(2,2)}{2}})^2 + P_A(1,2)^2} \tag{25}$$

RAIM is available if σ_V and d_{major} meet the bellow conditions:

$$\sigma_V \leq \sigma_{V-\max} \quad d_{major} \leq d_{major-\max} \tag{26}$$

In the calculation of P_A, σ_V and d_{major}, separate weight ratios are applied to vertical and horizontal.

2.2 Fault Detection

The optimal weighted average solution RAIM algorithm is based on the distance between GPS/BDS and the optimal weighted average solution as the fault detection test statistics. When both of the double-constellation appears faults, two test statistics will exceed the detection threshold at the time of fault occurrence. Because of fault detection based on position domain, the fault can be judged effectively when multiple failures occur in each constellation; when the single constellation fault, the fault constellation solution makes the distance between weighted average solution and fault-free constellation increasing, and this will lead to the occurrence of false alarms. In this case, the fault detection based on weighted least squares residuals can be used to assist the fault detection model in the OWAS localization domain to detect single constellation faults.

According to the linearization equation of GNSS pseudo-range observation, the equation is solved by weighted least squares estimation so that the solution is

$$\hat{X}_{WLS} = \left(H^T W H\right)^{-1} H^T W y \tag{27}$$

Among them, W is weight matrix. The pseudo-range residual vector v is:

$$v = y - H\hat{X}_{WLS} = \left(I_n - H\left(H^T W H\right)^{-1} H^T W\right) y \tag{28}$$

The sum of squares of the pseudo-range residuals is:

$$SSE_{WLS} = v^T v = y^T W y \tag{29}$$

When we assume that the observed noise is the normal distribution, SSE_{WLS} has a χ^2 distribution with (n − 4) freedom degrees. Therefore, make binary hypothesis for SSE_{WLS}. Give false alarm rate P_{fa} and probability density function, then we use Eq. (30) to obtain the detection threshold T_{DWLS}.

$$P_r(SSE_{WLS} > t_{WLS}) = 1 - \int_0^{t_{WLS}} f_{\chi^2(n-4)}(x)dx = \int_{t_{WLS}}^{\infty} f_{\chi^2(n-4)}(x)dx = P_{FA} \quad (30)$$

Set $T_{WLS} = \sqrt{SSE_{WLS}/(n-4)}$ as the detection statistic, corresponding to the detection threshold is $T_{DWLS} = \sqrt{t_{WLS}/(n-4)}$. Compare the detection statistic T_{WLS} and the detection threshold T_{DWLS}. If $T_{WLS} < T_{DWLS}$, we can judge that there is no fault satellite; otherwise, there is fault satellite.

3 Experiment and Simulation Analysis

This simulation is based on the GPS/BDS double constellation RINEX data of coordinates [−2279827.3156, 5004704.3094, 3219776.2093] collected at 0:00 on July 31, 2016, and sampled 900 times (15 min, once per second). In the RAIM algorithm availability analysis, the vertical guidance approach is emphatically verified. There is no failure satellite in the raw sampling. In order to verify fault detection algorithm, different deviations are intentionally added to the pseudo-range observations at different times.

3.1 RAIM Availability Analysis

Table 1 describes the ICAO requirements for vertical guidance approach, precision approach phase, alarm limits and availability [14]. The algorithm availability under different alarm limits is shown as Fig.1.

Table 1 Performance requirements for different flight phases

Flight phase	Vertical alarm threshold (m)	Algorithm availability (%)
APV-I	50	99–99.999
APV-II	20	99–99.999
CAT-I	10–15	99–99.999

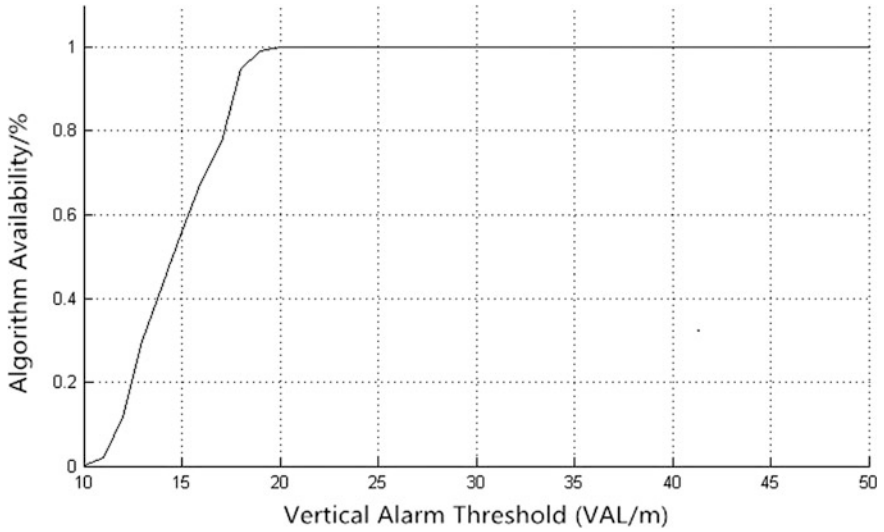


Fig. 1 Algorithm availability

Table 2 Algorithm availability statistics

Flight phase	Vertical alarm threshold (m)	Algorithm availability (%)
APV-I	50	100
APV-II	20	99.89
CAT-I	15	57

When simulating the availability of a vertical guidance approach and precision approach phase, the horizontal alarm thresholds are set to meet the minimum alarm threshold of 40 m for the three flight phases. As can be seen from Table 2, when the vertical alarm threshold is 20 m under the APV-II standard, the algorithm availability is 99.89%. With the increase of the vertical alarm threshold, the availability of the algorithm is improved. With 99% availability as the boundary, the vertical alarm threshold is at least 19 m.When the vertical alarm threshold is 19 m, the algorithm availability is 99.11%. Table 2 summarizes the availability of the algorithm in different flight phases. As Table 2 shown, under the GPS/BD dual constellation positioning mode, the optimal weighted average solution (OWAS) algorithm can support the APV-I and APV-II stages in the vertical guidance approach flight phase. CAT-I in the precision approach cannot support it, but it can be predicted. After adding the land-based augment, when this algorithm is applied to the precision approach stage, it will have better effects.

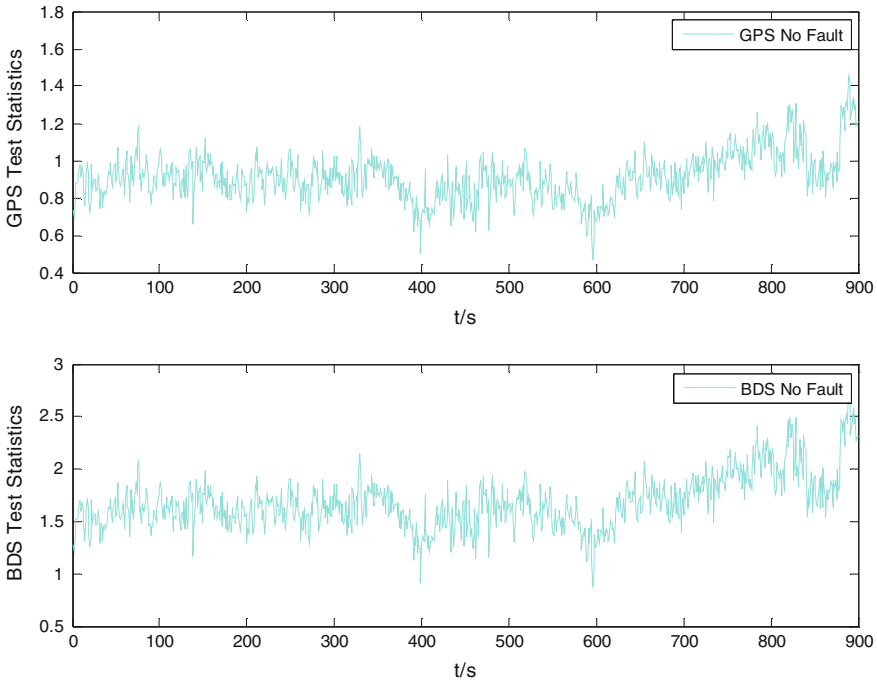


Fig. 2 GPS/BDS test statistics results under no pseudo-range bias

3.2 Fault Detection

In the progress of verifying the fault detection, the simulation environment is as described above. In order to verify the ability of the algorithm to detect double constellation faults, in the 16th satellite of GPS constellation and the 7th satellite of BDS constellation is added to 30, 50, 70 m pseudo-range deviation at 500–900 s.

As shown in Figs. 2 and 3, Fig. 2 shows the vertical position field test statistics of GPS/BDS without the pseudo-range deviation, and test statistics at this time is small and the change is not significant. Figure 3 shows the test statistics of GPS and BDS at the beginning of 500 s adding 30, 50 and 70 m pseudo-range deviation. It can be seen from the picture that when the pseudo-range bias is added, the test statistic increases instantaneously and exceeds the detection threshold. The greater the pseudo-range bias is, the more obvious the test statistic increases. According to above Figs. 2 and 3, we can know that the test statistic of the fault detection is stably maintained in a small range, which is far from the detection threshold. When the pseudo-range bias is added, and the test statistic exceeds the detection threshold, this way can be used to effective judge the fault.

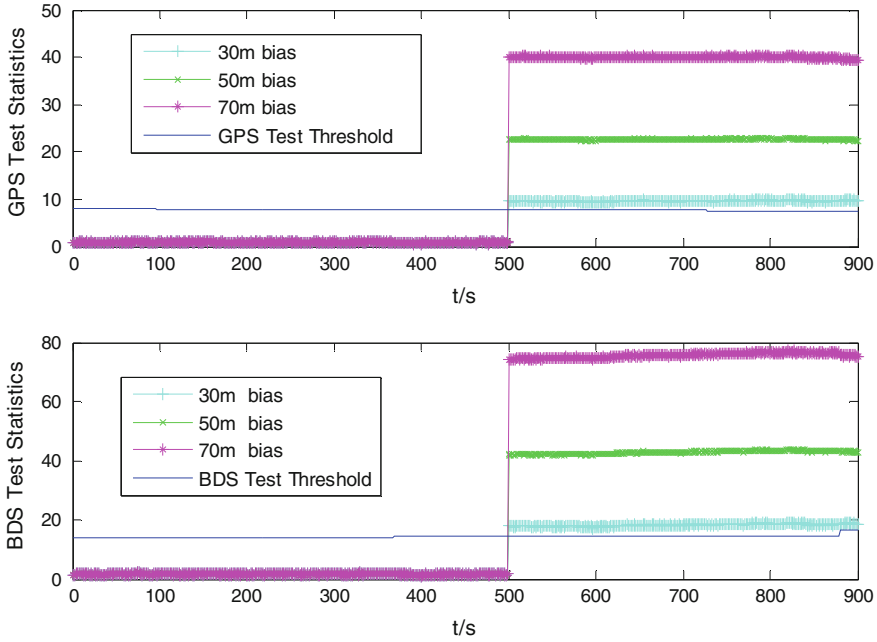


Fig. 3 GPS/BDS test statistics results under different pseudo-range bias

4 Conclusions

In the paper, a kind of double-constellation RAIM algorithm based on optimal weighted average solution (OWAS) is presented. According to the constellations, this algorithm divided visible satellites into two groups. In the premise of ensuring the vertical accuracy, the algorithm minimizes the vertical protection level, and makes it meet the requirements of the vertical guidance approach flight phase. Among them, the availability of APV-I phase algorithm is 99.89%, and the availability of APV-II phase algorithm is 100%. Through the simulation results based on the raw measurement data, this algorithm can be used to process the fault detection of double constellation. When the pseudo-range bias is more than 30 m, it can effectively detect the fault satellites. Meanwhile, the results are instructive for the study of integrity monitoring for the BDS and other satellite navigation system in civil aviation.

Acknowledgements This study is funded by the National Natural Science Foundation of China (61571309 & 61101161), and the Liaoning BaiQianWan Talents Program.

References

1. Wang E, Yang Y, Zhang Z et al (2015) Availability analysis of GPS receiver autonomous integrity monitoring algorithm. J Dalian Marit Univers 01:91–95
2. Walter T, Enge P (1995) Weighted RAIM for precision approach. In: Proceedings of ION GPS-95, Palm Spings, CA, Sept 1995, pp 12–15
3. Hewitson S, Wang J (2006) GNSS receiver autonomous integrity monitoring (RAIM) performance analysis. GPS Solutions 10(3):155–170. doi:10.1007/s10291-005-0016-2
4. Hewitson S, Wang J (2006) GNSS receiver autonomous integrity monitoring (RAIM) for multiple outliers. Eur J Navig 10:155–170
5. Ober PB, Harriman D (2006) On the use of multi-constellation-RAIM for aircraft approaches. In: Proceedings of 2006 ION GNSS, ION, Fort Worth, TX, pp 2587–2596
6. ENE A (2006) Further development of Galileo-GPS RAIM for vertical guidance. In: Proceedings of ION GNSS 2006, Ft. Worth, TX, vol 52(1), 26–29 Sept 2006, pp 432–440
7. Een A, Blanch J, Walter T (2006) Galileo-GPS RAIM for vertical guidance. In: Proceedings of the ION NTM. USA, 2006, vol 52(1), pp 432–440
8. Martineau A, Macabiau C, Mabilleau M (2009) GNSS RAIM assumptions for vertically guided approaches. In: Proceedings of GNSS 2009, 22nd international technical meeting of the satellite division of the institute of navigation. USA, 2009, vol 5538(1), pp 2791–2803
9. Xu J, Yang Y, Li J et al (2013) Integrity analysis of COMPASS and other GNSS combined navigation. Sci China Ear Sci 56:1616–1622
10. Lee Y, Braff R, Fernow J. et al (2005) GPS and Galileo with RAIM or WAAS for vertically guided approaches. In: Proceedings of the institute of navigation GNSS 2005. Institute of Navigation, Long Beach, USA, pp 1801–1825
11. Guo J, Cui X, Lu M, Feng Z (2011) Multi-constellation RAIM for approach with vertical guidance. J Tsinghua Univ (Sci Technol) 2:157–160
12. Lee Y (2004) Investigation of extending receiver autonomous integrity monitoring (RAIM) to combined use of Galileo and modernized GPS. In: Proceedings of the 17th international technical meeting of the satellite division of the institute of navigation, ION GNSS 2004, Long Beach, California, 21–24 Sept 2004, pp 1691–1698
13. Hwang P, Brown R (2006) RAIM FDE revisited: a new breakthrough in availability performance with NIORAIM (novel integrity-optimized RAIM). In: Proceeding of the institute of navigation NTM 2005. Institute of navigation, San Diego, USA, vol 53(1), pp 41–51
14. Hegarty C (2008) Evolution of the global navigation satellite system (GNSS). Proc IEEE 96(12):1902–1917

RAIM Algorithm Based on Residual Separation

Zhaoyang Li, Qingsong Li and Jie Wu

Abstract With the rapid development of Global Navigation Satellite Systems (GPS, BDS, GALILEO and GLONASS), the number of available satellites is increasing. The redundancy of GNSS is enhanced. This will offer the possibility of realizing receiver fault detection and exclusion (FDE). Due to factors such as satellite fault and strong electromagnetic interference, measurements from some satellite are vulnerable to become outliers. Unacceptable positioning errors will occur unless such faulty measurements are detected and excluded. Therefore, it is very important to apply RAIM (Receiver Automatic Integrity Monitoring) algorithm for improving the positioning precision. The main function of RAIM algorithm includes two aspects: fault detection and fault exclusion. This paper presents a RAIM algorithm based on residual separation. The algorithm uses the residual sum of squares of observation equations as test statistics. When test statistics exceed alarm threshold, it can be considered that there are faults in measurements. And then the visible satellites are excluded at every turn, while multi-combinations of remaining measurements are obtained. After SSE (Sum of Squared Error) calculated, the combination of remaining measurements with smallest SSE is the optimal result. Field test data and simulation data are processed in this work. The RAIM algorithm based on residual separation can not only avoid excluding too many satellites, but also improve the positioning accuracy in the case of containing faults.

Keywords Residual-based separation · Fault detection · Fault exclusion · RAIM

1 Introduction

Single point positioning (SPP) is the simplest and most convenient GNSS positioning technique. However, the positioning accuracy is poor and sometimes even unacceptable, especially when measurements contain faults [1]. Therefore, it is

Z. Li (✉) · Q. Li · J. Wu
National University of Defense Technology, College of Aerospace Science and Engineering, Changsha 410073, China
e-mail: 18684890459@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_19

essential to detect and exclude faults with integrity monitoring algorithm. At present, there are mainly two methods to monitor the integrity of navigation system. One method is using the Integrity Augmentation System, such as the WAAS (Wide Area Augmentation System) in the USA, the EGNOS (European Geostationary Navigation Overlay Service) in Europe and the BeiDou Augmentation System in China. The other is the Receiver Automatic Integrity Monitoring (RAIM), which exploits the measurement redundancy from receivers. The Integrity Augmentation System cannot monitor the faults in receivers and the alarm message may not send to users in time. Therefore, it is very important to apply RAIM algorithm in SPP, which is essential to ensure the availability of the positioning results [2].

RAIM was first proposed by Kalafus [3] in 1987. The RAIM algorithm mainly completes two tasks: detecting whether there are faults in the measurements, identifying and excluding the faulty satellites. At present, the fault exclusion methods such as data snooping [3] and parity space vector [3] are extensively used. And they are equivalent [3]. The data snooping presented by Baarda constructs test statistics with elements of residual vector and compares them with threshold one by one. If some of the test statistics exceed the threshold, it can be considered that the corresponding measurements contain faults, which should be excluded. The Baarda method tests the measurements individually, which may exclude the non-faulty satellites, leading to decreased positioning precision [4]. Furthermore, it may even cause the available satellites less than four, resulting in no positioning solutions [5].

This paper proposes a residual-based separation RAIM. First, the fault detection algorithm gives an alarm. Then an available satellite is excluded at every turn, and the SSE of remaining satellites is calculated. If the smallest SSE is less than alarm threshold, the corresponding remaining satellites combination is selected to carry out positioning calculation. Otherwise the next exclusion turn will be executed, until the SSE of remaining satellites is less than alarm threshold. Because only one available satellite is excluded at every turn, it is impossible to exclude too many satellites, which can effectively decrease faulty exclusion frequency and improve the positioning accuracy.

The rest of the paper is organized as follows. Section 2 gives the mathematical model of single point positioning, fault detection method based on least square residuals, fault exclusion method including Baarda method and Residual-based method. Section 3 verifies the performance of Baarda method and RBS method. Section 4 gives the conclusion.

2 Model and Algorithms

2.1 Mathematical Model of SPP

The pseudorange measurement equations containing the receiver clock bias, satellite clock bias and atmospheric propagation delay can be expressed as follows [2]:

$$\rho_k^j = \left|\boldsymbol{r}^j - \boldsymbol{r}_k\right| + c \cdot \delta t_k - c \cdot \delta t^j + \delta\rho_{kion}^j + \delta\rho_{ktro}^j + \varepsilon_k^j \tag{1}$$

where, superscript j is the identifier of satellites, $j = 1, 2, \ldots, n$, n is the number of satellites, and $n \geq 4$; subscript k is the identifier of receivers; ρ_k^j is the pseudorange measurement from receiver k to satellite j; $\boldsymbol{r}^j$ is the position vector of the satellite j, which can be obtained from satellite ephemeris and launch time of signal; $\boldsymbol{r}_k$ is the unknown position vector of receiver k; c is the speed of light; δt_k is the clock bias of receiver k, which is a unknown parameter; δt^j is the clock bias of satellite j; $\delta\rho_{kion}^j$ is the ionospheric delay; $\delta\rho_{ktro}^j$ is the tropospheric delay; ε_k^j is random pseudorange measurement noise.

To solve the pseudorange measurement equations, the equation need to be linearized firstly. Assuming that the initial value of the receiver's position vector is $\boldsymbol{r}_{k0}$, the pseudorange measurement equations can be linearized as [2]:

$$\rho_k^j = |\boldsymbol{r}^j - \boldsymbol{r}_{k0}| - \boldsymbol{r}_k^{j0} \cdot \mathrm{d}\boldsymbol{r}_k - c \cdot \delta t^j + c \cdot \delta t_k + \delta\rho_{kion}^j + \delta\rho_{ktro}^j + \varepsilon_k^j \tag{2}$$

where $\mathrm{d}\boldsymbol{r}_k$ is the increment of the receiver's position vector which is unknown; $\boldsymbol{r}_k^{j0}$ is the unit vector from k receiver to j satellite.

2.2 Fault Detection Method Based on Residual Vector

The position of the receiver and receiver clock bias are obtained using above standard SPP algorithm. However,in case of fewer visible satellites, the faulty measurement of the individual satellite will have a large impact on the positioning results. In this paper,the residual vector of least squares estimation is used to detect fault.

The pseudorange residual vector of least square estimation is calculated as:

$$\boldsymbol{v} = \boldsymbol{Y} - \boldsymbol{A}\hat{\boldsymbol{X}} = (\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{P}\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{P})\boldsymbol{E} \tag{3}$$

Define the cofactor matrix as:

$$\boldsymbol{Q}_{\boldsymbol{v}} = \boldsymbol{P}^{-\mathbf{1}} - \boldsymbol{A}(\boldsymbol{A}^{\mathbf{T}}\boldsymbol{P}\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{A}^{\mathbf{T}} \tag{4}$$

Thus, the pseudorange residual vector can be rewritten as:

$$\boldsymbol{v} = \boldsymbol{Q}_{\boldsymbol{v}}\boldsymbol{P}\boldsymbol{E} \tag{5}$$

And the corresponding mean square error:

$$\hat{\sigma} = \sqrt{\boldsymbol{v}^{\boldsymbol{T}}\boldsymbol{P}\boldsymbol{v}/(n-4)} = \sqrt{SSE/(n-4)} \tag{6}$$

where *SSE* means sum of squared error.

When the system is normal, all the elements of the pseudorange residual vector should follow Gauss distribution with zero mean and σ_0^2 variance; when the measurement of a satellite is abnormal, some elements of pseudorange residual vector will no longer follow zero mean, which is the case that need to be detected.

Therefore, according to the statistical theory, SSE/σ_0^2 should follow χ^2 distribution with $n-4$ degree of freedom in the absence of fault. Dualistic hypothesis can be proposed [3]:

Hypothesis with fault H_0: $E(\varepsilon) = 0$, thus, $SSE/\sigma_0^2 \sim \chi^2(n-4)$;

Hypothesis without fault H_1: $E(\varepsilon) \neq 0$, thus, $SSE/\sigma_0^2 \sim \chi^2(n-4, \lambda)$;

In the absence of fault, the *SSE* of the pseudorange residual should maintain at a minor value. If there are accidental factors which cause *SSE* too large, it should be false alarmed. With given false alarm rate P_{FA}, the following formula can be got:

$$P(SSE/\sigma_0^2 \leq T^2) = \int_0^{T^2} f_{\chi^2(n-4)}(x)dx = 1 - P_{FA} \tag{7}$$

The alarm threshold can be obtained from the above formula, in which the probability density function of the χ^2 distribution with $n-4$ degree of freedom is expressed as $f_{\chi^2(n-4)}(x)$. If mean square error $\hat{\sigma} < \sigma_0 * T/\sqrt{n-4}$, it can be considered that the system is working normally [6]. Otherwise, it is considered that there is fault in measurements which should be alarmed [7].

2.3 Fault Exclusion Method

When the GNSS navigation system is used as an auxiliary, the GNSS measurement information should be no longer used after alarming. But if there is no other auxiliary system, the fault in GNSS system needs to be excluded. The faults among measurements needs to be detected with appropriate algorithm, and the corresponding satellite has to be removed from the SPP observation model. At last, navigation results which meet the users' demand can be obtained.

2.3.1 Baarda Method

Baarda [3] proposed a method of fault exclusion based on elements of least square residual. The test statistics are constructed by elements in residual vector. Then the

identifier of satellite with fault is determined by the size of statistics. After that, measurement of the satellite with fault is removed. Finally, the positioning result without fault is obtained after iteratively calculating.

By using pseudorange residual error vector and the coordinated factor matrix, Baarda's method constructs the statistics as:

$$d_i = \frac{v_i}{\sigma_0 \cdot \sqrt{Q_{v_{ii}}}} \tag{8}$$

If the measurement of the ith satellite is normal, ε_i should follow Gauss distribution with zero mean and σ_0^2 variance. Thus, v_i and d_i should follow the Gauss distribution. If the absolute value of d_i is too large, it can be considered that the ith satellite contains fault, and should be excluded from the measurement equations.

Dualistic hypothesis about statistics d_i corresponding to the ith satellite can be proposed as:

Hypothesis with fault H_0: $E(\varepsilon) = 0$, $d_i \sim N(0, 1)$

Hypothesis without fault H_1: $E(\varepsilon) \neq 0$, $d_i \sim N(\delta_i, 1)$

where, δ_i is the zero offset of statistic d_i. Assume that false alarm rate of the system is P_{FA}, and false alarm rate of each satellite is P'_{FA}. Then the entire system without alarming is equivalent to all of the satellites without alarming, that is $1 - P_{FA} = (1 - P'_{FA})^n$. Thus, $P'_{FA} \approx P_{FA}/n$. Assume that the fault threshold T_d of d_i meets the need of false alarm rate P'_{FA}. Therefore,

$$P(|d_i| > T_d) = \frac{1}{\sqrt{2\pi}} \int_{T_d}^{\infty} e^{-\frac{x^2}{2}} dx = \frac{1}{2} \cdot \frac{P_{FA}}{n} \tag{9}$$

In the course of the positioning calculation, if $|d_i| > T_d$, it is considered that the pseudorange measurement of the ith satellite contains fault, which should to be rejected. If $|d_i| \leq T_d$, it is considered that the satellite is normal and should to be retained. In the end, measurement equations are constructed by all the reserved satellites, which can be used to obtain a satisfactory navigation solution [3].

2.3.2 Residual-Based Separation (RBS) Method

The Baarda method is likely to exclude satellites without faults, which will decrease the positioning precision. Thus, this paper proposes a residual-based separation method, which determines the identifier of faulty satellites with residual sum of squares and excludes only one satellite each time.

After excluding the first satellite, the pseudorange measurement equations are changed to:

$$\underset{(n-1)\times 1}{\boldsymbol{Y}} = \begin{bmatrix} \rho_k^2 - |\boldsymbol{r}^2 - \boldsymbol{r}_{k0}| + c \cdot \delta t^2 - \delta\rho_{kion}^2 - \delta\rho_{ktro}^2 \\ \vdots \\ \rho_k^n - |\boldsymbol{r}^n - \boldsymbol{r}_{k0}| + c \cdot \delta t^n - \delta\rho_{kion}^n - \delta\rho_{ktro}^n \end{bmatrix} \tag{10}$$

$$\underset{(n-1)\times 4}{\boldsymbol{A}_1} = \begin{bmatrix} (-\boldsymbol{r}_k^{20})_{1\times 3}^T & 1 \\ \vdots & \vdots \\ (-\boldsymbol{r}_k^{n0})_{1\times 3}^T & 1 \end{bmatrix}, \quad \underset{4\times 1}{\boldsymbol{X}_1} = \begin{bmatrix} \mathrm{d}\boldsymbol{r}_k \\ c \cdot \delta t_k \end{bmatrix}, \quad \boldsymbol{E}_1 = \begin{bmatrix} \varepsilon_k^2 \\ \vdots \\ \varepsilon_k^n \end{bmatrix} \tag{11}$$

The estimated value of the unknown parameters is:

$$\hat{\boldsymbol{X}}_\mathbf{1} = (\boldsymbol{A}_\mathbf{1}^\mathrm{T}\boldsymbol{P}_\mathbf{1}\boldsymbol{A}_\mathbf{1})^{-\mathbf{1}}\boldsymbol{A}_\mathbf{1}^\mathrm{T}\boldsymbol{P}_\mathbf{1}\boldsymbol{Y} \tag{12}$$

where,

$$\boldsymbol{P}_1 = diag(\sigma_2^{-2}, \sigma_3^{-2}, \ldots \sigma_n^{-2}) \tag{13}$$

The residual vector becomes:

$$\boldsymbol{v}_\boldsymbol{1} = \boldsymbol{Y}_\boldsymbol{1} - \boldsymbol{A}_\boldsymbol{1}\hat{\boldsymbol{X}}_\boldsymbol{1} = (\boldsymbol{I} - \boldsymbol{A}_\boldsymbol{1}(\boldsymbol{A}_\boldsymbol{1}^\boldsymbol{T}\boldsymbol{P}_\boldsymbol{1}\boldsymbol{A}_\boldsymbol{1})^{-\boldsymbol{1}}\boldsymbol{A}_\boldsymbol{1}^\boldsymbol{T}\boldsymbol{P}^\boldsymbol{1})\boldsymbol{E}_\boldsymbol{1} \tag{14}$$

After rejecting the first satellite, the residual sum of squares is:

$$SSE_1 = \boldsymbol{v}_1^T\boldsymbol{P}_1\boldsymbol{v}_1 \tag{15}$$

Similarly, after rejecting the second to nth satellite, all of the residual sum of squares can be obtained: $SSE_2, SSE_3, \ldots SSE_n$. Selecting the smallest *SSE*, it can be considered that the corresponding measurement of satellite contains fault. Then, a satisfactory positioning result can be obtained after excluding the satellite.

If the smallest *SSE* still exceeds the alarm threshold after excluding one faulty satellite, the next exclusion round should be carried out as aforementioned method, until the *SSE* is smaller than alarm threshold.

3 Algorithms Verifying

3.1 Simulation Data Processing

The Baarda method and RBS method are both fault exclusion algorithm and fault detection method are both based upon residual vector. Therefore, the miss detection

Table 1 Comparison of faulty exclusion frequency between two methods

Fault (m)	8	9	10	11	12	13
RBS (epochs)	0	5	7	17	24	33
Baarda (epochs)	0	5	25	117	320	660
Fault (m)	14	16	18	20	22	40
RBS (epochs)	24	16	5	2	1	0
Baarda (epochs)	1126	2235	3234	3960	4868	10,506

frequencies of the two methods are equivalent and the faulty exclusion frequencies are compared in this section.

In order to verify the effect of the two methods systematically, a part of normal data including 11,904 epochs was intercepted. Fault was artificially added in the measurement of a satellite. If measurements without faults are excluded in one epoch, it is considered that fault exclusion appears once. The faulty exclusion frequency of two methods was counted. Increase the faults added at every round, and the results are shown in Table 1.

From the table we can see that fault detection algorithm doesn't alarm when the fault added is tiny, so that the faulty exclusion frequency of the two methods are both close to zero; When the fault added is at around 13 m and closes to the normal measurement error, the faulty exclusion frequency of RBS method reaches maximum; When the fault added is larger than 13 m, the faulty exclusion frequency of RBS method is gradually reduced to zero, while the faulty exclusion frequency with Baarda method increases significantly. As a result, it can be considered that RBS method can reduce the faulty exclusion frequency.

3.2 Field Data Processing

From October 20 to 21, 2015 and in Changsha City, Hunan Province, static observation data of 23.7 h was collected. By processing the real measured data, a group of SPP solutions were obtained. Regarding the result of PPP as true value, the output is the error of SPP under three directions in the local geographic coordinate system. The positioning accuracy is evaluated with the RMSE (Root mean square error).

In this paper, the SPP program uses the modified Hopfield model to compensate the tropospheric error, uses the double frequency measurements to compensate the ionospheric error, and rejects the satellite with elevation less than 5°.

From the error curve in Fig. 1, it can be seen that in most of the time single point positioning error results in zero nearby, but in some period of time the error is

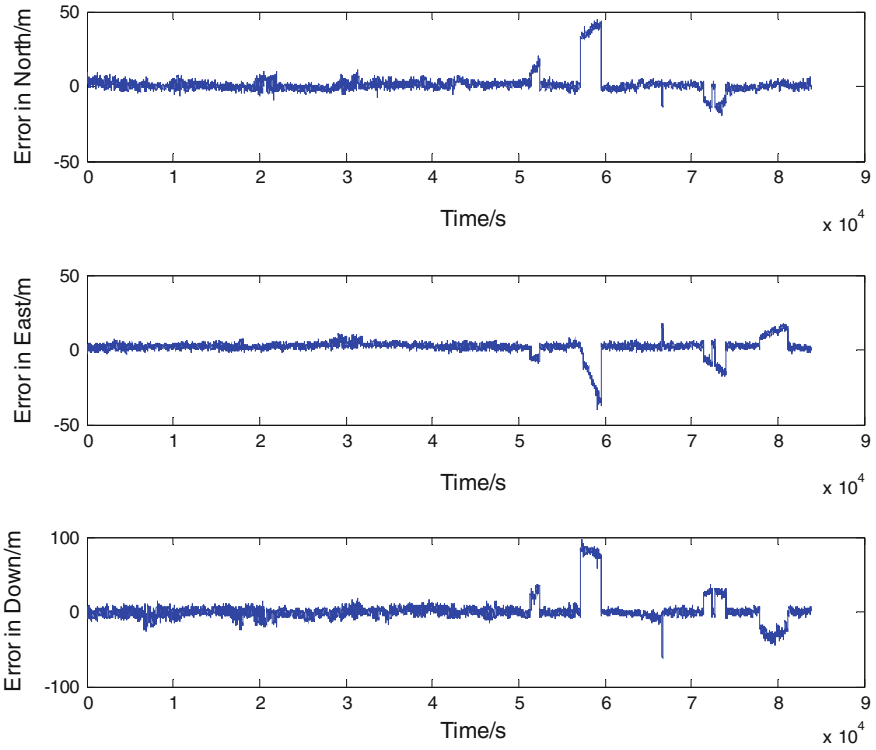


Fig. 1 Position error curve obtained by using standard SPP

unacceptable. After analysing the data, it is found that measurement of a low elevation satellite value is abnormal. The positioning results will meet the requirements after rejecting the satellite with fault.

The false alarm rate P_{FA} is set as 1×10^{-3}. After fault detection and exclusion with Baarda method and RBS method, the following results are obtained (Fig. 2).

According to the above results,it is shown that Baarda method can realize the fault exclusion and improve the accuracy of the SPP. However, it is obvious that the effect of the method is not satisfactory in some periods of time. Sometimes the measurements with faults are not correctly excluded. Therefore, the positioning error is still too large. Sometimes too many satellites are excluded that the number of remaining satellites is less than four. As a result, it's impossible to carry out single point positioning.

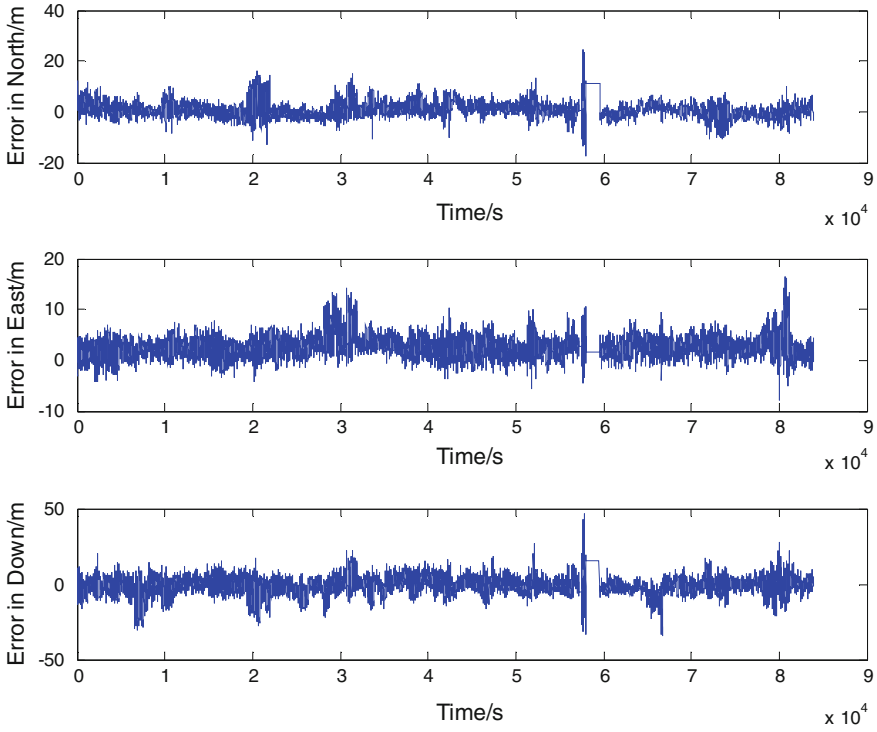


Fig. 2 Position error curve obtained by using Baarda method

With the same parameters, using the RBS method, the obtained results are shown as follows (Fig. 3):

By contrast it's obvious that the positioning accuracy was improved without large error. And all the time positioning result is able to be obtained with only the GPS satellites. So that it can be considered that positioning result is satisfactory (Table 2).

By comparing the positioning error in three directions, it can be concluded that FDE algorithm is able to improve the positioning precision significantly. As to the fault exclusion algorithm, the performance of RBS is better than Baarda in three directions.

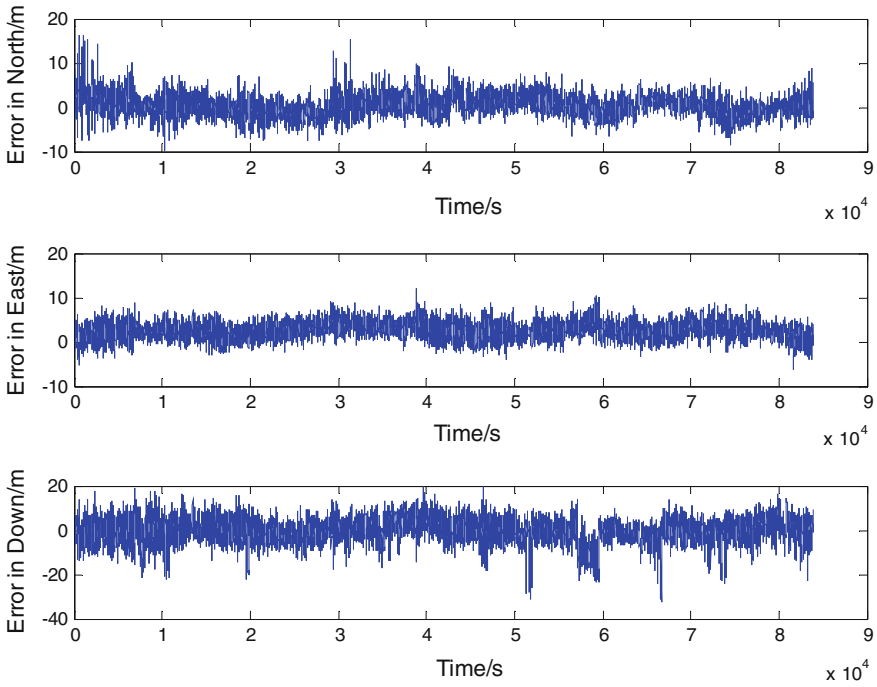


Fig. 3 Position error curve obtained by using the RBS method

Table 2 Positioning error comparison among three methods

Local direction		Mean error (m)	Standard deviation (m)	RMSE (m)
North	Origin	1.534	6.915	7.083
	Baarda	0.968	3.072	3.221
	RBS	0.356	1.982	2.014
Earth	Origin	1.890	5.189	5.523
	Baarda	2.544	1.859	3.151
	RBS	2.595	1.559	3.027
Down	Origin	1.349	16.100	16.156
	Baarda	−0.720	6.096	6.138
	RBS	−0.414	4.724	4.742

4 Conclusion

The RAIM algorithm is essential for detecting and excluding faults. Otherwise, the positioning precision will be unacceptable. Data snooping method proposed by Baarda is a classic method to exclude faults. However, healthy measurements are

possible to be excluded with Baarda method, which will decrease the positioning precision.

This paper proposes a residual-based separation method which is used to exclude faulty measurements. Because only one satellite is excluded at every round, the faulty exclusion frequency will be much smaller than Baarda method.

After comparing and analysing the positioning results, the conclusions are as follows: In case of few available satellites, the accuracy of SPP is greatly affected by the faults among measurements. And fault detection and exclusion algorithm can significantly improve the positioning accuracy. Comparing to the Baarda method proposed in this paper, the RBS method can not only avoid rejecting too many satellites, but also obviously improve the positioning accuracy.

References

1. Xu S, Zhang H, Yang Z, Wang Z (2015) Theory and application of GPS measure. Wuhan University Press, Wuhan
2. Wu J, An X, Zheng W (2015) Air-vehicle position and navigation technology. National Defense Industry Press, Beijing
3. Chen J (2001) Research of GPS integrity augmentation. Information Engineering University, Zhengzhou (in Chinese)
4. Joerger M, Chan F-C, Pervan B (2014) Solution separation versus residual-based RAIM. J Inst Navig U S A
5. Chan F-C, Pervan B (2010) A practical approach to RAIM-based fault-tolerant position estimation In: Proceedings of 23rd international technical meeting of the satellite division of the institute of navigation, ION GNSS 2010, Portland, OR, Sept 2010, pp 3181–3190
6. Blanch J, Walter T, Enge P (2013) Result on the optimal detection statistic for integrity monitoring In: Proceedings of the 2013 international technical meeting of the institute of navigation, San Diego, California, Jan 2013, pp 263–273
7. Joerger M, Stevanovic S, Chan F-C, Langel S, Pervan B (2013) Integrity risk and continuity risk for fault detection and exclusion using solution separation ARAIM In: Proceedings of the 26rd international technical meeting of the satellite division of the institute of navigation, ION GNSS+2013, pp 2702–2722

BDS Code Bias and Its Effect on Wide Area Differential Service Performance

Sainan Yang, Junping Chen, Yize Zhang, Chengpan Tang, Yueling Cao, Qian Chen and Wei Chen

Abstract The integrated wide-area differential system of BDS broadcasts differential corrections to compensate broadcast ephemeris errors and enhance user positioning accuracy. It uses real-time pseudo-range observations from ground monitoring stations and calculates differential corrections, which are then broadcasted to users through the GEO satellites. Previous studies show that apparent code biases exist for BDS tracking stations, which in turn will affect the accuracy of differential corrections. This paper analyses the long-term pseudo-range biases based on the residuals in the Precise Orbit Determination (POD), the characteristics of the mean and standard deviation of the pseudo-range bias are studied. A code bias correction model is proposed and applied to BDS wide-area differential calculation. Performance of the code bias model is analysed in real-time pseudo-range based positioning. Results show that: UDRE and user positioning accuracy are improved, where the satellite UDRE is reduced from 0.43 to 0.35 m, and the user positioning accuracy is increased by 10.6% and 14.6% in horizontal and height directions, respectively.

Keywords BDS · Code bias · Differential correction · UDRE · Positioning

S. Yang · J. Chen (✉) · Y. Zhang · C. Tang · Y. Cao · Q. Chen
Shanghai Astronomical Observatory, Chinese Academy of Sciences,
Shanghai, People's Republic of China
e-mail: junping@shao.ac.cn

S. Yang · J. Chen · C. Tang · Q. Chen
University of Chinese Academy of Sciences,
Beijing, People's Republic of China

J. Chen · Y. Cao
Shanghai Key Laboratory of Space Navigation and Positioning Techniques,
Shanghai, People's Republic of China

Y. Zhang
Department of Surveying and Geo-Informatics, Tongji University,
Shanghai, People's Republic of China

W. Chen
State Grid Taixing County Electric Power Supply Company,
Taixing, People's Republic of China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_20

1 Introduction

The BeiDou regional satellite navigation system (BDS) provides legacy PNT service to open users and wide-area differential service to authorized users since October 25, 2012. The BDS satellite constellation is combined of 5 Geostationary Earth Orbit (GEO) satellites, 5 Inclined Geostationary Earth Orbit (IGSO) satellites, and 4 Medium Earth Orbit (MEO) satellites [1–3]. Researches have been carried out in precision orbit determination (POD), precision positioning and wide-area differential positioning [4–8]. Previous researches found the existence of code bias in BDS pseudo-range observations, which is of meter level and is elevation dependent with significant differences between two groups of satellites [9, 10].

As an augmentation service to the legacy navigation message parameters, BDS wide-area differential system calculates the differential corrections using real-time observations of ground monitoring stations. These differential corrections are then broadcasted to the user through the GEO satellites [7, 8]. Since pseudo-range observations are used in the wide-area differential system [8], the code bias will also influence the accuracy of wide-area differential corrections.

In this paper, we analyze the long-term pseudo-range biases based on the residuals in the Precise Orbit Determination (POD). The characteristics of mean and standard deviation of code bias are studied. We developed a code bias correction model. The model is then applied to the estimation of BDS wide-area differential corrections. Real-time pseudo-range observations are used to evaluate the performance of the correction model. Results show that the accuracy of the UDRE and user differential positioning are improved using the code bias correction model, where UDRE is decreased from 0.43 to 0.35 m, and user differential positioning accuracy is increased by 10.6% and 14.6% in the horizontal and height directions respectively.

2 Analysis of BDS Code Bias

Precise orbit determination of BDS makes use of observations of the monitoring stations in China, where the internal 3D precision of the orbit determination is better than 1 m and the radial accuracy of the orbit is better than 0.2 m [6]. The strategies used in BDS POD are shown in Table 1 [6]:

Precise satellite orbits and clocks are the main products in precise orbit determination. Additionally, we can retrieve pseudo-range residuals for each station/satellite pair. Wanninger L et al. [9] analyzed BeiDou code bias derived from BeiDou linear combination observations. In this paper, BeiDou code biases are obtained from POD process. As phase observation has much higher weights in POD, pseudo-range residuals in POD thus reflect the effects of multipath, receiver hardware delays, noise etc., which is used in the next session for code bias analysis and modeling (Table 1).

Table 1 The strategies used in BeiDou precise orbit determination

Items	Description
Arc length	3 days
Data sampling rate	60 s
Estimated solar parameters	Scalar factor and Y-bias
Empirical parameters	Only cosine empirical accelerations in the along-track and cross-track directions to avoid correlation
Troposphere delay factor	One parameter per 4 h
Clock	The master station clock fixed to solve the other clock offsets
Weight rate	Pseudo-range/Phase(1/10000)

2.1 Code Bias Time Series

Based on the pseudo-range residuals from POD in DOY 316–345 of 2016, we calculate the mean and standard deviation of code bias. Figure 1 shows the daily mean and standard deviation of the code bias for each satellite of the station CDJL. We see that the daily mean code bias of different satellites is relatively stable, and they differ among satellites. The code bias standard deviation is stable for most satellites, while the variation of MEO satellites (sat11, sat12, and sat14) is much more obvious with a 7-day period. The variation of daily standard deviation of

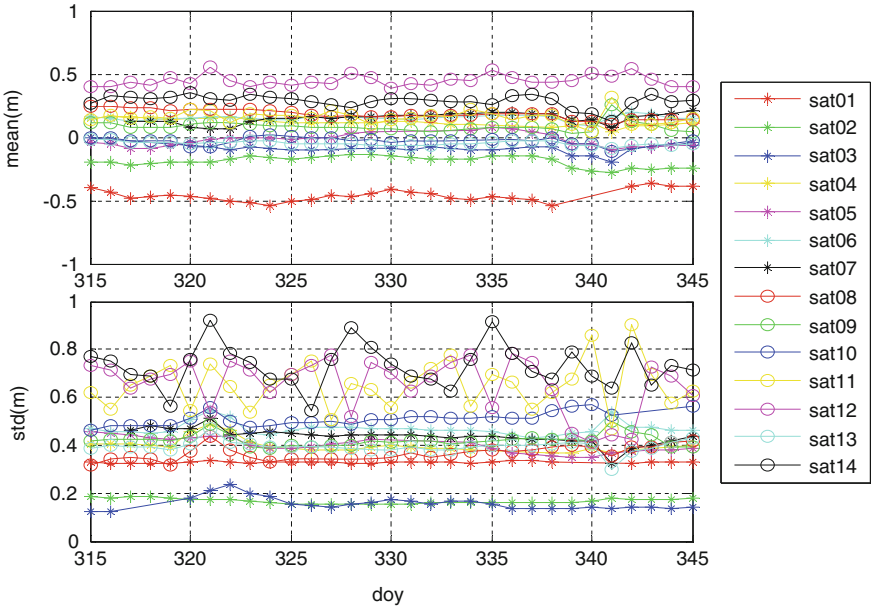


Fig. 1 The daily mean, standard deviation of the pseudo-range residuals

MEO satellites may be due to the different visible MEO arc, while the period of MEO satellite trajectory track relative to the ground station is 7 days.

2.2 Repeatability of Code Bias

Figure 2 shows the daily pseudo-range residual variations for the GEO-02, IGSO-06 and MEO-14 satellites at station BJDN over one-month. The top subplot in (a) shows the time-series of original pseudo-range residuals of GEO-02 for 1 month, where the missing 3 days data is due to satellite maneuvering, and the time-series of the GEO-02 is smoothed over one day and shown in the bottom subplot in (a). Plots (b) and (c) show the original and smoothed residuals of the IGSO/MEO satellites in top and bottom subplots respectively, which show their variation with the elevation angle. In plots (b) and (c), different color curves represent residuals in different days. It is found that the shape of pseudo-range residuals of each satellite is similar and very reproducible. In the range where the satellite elevation angle less than 30°, the satellite pseudo-range bias is relative large, and there is a certain range of tremble accompanied by a large noise. The code bias noise of IGSO/MEO satellites decrease with the elevation angle increase, but the pseudo-range bias still exists, and it is correlated with the satellite elevation angle.

2.3 Code Bias of Different Groups of Satellites

Further analysis of the pseudo-range bias was performed for different groups of satellites, namely the GEO, IGSO, and GEO group. Figure 3 shows pseudo-range residuals of all satellites of the station BJDN in DOY 345 of 2016. It is found that there are differences among different GEO satellites, where the code bias of each

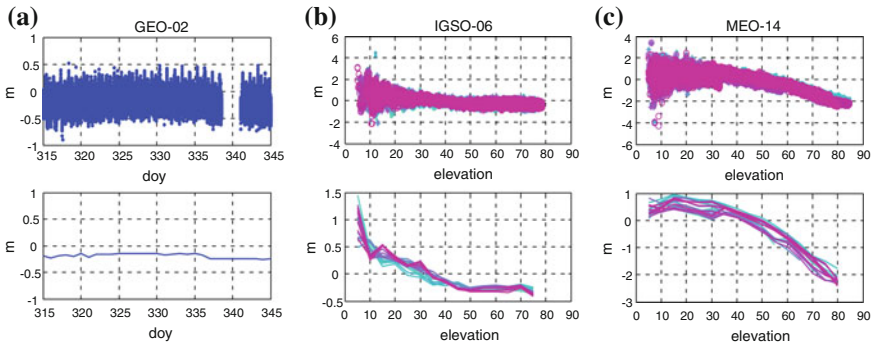


Fig. 2 Pseudo-range residual variations for the GEO-02, IGSO-06 and MEO-14 satellite over one month, where X-axis represents DOY in (**a**); X-axis represents elevation angle in (**b**) and (**c**); *top subplot* represents original pseudo-range residuals, *bottom subplot* represents smoothed pseudo-range residuals, and *different colors* represent different days

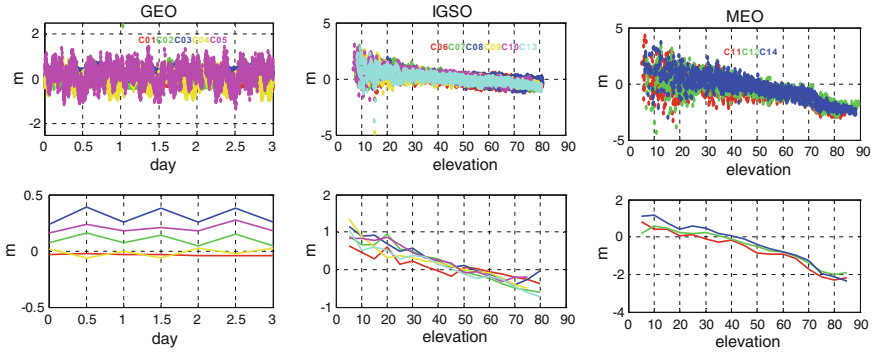


Fig. 3 Code bias of GEO/IGSO/MEO satellites, where *X-axis* represents time in 3 days for GEO; and *X-axis* represents elevation angle for IGSO/MEO; *different colors* represent different satellites

satellite differs in amplitude. The code biases of the same type of MEO or IGSO satellite are more consistent. Furthermore, the code bias of IGSO and MEO satellites differ in the magnitude of pseudo-range bias. The deviation of code bias can reach 1 m between low and high elevation angle. The code bias of IGSO satellite is between 0 and −0.5 m when satellite elevation reaches 90°, while it can reach the amplitude of −2 m for MEO satellite (Fig. 3).

3 BDS Code Bias Correction Model

BeiDou wide-area difference system utilizes the real-time pseudo-range observation of monitoring stations to calculate the wide-area differential correction for users. Based on the above analysis of the BeiDou code bias, we see that code biases are different for satellites at the same station which means that the code bias will not be assimilated in station clock; Also, the code biases are different for stations tracking the same satellite, which means that the code bias will also not be assimilated in satellite clock. Consequently, the code bias will affect the accuracy of wide-area differential corrections.

Wanninger L et al. analyzed BeiDou code bias derived from linear combination observations and set up correction models for BeiDou IGSO/MEO satellites. In this paper, we extend bias modeling for GEO satellites and set the elevation interval to 5° for the piecewise model of IGSO/MEO satellites.

According to the above discussion, the code bias of GEO satellites is long-term stable, constant code bias correction model could be established for GEO represented as Eq. (1):

$$CodeBias_{GEO} = C_i (i = 1, 2, 3, 4, 5) \tag{1}$$

where i denotes satellite PRN number, and C_j is constant term to be determined.

Table 2 Piecewise linear model of Beidou code bias (unit: m)

GEO		IGSO				MEO			
Sat No	Bias	Elevation	Bias	Elevation	Bias	Elevation	Bias	Elevation	Bias
1	−0.052	5	0.838	50	−0.206	5	0.753	50	−0.435
2	0.096	10	0.523	55	−0.216	10	0.488	55	−0.705
3	0.271	15	0.425	60	−0.282	15	0.539	60	−0.825
4	−0.008	20	0.41	65	−0.313	20	0.418	65	−1.01
5	0.176	25	0.355	70	−0.394	25	0.453	70	−1.507
		30	0.295	75	−0.413	30	0.357	75	−1.821
		35	0.162	80	−0.475	35	0.097	80	−2.121
		40	0.065	85	−0.665	40	0.049	85	−2.113
		45	−0.094			45	−0.107		

The code bias of IGSO/MEO satellites is correlated with satellite elevation angle, the correction model of IGSO/MEO satellite could be expressed as piecewise linear function model related to satellite elevation angle shown in Eq. (2):

$$CodeBias_{IGSO/MEO} = \begin{cases} C_j(j<5) \\ a_j + k_j \cdot (ele - ele_j)\ (j = 5, 10, \ldots, 85) \\ C_j(j > 85) \end{cases} \tag{2}$$

where *ele* is the elevation of the satellite, j represents the elevation angle index of the satellite; and C_j is constant term, a_j, k_j.are model parameters to be estimated.

The model can be provided to the users to interpolate the specific code bias with a given elevation angle.

The BDS code bias correction model is established by above data, and the constant bias value of each GEO satellite and model parameters of IGSO/MEO satellites are shown in Table 2. Figure 4 shows the code biases and their mean value for IGS/MEO satellites at each elevation node.

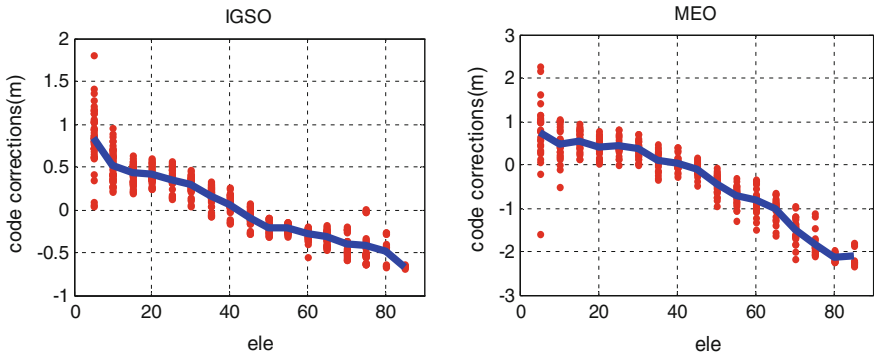


Fig. 4 The bias value of IGSO/MEO satellite at each elevation node, where *X-axis* represents elevation, the *blue* represents the mean bias value at elevation node, the *red* represents value of each satellite

4 Performance Evaluation of Code Bias Model

Applying the BeiDou code bias correction model in the BeiDou wide-area differential correction process, we evaluate the accuracy of resulting differential correction parameters and its performance in user differential positioning. In wide-area differential data processing, same strategies are implemented except the additional code bias model implement.

Experiment was carried out by using observation data in DOY 346 of 2016, two sets of differential corrections (namely "equivalent clock") are calculated with and without code bias corrections. The effect of code bias on differential correction is analyzed and further verified by the user's differential positioning.

4.1 Differences Between the Two Sets of Equivalent Clock

Figure 5 shows two sets of equivalent satellite clock results with and without correcting code bias. GEO-01, IGSO-07 and MEO-14 are analyzed. We found that GEO's equivalent satellite clock differs little with magnitude in 2 dm or less. The equivalent clock difference of IGSO/MEO satellites varies greatly, especially in the low elevation, where the difference may reach 1 m. Table 3 shows the RMS value of the difference for each satellite. It can be seen that the difference of the IGSO/MEO satellite is larger than that of the GEO satellite.

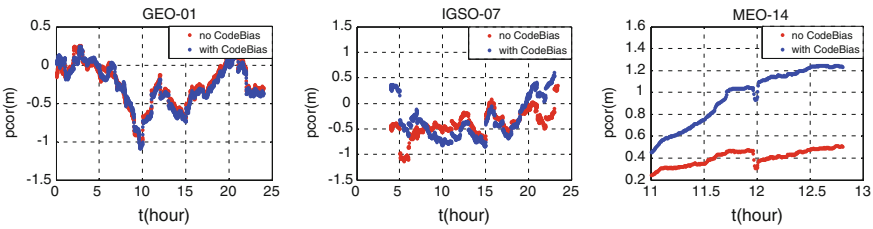


Fig. 5 Differences between the two sets of equivalent clock of C01/C07/C14

Table 3 RMS statistics of satellite equivalent clock differences before and after correction of code bias (unit: m)

satID	RMS	satID	RMS
1	0.066	8	0.336
2	0.044	9	0.337
3	0.029	10	0.34
4	0.022	11	0.332
5	0.135	12	0.327
6	0.35	13	0.34
7	0.356	14	0.59

4.2 UDRE Comparison

The differences between the two sets of equivalent clock may cause difference in the differential positioning accordingly. We calculate the satellite UDRE before and after the application of code bias correction model, and the results are shown in Fig. 6 and Table 4.

In the above figure, satellite 1–5 are GEO satellites, 6–10 and 13 are IGSO satellites, and 11–12 and 14 are MEO satellites. After correcting the code bias model, the GEO satellite UDRE decrease from 0.44 to 0.31 m, the IGSO satellite is from 0.44 to 0.39 m, the MEO satellite is from 0.42 to 0.32 m. The UDRE of all the satellites is improved.

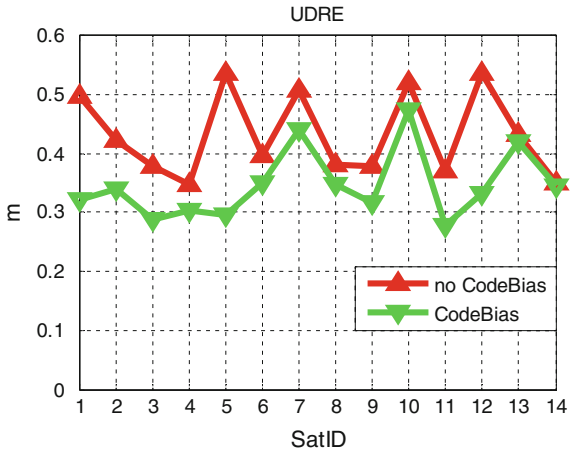


Fig. 6 Satellite UDRE comparison before and after the application of code bias model

Table 4 Satellite UDRE before and after the application of code bias model (unit: m)

satID	UDRE	UDRE (CodeBias)	satID	UDRE	UDRE (CodeBias)
1	0.50	0.32	10	0.52	0.48
2	0.42	0.34	11	0.37	0.28
3	0.38	0.29	12	0.54	0.33
4	0.35	0.30	13	0.43	0.42
5	0.54	0.30	14	0.35	0.35
6	0.40	0.35	GEO	0.44	0.31
7	0.51	0.44	IGSO	0.44	0.39
8	0.38	0.35	MEO	0.42	0.32
9	0.38	0.32	Mean	0.43	0.35

4.3 Comparison of User Differential Positioning Results

Based on the above analysis results, we use two sets of equivalent satellite clock parameters to perform kinematic positioning using dual-frequency pseudo-range data in DOY 346 of 2016. In positioning process, tropospheric delay, solid tide, relativity correction, satellite and receiver phase center offsets were corrected by model.

In the use of equivalent satellite clock parameters with code bias model applied, the same model was used in user station to correct pseudo-range data. Figure 7 shows the results of the two groups of user positioning, where the two sub-graphs show the statistical results of the positioning results of the different stations in the horizontal and height directions, respectively. Table 5 summarized the positioning results, where the user positioning accuracy in the horizontal improves from 1.04 m to 0.93 m, and in height component improves from 1.6 m to 1.36 m. The improvement is about 10.6 and 14.6% for horizontal and height component (Table 5; Fig. 7).

Table 5 The mean statistical results of dual-frequency pseudo-range kinematic differential positioning (unit: m)

	No CodeBias	CodeBias	Improvement percentage
Horizontal	1.036	0.926	10.63
Height	1.597	1.364	14.59

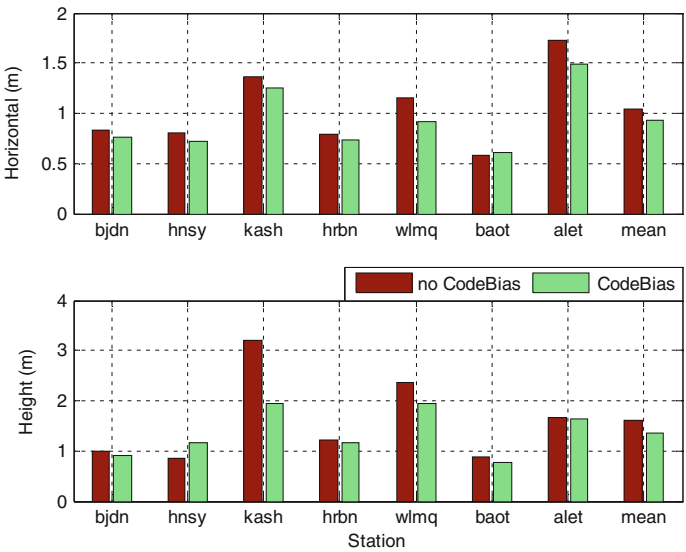


Fig. 7 Comparison of dual-frequency pseudo-range kinematic differential positioning for different stations

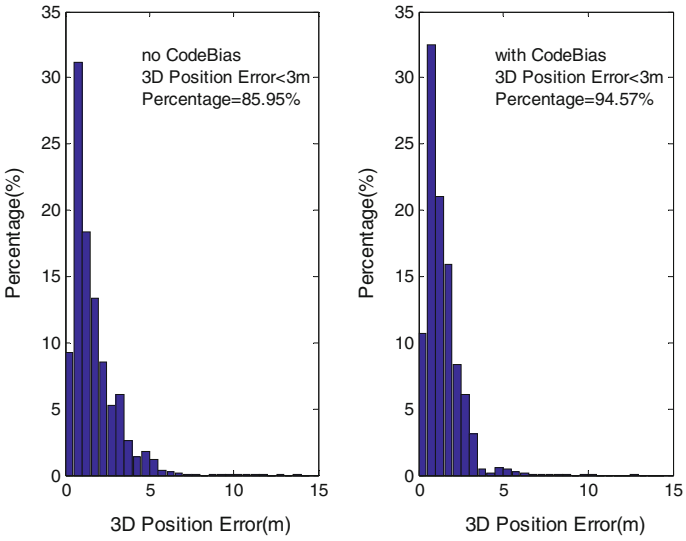


Fig. 8 The histogram of the statistics results of positioning error of all above stations

Figure 8 shows the histogram of the statistics results of positioning error of all above stations. It can be seen that the amount of 3D positioning error bigger than 5 m is reduced, and the percentage of positioning error less than 3 m is increased from 85.9% to 94.5%. This shows that the code bias correction model improves user positioning accuracy and the reliability simultaneously.

5 Conclusion

In this paper, we analyze the long-term pseudo-range residuals results of BeiDou POD. We find that the code bias of GEO satellite is relatively stable, and the code bias of IGSO/MEO satellite is the function of elevation angle. According to the analysis results, the BeiDou code bias correction model is established for different types of satellites, and the model is applied to BeiDou wide area differential processing. The performance of Beidou code bias model is evaluated by using the real-time pseudo-range observations. Results show that UDRE and user positioning accuracy are improved by using the Beidou code bias model, where the satellite UDRE decreased from 0.43 m to 0.35 m, and the user positioning accuracy is improved by 10.6% and 14.6% in horizontal and height directions, respectively.

Acknowledgements This paper is supported by the National 863 Program (Grant No. 2014AA123102), National Natural Science Foundation of China (NSFC) (Grant No. 11673050, 11273046, 41674041, 11403112), and the Opening Project of Shanghai Key Laboratory of Space Navigation and Positioning Techniques (ZZTX-201401).

References

1. Yang Y, Li J, Wang A et al (2014) Preliminary assessment of the navigation and positioning performance of BeiDou regional navigation satellite system. Sci China Earth Sci 57 (1):144–152
2. Zhigang HU (2013) BeiDou navigation satellite system performance assessment theory and experimental verification. Wuhan University, Wuhan
3. Beidou Satellite Navigation System [EB/OL], http://www.beidou.gov.cn/xtjs.html
4. Shi C, Zhao Q, Li M et al (2012) Precise orbit determination of beidou satellites with precise positioning. Sci China Earth Sci 6:854–861
5. Zhou SS, Hu XG, Wu B et al (2011) Orbit determination and time synchronization for a GEO/IGSO satellite navigation constellation with regional tracking network. Sci China Phys Mech Astron 54(6):1089–1097
6. Tang C, Hu X, Zhou S et al (2016) Improvement of orbit determination accuracy for Beidou navigation satellite system with two-way satellite time frequency transfer. Adv Space Res 58 (7):1390–1400
7. Cao YL, Hu XG, Wu B et al (2012) The wide-area difference system for the regional satellite navigation system of COMPASS. Sci China Phys Mech Astron 55(6):1307–1315
8. Cao Y, Hu X, Zhou J et al (2014) Kinematic wide area differential corrections for BeiDou regional system basing on two-way time synchronization. In: China satellite navigation conference (CSNC), pp 277–288
9. Wanninger L, Beer S (2015) BeiDou satellite-induced code pseudorange variations: diagnosis and therapy. GPS Solutions 19(4):639–648
10. Guo F, Li X, Liu W (2016) Mitigating BeiDou satellite-induced code bias: taking into account the Stochastic model of corrections. Sensors 16(6)

Ionospheric STEC and VTEC Constraints for Fast PPP

Yan Xiang, Yang Gao and Yihe Li

Abstract Significant initialization or convergence time of PPP is a limiting factor for many applications. Recent studies have indicated that external ionospheric information can be applied to provide a tight constraint to reduce the correlation between the position parameters and the ionosphere to shorten PPP long convergence time. However, receiver DCB is one challenge to separate from STEC. Receiver DCB is often neglected and taken for granted that it can be absorbed by the receiver clock parameter. In this paper, we critiqued the ionospheric STEC and VTEC constrained PPP models separating receiver DCB by taking advantage of the ionospheric information from a nearby station and GIM products. The results show that the improvements of positioning accuracy for ionospheric STEC and VTEC constrained PPP models are about 67 and 26% after a convergence time of 1 min. The convergence time is significantly reduced with the ionospheric constraints from a nearby station, while the performance of convergence time with GIM depends on the quality of GIM.

Keywords Uncombined precise point position (UPPP) · Ionospheric constraints · Slant total electron content (STEC) · Vertical total electron content (VTEC) · Receiver DCB · Convergence time

Y. Xiang (✉)
University of Calgary, 500 University Drive NW
Calgary, Calgary, AB T2N 1N4, Canada
e-mail: yan.xiang@ucalgary.ca

Y. Gao
University of Calgary, Calgary, Canada
e-mail: ygao@ucalgary.ca

Y. Li
Trimble Navigation Limited, Sunnyvale, USA
e-mail: yihli@ucalgary.ca

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_21

1 Introduction

Over the last two decades, precise point positioning (PPP) [10, 25] has become a valuable method in various kinds of applications, such as precise positioning, autonomous vehicle navigation, atmospheric sensing and seismology [1, 4, 6–8, 12, 15, 19, 23]. However, significant initialization or convergence time is required for PPP in order to reach the centimeter-level. It typically takes about half an hour or longer to reach the centimeter-level in three dimension (3D) [2, 8]. Because of the long convergence period, PPP is still mainly applied in places where network infrastructures are not available or long convergence time is not a major concern, such as offshore and glacier positioning. Reducing the convergence period is therefore essential in order to utilize the PPP technique to many other applications.

Recent studies have shown that the ionospheric information is helpful to shorten PPP long convergence time [12, 13, 18, 20, 24]. The conventional PPP model relies on the ionosphere-free combination observations [10]. This traditional PPP model amplifies the noises when dual frequency observations are combined to cancel the ionospheric effects. It is reported that eliminating the ionospheric delay by observation combination is equivalent to estimating the ionospheric effects as unknown parameters in the uncombined PPP (UPPP) model without constraint [3, 14, 16]. In contrast, the UPPP model based on uncombined measurements has been developed in which the ionospheric delays are to be estimated. The advantage of this model is that external accurate ionospheric information can be applied to provide a tight constraint to reduce the correlation between the position parameters and the ionosphere, and as a result, it can accelerate the convergence time.

Ionosphere constrained PPP models have drawn the attention from researchers in precise positioning community in recent years. Li et al. [12] attempted to fix the ambiguities instantaneously with a dense network. The authors also held that receiver DCBs from reference stations could be absorbed by receiver clock in the user ends. Shi et al. [19] modeled the ionosphere deterministically and stochastically with a polynomial function and a random part to improve PPP performance for single-frequency users. Tu et al. [20] implemented the combination of GPS and GLONASS using UPPP by modeling the ionosphere with a polynomial function in a similar manner. They found that the convergence time to sub-decimeter could be reduced to 10 min. Zhang et al. [24] discussed the convergence performance of ionosphere constrained PPP with receiver DCB considered, but only on code measurements. Rovira-Garcia et al. [18] applied the slant ionospheric corrections from reference stations, and achieved faster convergence. Lou et al. [15] demonstrated the single- and dual- frequency PPP using multi-system, and an improvement of 60% in the convergence time was found.

Although significant convergence time has been improved by adding ionospheric constraints, the receiver's bias effects on both code and phase measurements have been little discussed. In this paper, we display an ionosphere constrained UPPP model to accelerate the positioning convergence. In the proposed UPPP model, the ionospheric constraints are in the form of either slant total electron

content (STEC) or vertical total electron content (VTEC). Different from a basic UPPP model, the receiver DCB is also estimated separately from the ionospheric parameters in the constrained UPPP model. The paper is organized as follows. Section 2 provides a detailed description of the methodology developed for the ionosphere constrained UPPP models. Experiments and results are then given in Sect. 3 to assess the performance improvement of the ionosphere constrained UPPP models. A comparison to the basic UPPP model is also conducted. Section 4 summarizes the research work including conclusions and future work.

2 Methodology

2.1 Basic UPPP Model

We start with the UPPP model since it is the basic model for applying the ionospheric constraints. To clarify the receiver and satellite clocks that contain the bias terms, the traditional model is given first. The traditional PPP [10] is usually based on ionosphere-free (IF) L1/L2 combined observation equations as shown in Eq. (1), where the first-order ionosphere delays are removed. As can be seen from this equation, the receiver and satellite clocks are consistent with the ionosphere-free combination observations.

$$\begin{cases} P_{IF} = \rho + \underbrace{(cdt_r + b^r_{P_{IF}})}_{d\tilde{t}} - (cdt^s + b^s_{P_{IF}}) + T + \varepsilon_P \\ \Phi_{IF} = \rho + \underbrace{(cdt_r + b^r_{P_{IF}})}_{d\tilde{t}} - (cdt^s + b^s_{P_{IF}}) + d_{orb} \\ \qquad + T + \underbrace{N_{IF} + [(b^r_{\Phi_{IF}} - b^r_{P_{IF}}) - (b^s_{\Phi_{IF}} - b^s_{P_{IF}})]}_{\tilde{N}_{IF}} + \varepsilon_\Phi \end{cases} \tag{1}$$

where

P_{IF}	denotes the ionosphere-free code combination;
Φ_{IF}	is the ionosphere-free carrier phase measurements (m);
ρ	is the distance from receiver to satellite (m);
c	is the light speed;
dt_r	is the receiver clock (s);
dt^s	is the satellite clock (s);
T	is the troposphere delay (m);
N_{IF}	is the ionosphere-free ambiguity (m);
ε	is the multipath and measurements noise (m);
$b^r_{P_{IF}}, b^s_{P_{IF}}$	are the biases of IF combination at satellite and receiver (m);
$b^r_{\Phi_{IF}}, b^s_{\Phi_{IF}}$	are the phase bias of IF combination at satellite and receiver (m).

As the IGS satellite clocks are applied in the later UPPP model, the relationship between the biases at one frequency and IF biases needs to be established and is developed as below.

Since $DCB_{P1/P2} = b_{P1} - b_{P2}$, and $\begin{cases} b_{P1} = b_{P_{IF}} + 1/(1-\gamma_2) \cdot DCB_{P1/P2} \\ b_{P2} = b_{P_{IF}} + \gamma_2/(1-\gamma_2) \cdot DCB_{P1/P2} \end{cases}$, we have

$$\begin{cases} b^r_{P_1} - b^s_{P_1} = b^r_{P_{IF}} - b^s_{P_{IF}} + 1/(1-\gamma_2) \cdot (DCB^r_{P1/P2} - DCB^s_{P1/P2}) \\ b^r_{P_2} - b^s_{P_2} = b^r_{P_{IF}} - b^s_{P_{IF}} + \gamma_2/(1-\gamma_2) \cdot (DCB^r_{P1/P2} - DCB^s_{P1/P2}) \end{cases} \tag{2}$$

where $b^r_{P_j}, b^s_{P_j}$ are the code bias at frequency j at receivers and satellites (m); $\gamma_2 = f_1^2/f_2^2$.

After correcting the satellite clock errors using IGS precise clock products that are generated based on ionosphere-free observation equations, the basic UPPP model is expressed as the following observation equations.

$$\begin{cases} P_1 = \rho + (cdt_r + b^r_{P_{IF}}) - (cdt^s + b^s_{P_{IF}}) + T \\ \quad + (I_1 + \dfrac{1}{\gamma_2 - 1} DCB^s_{P1/P2} - \dfrac{1}{\gamma_2 - 1} DCB^r_{P1/P2}) + \varepsilon_P \\ P_2 = \rho + (cdt_r + b^r_{P_{IF}}) - (cdt^s + b^s_{P_{IF}}) + T \\ \quad + (\gamma_2 I_1 + \dfrac{\gamma_2}{\gamma_2 - 1} DCB^s_{P1/P2} - \dfrac{\gamma_2}{\gamma_2 - 1} DCB^r_{P1/P2}) + \varepsilon_P \\ \lambda_1 \Phi_1 = \rho + (cdt_r + b^r_{P_{IF}}) - (cdt^s + b^s_{P_{IF}}) + T - (I_1 + \dfrac{1}{\gamma_2 - 1} DCB^s_{P1/P2} \\ \quad - \dfrac{1}{\gamma_2 - 1} DCB^r_{P1/P2}) + (\lambda_1 N_1 + b^r_{\Phi_1} - b^s_{\Phi_1} - b^r_{P_{IF}} \\ \quad + b^s_{P_{IF}} + \dfrac{1}{\gamma_2 - 1} (DCB^s_{P1/P2} - DCB^r_{P1/P2})) + \varepsilon_\Phi \\ \lambda_2 \Phi_2 = \rho + (cdt_r + b^r_{P_{IF}}) - (cdt^s + b^s_{P_{IF}}) + T - (\gamma_2 I_1 + \dfrac{\gamma_2}{\gamma_2 - 1} DCB^s_{P1/P2} \\ \quad - \dfrac{\gamma_2}{\gamma_2 - 1} DCB^r_{P1/P2}) + (\lambda_2 N_2 + b^r_{\Phi_2} - b^s_{\Phi_2} - b^r_{P_{IF}} + b^s_{P_{IF}} \\ \quad + \dfrac{\lambda_2}{\gamma_2 - 1} (DCB^s_{P1/P2} - DCB^r_{P1/P2})) + \varepsilon_\Phi \end{cases} \tag{3}$$

where $I_1 = 40.3 \frac{STEc}{f_1^2}$ is the ionosphere delay along the line of sight between receiver and satellite at L1 frequency (m). N_j is the ambiguity at frequency j (cycle); $b^r_{\Phi_j}, b^s_{\Phi_j}$ are the phase bias at frequency j at receivers and satellites (m).

Based on Eq. (3), the following ionospheric terms I_1 grouped with DCB parameters are estimated as unknowns for each satellite as Eq. (4).

$$\tilde{I}_1 = I_1 + \frac{1}{1-\gamma_2} DCB^r_{P1/P2} - \frac{1}{1-\gamma_2} DCB^s_{P1/P2} \tag{4}$$

Generally speaking, the divergence between the UPPP model in Eq. (4) and the traditional PPP model in Eq. (1) lies in the way to treat ionosphere delays. The UPPP model estimates the ionosphere delay as unknown parameters, while the IF model eliminates the ionospheric delays by combination of dual-frequency observations. The UPPP model can also decrease the noise level with better internal consistency (smaller residuals). Regarding ionospheric models and constraints, they has been widely studied in long-baseline relative positioning [16], known as Stochastic Ionosphere Parameter (SIP) [3]. Recent studies have shown that the SIP model is also feasible and effective in PPP [11, 19–21, 23, 24].

If a receiver observes *n* satellites, the number of observations in the UPPP model is 4n, with 3n + 5 unknown parameters. Therefore, the redundancy in the UPPP model is n − 5, which is the same as the IF PPP model. Dach et al. [3] also reported that UPPP results with no ionospheric information constraints achieve the same effect like the IF combination PPP.

2.2 Ionosphere Constrained UPPP Model

When external ionospheric information is available, the UPPP model in Eq. (3) can be further extended. The external ionospheric information can be taken as a priori information or pseudo measurements. The estimation with external ionospheric information can be augmented as follows [17].

$$z^+ = \begin{bmatrix} z \\ x_0 \end{bmatrix} = \begin{bmatrix} H \\ I \end{bmatrix} x + \begin{bmatrix} v \\ v_x \end{bmatrix} \tag{5}$$

$$R^+ = \begin{bmatrix} R & 0 \\ 0 & P_0 \end{bmatrix} \tag{6}$$

where z^+ is the augmented measurement vector; R^+ is the corresponding augmented variance; v is the measurement errors; z is measurements; H is the design matrix and x is the state vector. x_0 and P_0 refer to the external ionospheric information (the ionospheric delays and their variance-covariance matrix) that can be obtained from the augmentation of either VTEC or STEC. The external information is the key to add constraints. We will explain the ways to obtain the priori information as STEC and VTEC in detail in this subsection.

The least square solution of the linear system of Eqs. (5) and (6) can be derived as Eq. (7).

$$\begin{cases} \mathbf{x} = (H^T R^{-1} H + P_0^{-1})^{-1} (H^T R^{-1} \mathbf{z} + P_0^{-1} \mathbf{x}_0) \\ P = (H^T R^{-1} H + P_0^{-1})^{-1} \end{cases} \tag{7}$$

The constrained solutions provide more precise solutions than that with no external ionospheric information and also accelerate the PPP convergence.

Please note that, as Eq. (3) shows, the $\widetilde{I}_1$ estimate contains not only the ionosphere but also the DCBs from satellites and receivers. Since the satellite DCBs are very stable, we can correct them in advance. Satellite DCB have been provided on a daily basis by International GNSS service (IGS) as a by-product of GIMs based on smoothed code. We also estimated the satellite DCBs by reducing the leveling errors using UPPP model [21]. In addition, if the P1 observations are available, we choose the P1 measurements. Otherwise, the C1 code measurements are used. Once the C1 type is used, the P1C1 biases are required to align to P1, and monthly basis products are available from CODE (ftp.unibe.ch. aiub/CODE/YYYY/P1C1YYMM.DCB). Therefore, the model can be further simplified as below by correcting the satellite DCBs before adding priori ionospheric constraints.

$$\begin{cases} P_1 = \rho + \left(cdt_r + b_{IF}^r\right) + T + (I_1 - \frac{1}{\gamma_2 - 1} DCB_{P1/P2}^r) + \varepsilon_P \\ P_2 = \rho + \left(cdt_r + b_{IF}^r\right) + T + \gamma_2 (I_1 - \frac{1}{\gamma_2 - 1} DCB_{P1/P2}^r) + \varepsilon_P \\ \lambda_1 \Phi_1 = \rho + \left(cdt_r + b_{IF}^r\right) + T - (I_1 - \frac{1}{\gamma_2 - 1} DCB_{P1/P2}^r) + \lambda_1 \widetilde{N}_1 + \varepsilon_\Phi \\ \lambda_2 \Phi_2 = \rho + (cdt_r + b_{IF}^r) + T - \gamma_2 (I_1 - \frac{1}{\gamma_2 - 1} DCB_{P1/P2}^r) + \lambda_2 \widetilde{N}_2 + \varepsilon_\Phi \end{cases} \tag{8}$$

where $\widetilde{N}_1$ and $\widetilde{N}_2$ are estimated ambiguities at L1 and L2 frequency.

However, the receiver bias is still an issue when adding the ionospheric constraints. After correcting the satellite DCBs, the ionosphere and receiver DCB remain in the UPPP model. As the receiver DCB is different for code and phase in sign, it is impossible to have the receiver DCB absorbed in the receiver clock parameter. Simply ignoring the receiver DCB will result in less precise solutions.

If the vertical ionosphere is available, like GIM, an additional parameter of receiver DCB must be considered. Li et al. [12] held that receiver DCBs from the reference stations could be absorbed by receiver clock parameter in the user ends. This is very important because the bias affects how the ionosphere errors being corrected.

2.2.1 Ionospheric STEC Constrained UPPP

If the external ionospheric information is obtained before mapping, we called it STEC. The ionospheric STEC constraints to user ends are highly correlated with that from the nearby stations. The closer the stations, the variation STEC is more similar. The accuracy level of the estimated ionospheric STEC parameter is

generally at a centimeter level, sometimes to a decimeter level. Therefore, applying the ionospheric STEC constraints from the nearby stations can compensate the ionospheric variations in the user end.

However, besides the ionosphere, the estimated ionospheric parameters also contain receiver DCBs from nearby stations. If we apply these constraints as pseudo measurements to the user end, the DCBs from both nearby stations and the user end need to be addressed. The deterministic and stochastic parts can be written as Eq. (9).

$$I_{1,0} = \widetilde{I}_{ex} = I - \frac{1}{(\gamma_2 - 1)} DCB_{ex}, \quad and \quad P_0 = \delta^2_{ex} \tag{9}$$

where $\widetilde{I}_{ex}$ is the estimated ionospheric parameters from nearby stations; DCB_{ex} is the receiver DCB from nearby stations; δ^2_{ex} is the variance of the estimated ionospheric parameters from the basic UPPP model.

2.2.2 Ionospheric VTEC Constrained UPPP

The ionosphere and the biases are highly correlated since they are both frequency-dependent. The conventional way to separate the ionosphere from biases is via ionospheric modeling. One typical example of VTEC products is global ionospheric TEC maps (GIM). GIM has been supported by IGS working group on ionosphere and available since June 1998. The primary objective of GIMs is to monitor the variation of ionosphere continuously and apply this information to single-frequency receivers who demand up-to-date ionosphere correction or scientist who are interested in highly accurate ionosphere models [5].

Fortunately, GIM is found helpful for fast PPP by being offered as a priori information. It is a pity that the accuracy of GIM is limited due to the number of stations used for modeling and the modeling errors [22]. Considering the errors from mapping function, the variance is doubled when the elevation is smaller than 20°. The deterministic and stochastic parts can be written as Eq. (10).

$$I_0 = I_{GIM}, \quad and \quad P_0 = \delta^2_{GIM} = \begin{cases} 4\delta^2_{GIM}, & ele < 20^\circ \\ \delta^2_{GIM}, & ele \geq 20^\circ \end{cases} \tag{10}$$

where I_{GIM} is interpolated from GIM model by mapping to the corresponding elevation; δ^2_{GIM} is the variance interpolated from the GIM.

3 A Case Study on PPP Performance

3.1 Data Processing

Two nearby stations SUTH and SUTM from the IGS network on May 11, 2014 are selected to validate the methods in Sect. 2. The two receivers are equipped with Ashtech UZ12 and Javad RINAGANT_G3T respectively. The receivers are with a distance of about 140 m, which is so close that the ionosphere information is assumed to be the same. We chose a short baseline instead of a network. In practice, the corrections are usually generated from a network at a relative long distance. However, in our experiment, we want to address the receiver DCB and study how fast it can be if precise ionosphere information is accessible. Therefore, we implemented the basic UPPP model, ionospheric STEC, and VTEC constrained UPPP models to analyze the performance of ionospheric constrained PPP.

Table 1 displays the parameter estimation strategies in data processing. Please note that the geometry-free (GF) threshold to detect cycle slips and the variance for receiver clock. A small GF threshold causes re-initialization of ambiguities due to the ionosphere disturbance, while a large threshold is risky miss cycle slips. Regarding variance of receiver clock, way too large variance than other variances is not efficient for numerical computation. Furthermore, an outlier detection is adopted to reject the satellite with any of the measurement residual larger than four times root square of its corresponding variance. In addition, the orbit, clock, and P1C1 products from Orbit Determination in Europe (CODE) are used for consistency.

Table 1 Data processing strategies for UPPP

Parameters	UPPP models
Observations	Raw observations
Sampling rate	30 s
Cut-off angle	10°
GF threshold for cycle slip	0.1 m
PCO/PCV	Corrected with the igs08_1886.atx
Phase windup	Yaw-attitude model from Kouba [9]
Solid earth/tides	IERS 2010
Receiver clock	Model as random walk, initialized from SPP[a] with variance of 60^2 (m^2)
Troposphere	Saastamoinen model for zenith hydrostatic delays and estimate the wet delays
Ionosphere	Estimate the slant delay for each satellite each epoch with constraints
Ambiguity	Model as random constant with initial variance of 30^2 (m^2)
Receiver DCB	Model as random constant with initial variance of 30^2 (m^2)

[a]SPP is stand point position with solutions estimated using pseudorange at L1 frequency

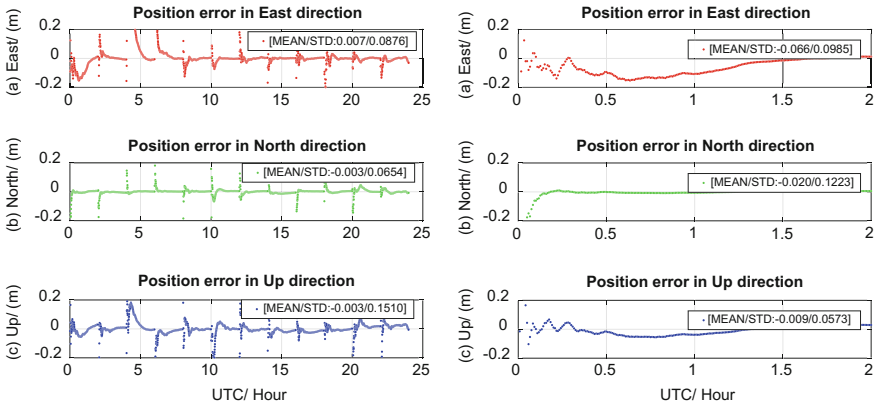


Fig. 1 Time series of position errors in east, north, and up direction using the basic UPPP model every two-hour interruption (*left*), and the first two-hour session (*right*)

3.2 *Basic UPPP Model*

Figure 1 shows the time series of position errors at IGS station SUTH in east, north, and up directions using the basic UPPP model, where the ionosphere is float weighted with dynamic variation of 5 mm/s. The errors are compared with week solutions in IGS08 reference frame. The left subplot is the time series of convergence process during the day by simulating interruption every two hours. The right figure shows the first two-hour session. It can be seen that the errors converge from a large error then smoothly converging to millimeter. Regarding the errors in east direction, they are slightly away from zero after the fluctuation at the beginning. It takes 62.5 min for the east direction converge to 0.1 m, even longer than the vertical direction in this case. We also checked the residuals, the standard deviation (std) for pseudorange is around at 0.4 m, and the std for phase is at about 3.0 mm.

3.3 *Ionosphere Constrained UPPP Model*

3.3.1 Ionospheric STEC Constrained UPPP

Similarly, position errors of SUTH in east, north, and up directions using the ionospheric STEC constrained UPPP model from the nearby station SUTM is shown in Fig. 2. The ionospheric STEC from SUTM is also from basic UPPP model including receiver bias. The left subplot is the convergence process during one day by simulating interruption every two hours. The right figure shows the second two-hour session. We chose the second one is because the first two-hour session is meaningless due to converging process that results in inaccurate

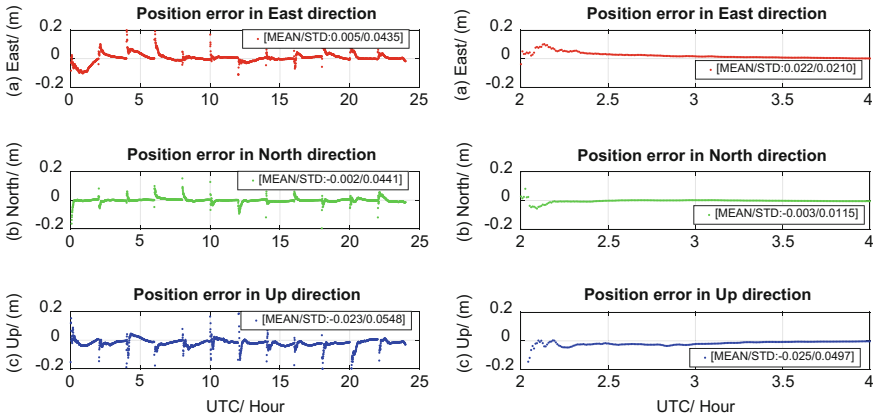


Fig. 2 Time series of position error in east, north, and up direction using the ionospheric STEC constrained UPPP model every two-hour interruption (*left*) and a two-hour session (*right*)

ionosphere corrections. Compared to the lower Fig. 1, lower Fig. 2 has a fast convergence process due to tight constraints on the ionosphere from nearby station SUTM. Compared with the basic UPPP model, the convergence process has smaller fluctuation, and the convergence time to 0.1 m for east direction is highly reduced to 8 min in this case.

3.3.2 Ionospheric VTEC Constrained UPPP

We also present the position errors of IGS station SUTH in east, north, and up directions using the ionospheric GIM constrained UPPP model in Fig. 3. Due to the accuracy of the ionosphere products, it can be seen that the first two-hour session has a little improvement in convergence time in right Fig. 3. If you observe the left figure of all the 12 sessions, it can be seen the convergence of other sessions sometimes are even worse than the basic UPPP model. The main reason is the poor accuracy of the ionosphere in GIM.

3.4 Statistics on Convergence Time for the Three Methods

To compare the difference among the basic, ionospheric STEC and VTEC constrained UPPP models, we summarized the statistic convergence time for 12 sessions. Table 2 shows the convergence time to 1 dm in three dimensions after a convergence of 1 and 5 min with 68 and 95%. From the first two rows, it can tell that the improvements of accuracy for ionospheric STEC and VTEC constrained UPPP models are about 67 and 26% after 1 min. When it comes to 5 min, there is no advantages for GIM constrained model due to the quality of GIM.

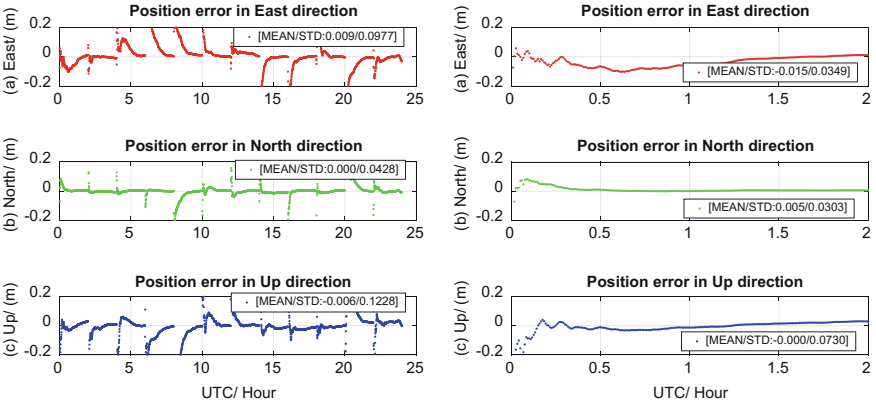


Fig. 3 Time series of position errors in east, north, and up direction using the ionospheric GIM constrained UPPP model with two-hour interruption (*left*), and a two-hour session (*right*)

Table 2 Statistics on the convergence time to reach 0.1 m in three dimensions

Error after	Basic UPPP (m)	GIM constrained UPPP	STEC constrained UPPP
1 min (68%)	0.73	0.54 (26.0%)	0.24 (67.1%)
1 min (95%)	0.97	0.71 (26.8%)	0.32 (67.0%)
5 min (68%)	0.13	0.28	0.10
5 min (95%)	0.21	0.35	0.15

4 Conclusions

In this paper, we enhanced the ionospheric STEC and VTEC constrained UPPP models by considering the receiver DCB on both code and phase measurements. The ionosphere constrained UPPP methods aim to emphasize the significance of receiver DCB. We also evaluated the performance of ionosphere constrained UPPP models by comparing with the basic UPPP model.

We first explained the estimated ionospheric parameters in the UPPP model to address the estimated ionospheric parameters contain DCBs as well. Secondly, we illustrated the ionospheric STEC constrained UPPP model by applying the slant ionospheric corrections from nearby stations. In this step, we showed that the receiver DCB from reference station must be considered in the constrained model. Thirdly, the ionospheric VTEC constrained UPPP model is implemented using GIM.

Based on the results, we can draw the critical conclusions as below.

- The estimated ionospheric parameters from UPPP contain DCB from satellites and receivers.

- The receiver DCB cannot be fully absorbed in the receiver clock. An extra receiver DCB parameter is required.
- The accuracy improvements after a convergence of 1 min for ionospheric STEC and VTEC constrained UPPP models are about 67 and 26%.
- There are no advantages of accuracy improvements at 5 min for ionospheric STEC and VTEC constrained UPPP models.

Research for this study was conducted using a single station under quiet ionospheric conditions. However, exposure to long-term data and a disturbed ionospheric environment would mean the statistics is not always the same. Our next step is working on analyzing globally distributed stations for a long-term data.

Acknowledgements Data from IGS and DCB products downloaded from CDDIS are acknowledged. Yan Xiang wishes to thank Xingyu Chen for valuable discussions and the funding from the Chinese Scholarship Council (CSC) and the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors thank the reviewers for their attention and valuable contributions during the review of this paper.

References

1. Bar-Sever YE, Kroger PM, Borjesson JA (1998) Estimating horizontal gradients of tropospheric path delay with a single GPS receiver. J Geophys Res Solid Earth 103:5019–5035
2. Bisnath S, Gao Y (2008) Current state of precise point positioning and future prospects and limitations. In: Observing our changing earth. Springer, Berlin, pp 615–623
3. Dach R, Hugentobler U, Fridez P, Meindl M (2015) Bernese GPS software version 5.2, vol 640. Astronomical Institute, University of Bern
4. Dixon K (2006) StarFire: a global SBAS for sub-decimeter precise point positioning. In: Proceedings of ION GNSS, pp 26–29
5. Feltens J, Schaer S (1998) IGS products for the ionosphere. In: Proceedings of the 1998 IGS Analysis Center Workshop Darmstadt, Germany
6. Gao Y, Chen K (2004) Performance analysis of precise point positioning using real-time orbit and clock products. Positioning 1
7. Gao Y, Shen X (2001) Improving ambiguity convergence in carrier phase-based precise point positioning. In: Proceedings of the 14th international technical meeting of the Satellite Division of the Institute of Navigation (ION GPS 2001), pp 1532–1539
8. Geng J (2010) Rapid integer ambiguity resolution in GPS precise point positioning. University of Nottingham
9. Kouba J (2009) A simplified yaw-attitude model for eclipsing GPS satellites. GPS Solut 13:1–12. doi:10.1007/s10291-008-0092-1
10. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solut 5:12–28. doi:10.1007/pl00012883
11. Li X (2012) Improving real-time PPP ambiguity resolution with ionospheric characteristic consideration. In: Proceedings of ION GNSS-12, Institute of Navigation, Nashville, Tennessee, September, 17–21
12. Li X, Zhang X, Ge M (2011) Regional reference network augmented precise point positioning for instantaneous ambiguity resolution. J Geod 85:151–158
13. Li Y, Li B, Gao Y (2015) Improved PPP ambiguity resolution considering the stochastic characteristics of atmospheric corrections from regional networks. Sensors 15:29772

14. Liu GCH (2001) Ionosphere weighted global positioning system carrier phase ambiguity resolution. University of Calgary, Department of Geomatics Engineering
15. Lou Y, Zheng F, Gu S, Wang C, Guo H, Feng Y (2016) Multi-GNSS precise point positioning with raw single-frequency and dual-frequency measurement models. GPS Solut 20:849–862. doi:10.1007/s10291-015-0495-8
16. Odijk D (2002) Fast precise GPS positioning in the presence of ionospheric delays
17. Petovello MG (2014) Estimation for navigation
18. Rovira-Garcia A, Juan JM, Sanz J, Gonzalez-Casado G (2015) A worldwide ionospheric model for fast precise point positioning. IEEE Trans Geosci Remote Sens 53:4596–4604
19. Shi C, Gu S, Lou Y, Ge M (2012) An improved approach to model ionospheric delays for single-frequency precise point positioning. Adv Space Res 49:1698–1708. doi:10.1016/j.asr.2012.03.016
20. Tu R, Ge M, Zhang H, Huang G (2013) The realization and convergence analysis of combined PPP based on raw observation. Adv Space Res 52:211–221
21. Xiang Y (2016) DCB estimation based on uncombined PPP. In: ION GNSS+, Portland
22. Xiang Y, Yuan Y, Li Z, Wang N (2014) Analysis and validation of different Global Ionospheric Maps (GIMs) over China. Adv Space Res
23. Zhang B, Ou J, Yuan Y, Li Z (2012) Extraction of line-of-sight ionospheric observables from GPS data using precise point positioning. Sci China Earth Sci 55:1919–1928. doi:10.1007/s11430-012-4454-8
24. Zhang H, Gao Z, Ge M, Niu X, Huang L, Tu R, Li X (2013) On the convergence of ionospheric constrained precise point positioning (IC-PPP) based on undifferential uncombined raw GNSS observations. Sensors 13:15708–15725
25. Zumberge J, Heflin M, Jefferson D, Watkins M, Webb F (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res Solid Earth (1978–2012), 102:5005–5017

Initial Assessment of BDS Zone Correction

Yize Zhang, Junping Chen, Sainan Yang and Qian Chen

Abstract Zone correction is a new type of differential corrections for BeiDou wide area augmentation system. As broadcasted together with the equivalent satellite clock and orbit corrections by BDS satellites, they enable user decimeter-level real-time positioning capability using the carrier-phase observations. In this paper, we give a brief introduction of zone corrections, and the function model of precise point positioning (PPP) for dual- and single-frequency users using the zone corrections. Tracking data of 30 stations in mainland China are used to evaluate the zone-divided PPP performance, and the handling of troposphere delay and ionosphere delay are discussed. Results show that the zone-divided PPP performance improves when fixing the troposphere delay. Model of UofC is much suitable for single frequency user. The dual-frequency PPP can convergences to 0.5 m in 25 min and the positioning accuracy are 0.15 m in horizontal and 0.2 m in vertical, respectively. As for single frequency PPP, the positioning accuracy convergences to 0.8 m in 20 min, while the positioning accuracy is 0.3 m in horizontal and 0.5 m in vertical.

Keywords BDS · Zone correction · Equivalent satellite clock · Convergence time · Precise point positioning

Y. Zhang · J. Chen (✉) · S. Yang · Q. Chen
Shanghai Astronomical Observatory, No. 80, Nandan Rd.,
Shanghai 200030, China
e-mail: junping.chen@shao.ac.cn

Y. Zhang
College of Surveying and Geo-Informatics, Tongji University,
No. 1239, Siping Rd., Shanghai 200092, China

J. Chen
School of Astronomy and Space Science,
University of Chinese Academy of Sciences, Beijing 100049, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_22

1 Introduction

In the BeiDou Navigation Satellite System (BDS) broadcast messages, the orbit parameters is generated through orbit fitting processing using the precise orbits of ODTS (Orbit Determination and Time Synchronization) process [1–3], while the satellite clock parameter is generated based on the TWTT (Two-Way satellite Time Transfer) technique [1, 3, 4]. BDS's Legacy PNT service performance is 10 m (95%) in positioning precision, and 50 ns (95%) timing precision at in Asia-Pacific area [5].

To improve positioning accuracy and integrity, BDS integrates the wide area differential services together with the Legacy PNT system. Current differential corrections include the Equivalent Satellite Clock for the combined correction for satellite radial orbit and clock errors, and ionospheric grids for the ionospheric correction improvement for single-frequency users [6, 7]. The current differential corrections provide for authorized users only, and the positioning precision can be improved by 50% for dual-frequency user and 30% for single-frequency user, respectively [7]. The Equivalent Satellite Clock correction principally includes the mean radial component of orbits error in BDS monitoring area and the real-time satellite clock errors, thus it could not fully reflect the orbit errors in the Along-track and Cross-track directions and its differences maybe at decimeter level for users at boundary areas of BDS service region. To correct for this effect, the orbit corrections based on the combined the epoch-differenced carrier phase and pseudo-range is proposed [8]. However, these models do not actually take advantage of the high precision carrier phase, as pseudo-range may be seriously biased by multipath and channel biases. Under such strategies, the User Differential Range Error (UDRE) is at a level of sub-meter level, which will limit the applications of higher accuracy requirements.

Supporting user decimeter level positioning accuracy requirements, a new type of differential correction, namely the zone correction, is proposed in the upgrading BDS wide area differential service system [9]. The new zone correction is generated using the high precision carrier phase and corrects for the remaining common errors for specific zones. In this paper, we firstly review the definition and calculation of zone correction. Precise Point Positioning (PPP) models for single- and dual-frequency observations using the zone corrections are introduced. Performance of the zone-divided PPP is evaluated for single- and dual-frequency observations, and strategies in handling of tropospheric and ionospheric delays are discussed.

2 Zone Correction and Zone-Divided PPP

2.1 Zone Correction

Most current wide area augmentation system uses smoothed pseudo-ranges in the calculation of wide area differential message, which is rather reliable and efficient in

real-time processing. However, due to the relative large noise of pseudo-range observations and the impact of multipath errors, the accuracy of the calculated differential correction is around 0.5 m in terms of UDRE [7, 8], which limits the current performance at the level of few meters and could not meet the requirement of real-time high precision applications.

To further improve the accuracy of differential corrections, the concept of zone correction is proposed and implement in the new BDS wide area differential service system [9]. The core of the algorithm is that both the observation error of a satellite and the remaining atmosphere correction error for a specified regional are assumed largely to be the same and are grouped into a time varying correction. Applying the new concept, service areas of BDS are divided into several regions. The zone correction is super-posed on broadcast ephemeris, the Equivalent Satellite Clock and the orbit corrections. It is represented by the combined residuals of carrier phase observations, where the receiver clocks and partial ambiguity are included, and are combined through a comprehensive combination process of multi-stations. The steps of zone correction are as follows:

(1) Calculation of Ionospheric-Free carrier phase residuals for each station in one zone:

$$dL(i,t) = LC - \rho - \delta_{rec} + \delta^{sat} - \delta_{trop} - \delta_{rela} - \delta_{amb} - \delta_{ESC} - \delta_{orb} + \varepsilon \quad (1)$$

In Eq. (1), L is the phase residuals of frequency i at time t, ρ is the geometric distance between tracking station and satellite computed from broadcast ephemeris, δ_{rec} is the appropriate station clock, δ_{sat} is the satellite clock computed from broadcast ephemeris, δ_{trop} is the tropospheric delay, δ_{rela} is the correction of relatively, δ_{amb} is the appropriate satellite ambiguity through pseudo-range minus carrier phase, δ_{ESC} is the Equivalent Satellite Clock Correction, δ_{orb} is the orbit correction, ε contains the phase-windup corrections, solid tide correction, ocean tide correction and the observation noise. It should be pointed out that the satellite clock in broadcast ephemeris is based on B3 frequency and the satellite clocks differs between different frequencies [10, 11], and TGD correction should be corrected for observations of other frequency combination.

From Eq. (1) we can see that the residual of carrier phase contains the residual of station clock and partial ambiguity, together with observation noise, residual of satellite orbit error, satellite clock error and tropospheric model error.

(2) Calculation of epoch-differenced carrier phase residual for each station in one zone.

$$\Delta L(i,t,t-1) = \begin{cases} 0, & t = 1 \\ dt(i,t) - dt(i,t-1), & t > 1 \end{cases} \quad (2)$$

(3) Combination of zone correction

$$dL(t) = f(\Delta L(t, t-1)) + dL(t-1) \tag{3}$$

In Eq. (3), f means the combination function of comprehensive carrier phase correction. In most situations, f is the weighted average of each station. During the case of satellite disappearing, satellite arising, cycle slips or station clock jumps, the process of cycle slip repairing and clock jump repairing should be performed, otherwise the data should be abandoned.

2.2 Model of Zone-Divided PPP

Based on the broadcast ephemeris, the Equivalent Satellite Clock, the orbit correction and precise phase zone correction, real-time PPP can be realized for both dual-frequency and single frequency users.

(1) Dual-frequency user PPP

For B1B2 or B1B3 frequency user, the positioning model can be expressed as follow by using ionospheric-free combination.

$$\begin{cases} PC = \rho + \delta_{rec} - \delta^{sat} + \delta_{trop} + \delta_{rela} + \delta_{ESC} + \delta_{orb} + \varepsilon_{PC} \\ LC = \rho + \delta_{rec} - \delta^{sat} + \delta_{trop} + \delta_{rela} + \delta_{amb} + \delta_{ESC} + \delta_{orb} + dL + \varepsilon_{LC} \end{cases} \tag{4}$$

In Eq. (4), PC and LC are the ionospheric-free pseudo-range and carrier phase combination of B1B2 or B1B3. Comparing Eq. (1) with Eq. (4), we can see that although the broadcast orbit and clock is not precise after Equivalent Satellite Clock correction and orbit correction, the residual error will be further corrected using the zone correction. The station clock residuals contains in the zone correction will absorbed by user station clock parameters, while the partial ambiguity in the zone correction will absorbed in user ambiguity parameter. The change of satellite orbit and clock remaining errors vary little during short time (90 or 180 s).

In the above model, the residuals mainly come from the carrier phase observation noise and the remained differences between station depended troposphere delay and the troposphere delay contained in zone correction. As for pseudo-range observations, the zone correction can't be used because the containment of partial ambiguity, so only the Equivalent Satellite Clock and orbit correction can be applied.

To study the impact of the troposphere delay difference on dual-frequency zone-divided PPP, tropospheric delay parameters could be theoretically set up in Eq. (4).

(2) Single frequency user PPP

For single frequency user, the ionosphere error is the main factor that affects the positioning precision. Currently, the BDS provide 8 parameter ionosphere model for legacy navigation, as for augmentation positioning, the 14 parameter ionosphere model or the ionosphere grids is used [10]. However, the correction precision is about 0.5 m [12, 13]. In 2002, GAO proposed a positioning model called UofC [14]. In this model, due to the opposite of ionospheric delay in pseudo-range and carrier phase, the ionospheric error can be eliminated by averaging pseudo-range and carrier phase. Based on UofC, we establish a single frequency zone-divided PPP model.

$$\begin{cases} P = \rho + \delta_{rec} - \delta^{sat} + \delta_{trop} + \delta_{\text{iono}} + \delta_{rela} + \delta_{ESC} + \delta_{orb} + \varepsilon_P \\ (P+L)/2 = \rho + \delta_{rec} - \delta^{sat} + \delta_{trop} + \delta_{rela} + \delta_{amb} + \delta_{ESC} + \delta_{orb} + dL + \varepsilon_L \end{cases} \tag{5}$$

In Eq. (5), P and L is the single frequency pseudo-range and carrier phase observation, $\delta_{iono}$is the ionospheric delay using 14 parameter ionosphere model or ionosphere grids. The other parameters are the same with Eq. (4). Which should be pointed out is that δ_{amb} is actually half of the carrier phase ambiguity at single frequency.

To study the impact of the ionosphere delay on single-frequency zone-divided PPP, phase observations corrected for ionospheric delay using 14 parameter ionosphere model or ionosphere grids could also be directly combined with pseudo-range observations in Eq. (5).

3 Data Processing Strategy

In the implementation of the zone correction for BDS wide area differential service system, the service area is divided into 18 zones, where the zone correction for each zone is assumed to be calculated based on a pseudo reference station at the center of the zone. Users first calculate its approximate coordinates and searching for the nearest zone according to the given table of coordinates of each zone center. The zone corrections are broadcasted to user through GEO satellite at given defined epoch. To evaluate the precision and reliability of zone correction, the zone-divided PPP performance is analyzed.

Figure 1 shows the region of each zone. About 30 BDS tracking stations distributed in China are selected for zone-divided PPP assessment, which are also showed in Fig. 1. For each station, the zone centers within 1000 km are selected orderly. The mean distance for all stations and zone centers is 597 km. Table 1 gives the processing strategy of zone-divided PPP.

Data from DOY 346–348 in 2016 is chose. For each day, the data is divided into 4 parts every 6 hours. So there are 12 arcs processed in total. For each arc, the B1B2, B1B3, B1, B2, B3 zone-divided PPP are computed in kinematic mode.

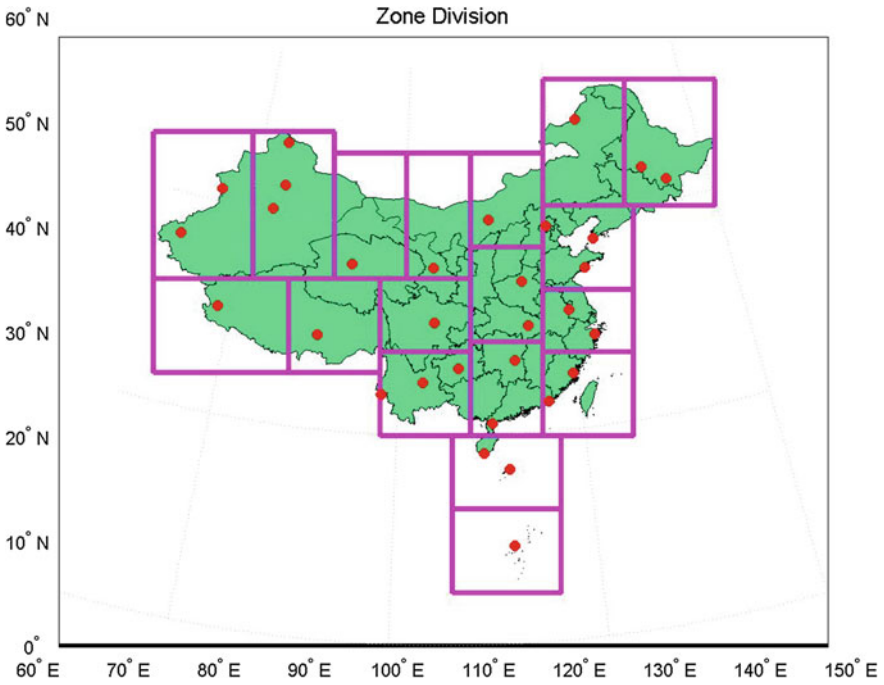


Fig. 1 Zone area and test stations

Table 1 Zone-divided PPP strategy

Type	Strategy
Estimator	Kalman filter
Satellite orbit and clock	Broadcast ephemeris
Augmentation message	ESC correction, orbit correction, zone correction
Data sampling	30 s
Limit elevation	10°
Ionospheric delay	Dual-frequency: Ionospheric-free combination; Single frequency: BDS ionosphere model
Tropospheric delay	GPT2w + SAAS + VMF1
Solid tide, ocean tide	IERS convention
Station coordinate	Estimated, white noise
Station clock	Estimated, white noise
Ambiguity	Estimated

4 Analysis of Zone-Divided PPP Performance

4.1 Effect of Troposphere Delay Parameter

Troposphere delay is one of the errors in GNSS observation. Troposphere models such as UNB3, EGNOS, GPT2, IGGtrop can estimate the hydrogen delay and most of wet delay, the precision of these models have been demonstrated to be 4–6 cm [15, 16]. The rest of the wet tropospheric delay is estimated as a random walk parameter in traditional PPP. As mentioned above, the zone correction contains tropospheric model error. After applying zone correction, the troposphere error would well be eliminated in areas near zone center.

To verify this, we make a comparison of zone-divided PPP with and without troposphere parameter estimation. The troposphere model of GPT2w is set as the initial value. Figure 2 shows the comparison of horizontal and vertical RMS distribution for B1B2 combined zone-divided PPP and Table 2 gives the statistical RMS of different data length. From the figures we can see that by fixing the troposphere delay, the precision of zone-divided PPP in height gets better. After fixing the troposphere delay, 87% of PPP result is better than 0.2 m in horizontal and 83% is better than 0.6 m in vertical for dual-frequency user. We can also find that there is no much difference on the precision during 2–4 and 4–6 h, which means the positioning error is already convergence (Table 2).

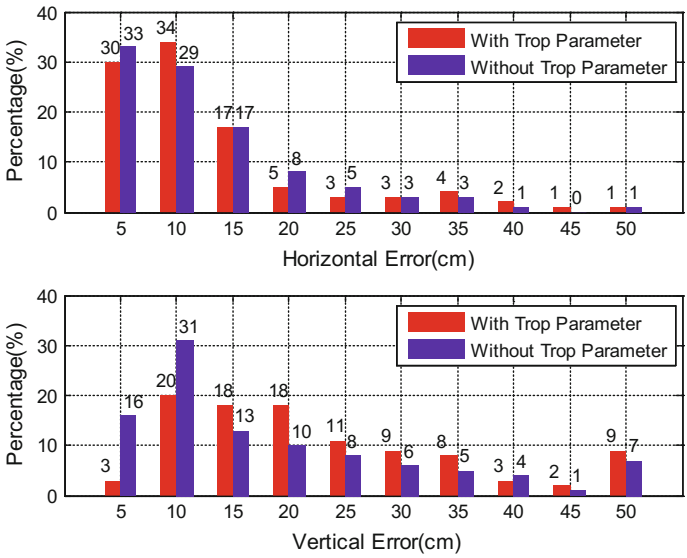


Fig. 2 Zone-divided PPP precision distributions at horizontal (*up*) and vertical (*down*) with and without troposphere parameter estimation

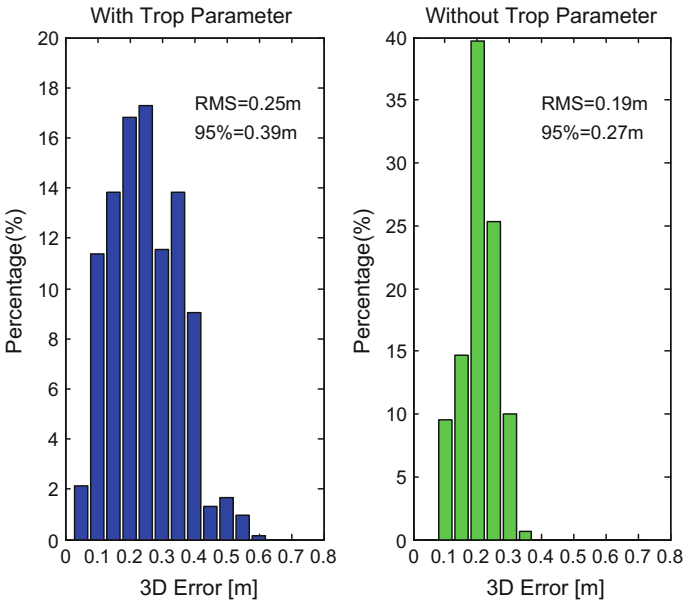


Fig. 3 Positioning error distribution of zone-divided PPP for B1B2 with and without troposphere parameter

Table 2 RMS of zone-divided PPP results at different data length

Type (h)	With troposphere parameter		Without troposphere parameter	
	Horizontal (m)	Vertical (m)	Horizontal (m)	Vertical (m)
1–2	0.17	0.28	0.17	0.24
2–4	0.11	0.22	0.11	0.18
4–6	0.10	0.22	0.10	0.17

To take a specific observation of positioning error distribution, the 3D positioning error distribution of zone-divided PPP at one station is compared by histogram statistics after one hour of convergence, showed in Fig. 3. The station is about 700 km away from the zone center. From the figure we can see that the 3D positioning RMS is 0.24 m and 95% is better than 0.39 m when estimating the troposphere parameter. While the 3D positioning RMS is 0.1 m and 95% is better than 0.27 m by fixing it. This proves that the zone correction contains part of tropospheric model error at a regional area.

4.2 *Effect of Single Frequency PPP Model*

As for single frequency user, one can use the traditional single frequency PPP model or the UofC model mentioned above. To evaluate the different of these two models, a comparison test is conducted. Figure 4 is a typical zone-divided PPP results for B1 frequency user of two models. From the figure we can see that the traditional model is much worse than UofC, which is due to that the ionosphere error is eliminated in UofC model by averaging pseudo-range and carrier phase.

To further evaluate the effect of zone correction on single frequency PPP, we also test the traditional single frequency PPP without zone correction. The statistical results for all stations are listed in Table 3. From the table we can see the advantage of UofC model. The RMS is 0.22 m in horizontal and 0.43 m in vertical. As for traditional PPP, the positioning precision also improves after applying zone correction, from 0.67/0.99 in horizontal and vertical to 0.42/0.87. To take a specific observation, the 3D positioning error distribution at one station is compared in Fig. 5. After applying zone correction, the 3D positioning RMS improves from 1.38 to 1.04 m, as for 95% positioning error, it improves from 2.03 to 1.59 m.

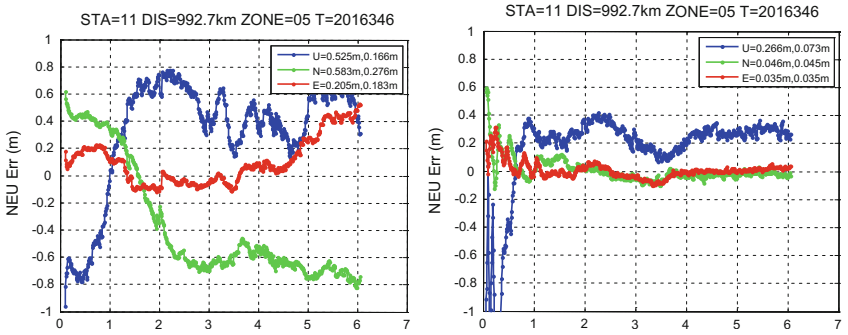


Fig. 4 Traditional single frequency PPP (*left*) versus UofC PPP (*right*)

Table 3 Single-frequency PPP for different model

Type	Traditional PPP without zone correction	Traditional PPP with zone correction	UofC PPP with zone correction
Horizontal (m)	0.67	0.42	0.22
Vertical (m)	0.99	0.87	0.43

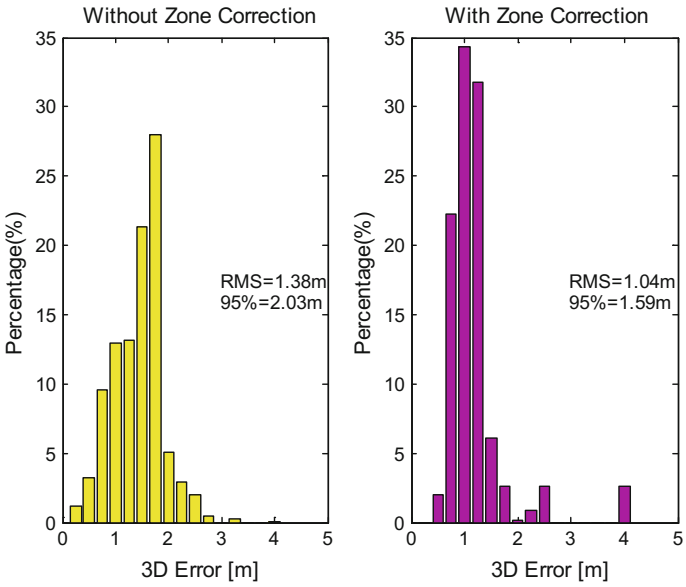


Fig. 5 Positioning error distribution of zone-divided PPP for B1 with and without zone correction

4.3 Overview of Zone-Divided PPP Performance

Based on the analysis of effect of troposphere delay parameter and single frequency PPP model analyzed above. We conduct the full experiment of zone-divided PPP for different frequencies. In the experiment, the troposphere error is fixed through GPT2w model. For single frequency positioning, the UofC model is applied. The RMS of zone-divided PPP is calculated from 4th to 6th hour. Table 4 gives the mean value of all stations and zones of different frequencies. We can see that for dual-frequency user, the mean kinematic positioning precision is below 0.15 m in horizontal and 0.20 m in vertical. As for single frequency user, the mean kinematic positioning precision is below 0.30 m in horizontal and 0.50 m in vertical, respectively. The single frequency zone-divided PPP performance is much worse than that of dual-frequency. The precision can be regarded as the zone-divided precision within 600 km.

For real-time kinematic user, the convergence performance is more concerned. To evaluate the convergence performance of all stations, we get the mean 3D positioning error of zone-divided PPP every 5 min at first one hour. Figure 6 gives

Table 4 RMS of zone-divided PPP results at different frequencies

	B1B2	B1B3	B1	B2	B3
Horizontal (m)	0.11	0.14	0.22	0.23	0.27
Vertical (m)	0.18	0.19	0.43	0.43	0.46

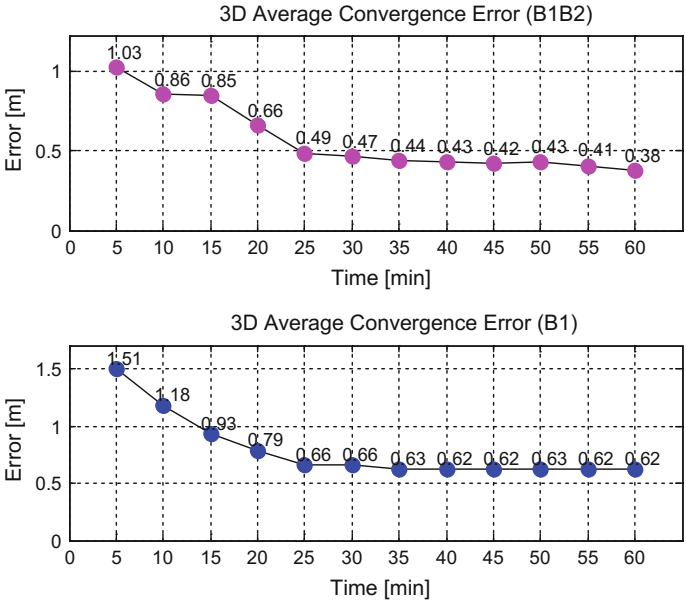


Fig. 6 Convergence statistics of zone-divided PPP for B1B2 (*up*) and B1 (*down*)

the convergence performance of zone-divided PPP for B1B2 and B1. From the figure we can see that it can convergence to 0.5 m in 25 min for dual-frequency user and to 0.8 m in 20 min for single frequency user.

5 Conclusions

In this paper we give a brief introduction of zone correction and the zone-divided PPP model for dual- and single-frequency user. Based on national distributed BDS tracking station in mainland China we evaluate the performance of dual- and single-frequency zone-divided PPP in different aspects. We conclude that:

(1) The precision of zone-divided PPP improves after fixing the troposphere delay.
(2) Zone correction improve the user positioning performance. UofC model is much better for single frequency user.
(3) For dual-frequency user, the mean kinematic positioning precision is below 0.15 m in horizontal and 0.20 m in vertical. As for single frequency user, the mean kinematic positioning precision is below 0.30 m in horizontal and 0.50 m in vertical, respectively.
(4) The zone-divided kinematic PPP can convergence to 0.5 m in 25 min for dual-frequency user and to 0.8 m in 20 min for single frequency user within 600 km. In general, the user can achieve a positioning precision better than 1 m within the distance of 1000 km from zone center.

Acknowledgement This work is support by 863 project (2014AA123102) and NSFC 11673050 and 11273046.

References

1. Zhou SS, Hu XG, Wu B et al (2011) Orbit determination and time synchronization for a GEO/IGSO satellite navigation constellation with regional tracking network. Sci China Phys Mech Astron 54:1089–1097
2. Zhou SS, Cao YL, Zhou JH et al (2012) Positioning accuracy assessment for the 4GEO/5IGSO/2MEO constellation of COMPASS. Sci China Phys Mech Astron 55:2290–2299
3. Tang CP, Hu XG, Zhou SS et al (2016) Improvement of orbit determination accuracy for BeiDou Navigation Satellite System with two-way satellite time frequency transfer. Adv Space Res 58(7):1390–1400
4. He F, Zhou SS, Hu XG et al (2014) Satellite-station time synchronization information based real-time orbit error monitoring and correction of navigation satellite in BeiDou System. Sci China Phys Mech Astron 57:1395–1403
5. Yang Y, Li J, Xu J et al (2011) Contribution of the compass satellite navigation system to global PNT users. Sci Bull 56(26):2813–2819
6. Wu XL, Zhou JH, Wang G et al (2012) Multipath error detection and correction for GEO/IGSO satellites. Sci China Phys Mech Astron 55:1297F02D1306, doi:10.1007/s11433-012-4741-6
7. Cao YL, Hu XG, Wu B et al (2012) The wide-area difference system for the regional satellite navigation system of COMPASS. Sci China Phys Mech Astron 55:1307–1315
8. Chen J, Yang S, Zhou J (In Press) A new pseudo-range/phase combined SBAS differential correction model. Acta Geod Cartogr Sin
9. Chen J, Zhang Y, Yang S, Wang J (2015) A new approach for satellite based GNSS augmentation system: from sub-meter to better than 0.2 meter era. In: Proceedings of the ION 2015 Pacific PNT meeting, pp 180–184
10. China Satellite Navigation Office (2013) BeiDou Navigation Satellite System signal in space interface control document open service signal (Version 2.0), 2013.12
11. Li H, Zhu W (2014) Inter frequency clock bias of BeiDou. Acta Geod Cartogr Sin 43 (11):1127–1131
12. Wu XL, Hu XG, Wang G et al (2013) Evaluation of COMPASS ionospheric model in GNSS positioning. Adv Space Res 51(6):959–968
13. Wu XL, Zhou JH, Tang B et al (2014) Evaluation of COMPASS ionospheric grid. GPS Solut 18(4):639–649
14. Gao Y, Shen X (2002) A new method of carrier phase based precise point positioning. J Navig 49(2):109–116
15. Bohm J, Moller G, Schindelegger M (2014) Development of an improved empirical model for slant delays in the troposphere (GPT2w). GPS Solut
16. Li W, Yuan YB, Ou JK et al (2012) A new global zenith tropospheric delay model IGGtrop for GNSS applications. China Sci Bull 57:2132–2139. doi:10.1007/s11434-012-5010-9

Part III
Multi-source Fusion Navigation Technology

Test Results of HiSGR: A Novel GNSS/INS Ultra Tight Coupled Spaceborne Receiver

Xiaoliang Wang, Deren Gong, Yanguang Wang, Bo Qu, Longlong Li, Xingyuan Han and Yansong Meng

Abstract With supporting of National Natural Science Foundation, Academy of Space Electronic Information Technology (ASEIT) is developing a novel compact spaceborne GNSS receiver; we call it as High Sensitive GNSS Receiver (HiSGR), that can operate effectively in the full range of Earth orbiting missions from Low Earth Orbit (LEO) to geostationary and beyond. Improved signal detection algorithms have been used in signal process section of HiSGR and inertial sensor is used for GNSS/INS ultra-tight coupled design that speed up the acquisition process and provide improved tracking performance for weaker GPS signals and in the presence of high dynamics shifts. HiSGR has experienced extensive tests to demonstrate the promising performance of some crucial specifications, by using real GNSS signal receive on open field and Hardware In-the-Loop (HIL) simulation. Receiver performance in LEO scenarios is provided. Finally, a ground vehicle running test is introduced for the demonstration of fast acquisition and reacquisition capability under conditions of signal lost. HiSGR demonstrated promising performance and acted stable during all those simulations and tests, which proved the capability for future space applications.

Keywords Ultra-tight coupled · Receiver · Spaceborne · Weak signal · Acquisition and tracking

X. Wang · Y. Wang · B. Qu · L. Li · X. Han · Y. Meng
Institute of Satellite Navigation and Intra-Satellite-Link Technology, Academy of Space Electronic Information Technology, Xi'an, China
e-mail: xlwang12321@gmail.com

D. Gong (✉)
School of Aeronautics and Astronautics, Shanghai Jiaotong University, Dongchuan Road No.800, Shanghai 200230, China
e-mail: drgong@sjtu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_23

1 Introduction

For recent decades, the use of GNSS receiver for space missions has becoming quite a common technique. Some major applications include [1]: Real-time Precise Orbit Determination (POD), attitude determination and raw measurement provider.

Since so many kinds of scientific demanding for space applications of GNSS receiver, several organizations have undertaken efforts to develop spaceborne GNSS receivers from Low Earth Orbit (LEO) to lunar missions in past two decades. Some famous products includes BlackJack (or TurboRogue Space Receiver, TRSR-2) [2, 3] receiver and some follow-on version developed by NASA's Jet Propulsion Laboratory (JPL) and Broadreach Engineering (BRE) [4], commercial-off-the-shelf (COTS) dual-frequency receiver OEM4-G2L from NovAtel Inc. for space applications, Septentrio's PolaRx2 receiver for LEO missions [5, 6].

ASEIT is currently developing a novel compact spaceborne GNSS receiver "HiSGR" for future space applications, under the supporting of National Nature Science Found, as shown in Fig. 1. This paper gives the introduction of HiSGR receiver in detail. The rest of this article is organized as follows: Sect. 2 provide the introduction of hardware and software architectures of HiSGR receiver. The simulation results using HIL and real signals are given in Sect. 3.

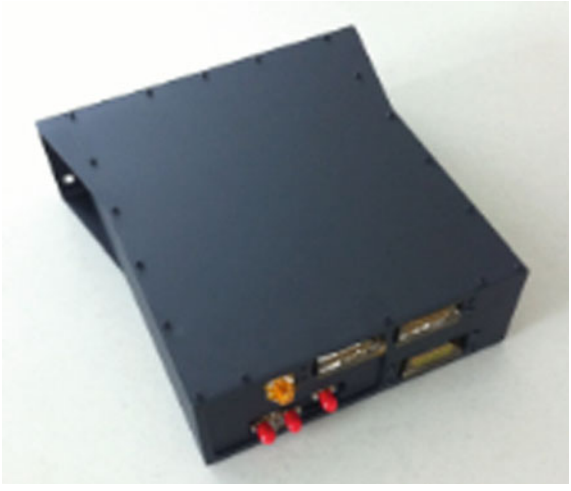


Fig. 1 HiSGR receiver with painted encapsulation

2 Receiver Architecture

2.1 Hardware

As shown in Fig. 2, the total receiver components are carefully layout in two Printed Circuit Boards (PCBs). The bottom layer includes (1) high sensitive front-end with low noise amplifier (LNA) that provides the down-converted GNSS signal. (2) DC-DC converter power supply module which provide the necessary second voltages to the other components from the primary 28 V DC power bus. (3) High stability frequency reference and (4) inertial measurements sensor. All the information will be transferred to top layer for post process. The output navigation solutions and necessary raw measurements will be delivering to remote-terminal Command & Telemetry (C&T) interface using MIL-STD-1553B communication bus, and finally proved to spacecraft GNC system through internal connector.

Two GNSS receive antennas are used, one for zenith pointing and one for nadir. Clearly this is from the requirement of high altitude spacecraft that get signal acquisition and tracking of GPS satellites through a down-looking direction. The two antennas can be switched on and off in front-end device automatically, according to the altitude of receiver [7], and the intermit frequency signals are finally provided. The current version of HiSGR is a GPS L1/L2 due frequency receiver that capable of receiving L1 C/A, L2C, L2 P(Y) code. Receiving BeiDou B1 and B3 signals will be realized in near future.

The FPGA1 in top layer is acquisition and tracking correlator FPGA controlled by an internal microprocessor, which is used for fast weak signal acquisition and tracking. This device undertaken the most computationally demanding task since it requiring a search across a three dimensional space of unknown time delay, doppler

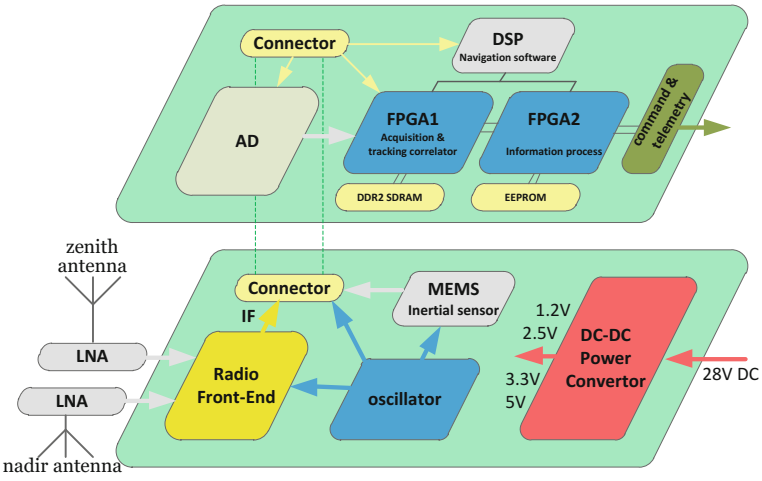


Fig. 2 HiSGR receiver hardware structure

shift, and satellite pseudo random noise number. It is also the critical factor to sensitivity of traditional receiver design. Improved signal detection algorithms have been used in HiSGR that speed up the acquisition process and provide improved performance for weaker GPS signals in presence of high dynamics. Typically, tracking of signals down to 25 dB-Hz is possible with standard GPS methods [8], while acquisition is limited to about 35 dB-Hz. Some special treatments of acquisition scheme are used for the HiSGR. First, the satellite selection logic from software section can predict the GNSS satellite passes and coarse doppler shift for any altitude and antenna orientations. Second, an ultra-tightly coupled GNSS/Inertial Kalman Filter is used in occasions as (or some scenarios during) dynamic maneuver and in challenging signal environment (during launch, orbit transitions and reentry), which also aid the signal fast acquisition or reacquisition.

The tracking correlator implements the code delay locked loop and carrier tracking functions on each of the channels with data output by the acquisition section. Those data are stored in a fast DDR2 SDRAM connected to FPGA1. By optimal design of the code and carrier tracking loops for the expected signal levels and dynamics in space, it is possible to increase the sensitivity of the receiver for tracking weaker GPS signals, at or just below the tracking threshold of typical receivers. Conventional GPS acquisition and tracking algorithm generally use 1 ms accumulated I and Q values, as this is convenient for detecting the data bit transitions. The longer the accumulation time, the lower the satellite C/N0 that can be detected. Detection thresholds are generally set at around 95% probability. This equates to a post-accumulation signal/noise level of around 8 dB. For the improvement of weak signal tracking performance in HiSGR, two methods are used: One by longer non-coherent accumulations and the other is adaptive tracking loop filter. The first is using 20 ms correlations for tracking weak signal, and even 50 non-coherent accumulations over the 1 s time interval using 20 ms coherent accumulations. The C/N0 detection threshold is achieved about 15 dB-Hz and even below in test. The adaptive tracking loop filter is also used for the reliable tracking and data demodulation of C/A signal, by estimation of best gain settings of tracking loops in real-time [9]. However, the filter method is far too computationally intensive for routine use, especially in strong signal scenarios.

FPGA2 is used for inertial measurements and aided almanac information collection from MEMS sensor in bottom layer and attached EEPROM. It also controls the navigation solutions output from DSP and handles commanding and telemetry messages thought MIL-STD-1553B interface ports.

2.2 Software

The DSP in the top layer of Fig. 2 is in charge of basic navigation software, which processes the received signal measurements to compute the navigation solutions. The core of navigation algorithm is achieved by Real-Time Precise Orbit Determination software (RTPODs). RTPODs is a compact and portable software

package optimized for real-time processing and is designed for use on embedded DSP systems. It makes use of an extended Kalman filter as well as precise dynamic models for orbiting receivers. The models used in RTPODs include a full EGM-96 70 by 70 gravity field, the DTM 94 atmospheric drag model [10], a solar radiation pressure model, Earth orientation and polar motion models as well as a relativity model. In addition, RTPODs has the capability to utilize the reduced dynamic technique [11] in which empirical accelerations are estimated in order to account for any dynamics left unmodeled. Finally, the Position, Velocity, and Timing (PVT) point solution will be provided when four or more satellites are being tracked, via a single-shot nonlinear least-squares solver, and also provides onboard orbit determination capabilities in sparse. When RTPODs operating in dual frequency mode, more observations are used to compute the ionospheric delay and ionosphere-free pseudoranges for each satellite in view.

The RTPODs mainly operate in GNSS stand-alone navigation mode, and can also be switched to integrated navigation mode that incorporate inertial measurements and on-line processed using ultra-tight coupled GPS/inertial Kalman Filter for predicted orbit position, velocity and attitude, which is used to remove the Doppler effects to allow weak signal tracking in some scenarios.

The availability of the source code allows customization of RTPODs for space applications in difference altitude. So far, several modifications have been made for RTPODs which covers LEO to Lunar missions (Table 1).

3 Tests and Simulations

HiSGR has experienced extensive tests to demonstrate the promising performance of some crucial specifications, by using Hardware In-the-Loop (HIL) simulation and real GNSS signal receive from moving vehicle in open field and underground.

The following subsections describe the HIL simulation configurations, general HiSGR receiver performance, data processing and preliminary results.

Table 1 Some RTPODs versions

RTPODs versions	Applications
RTPODs-LEO	Low earth orbit
RTPODs-R	Reflected signal receiving
RTPODs-HEO	High eccentric orbit
RTPODs-GEO	Geostationary orbit
RTPODs-LUNAR[a]	Cislunar/orbiting moon
RTPODs-L2[a]	Helo Lagrange L2
RTPODs-T[a]	Test for ground vehicle

[a]Under development

3.1 HIL Simulation for LEO

HIL testing is useful to evaluate actually performance of receiver under varies conditions including dynamics, signal levels, and error sources on orbit. The RF input of the receiver is connected to a GPS simulator in such a test rather than a real antenna. The specified trajectory and attitude could also be provided from simulator that simulates the receiver was actually in motion. Aspects of the receiver performance could be evaluate repeatly that would be impossible through terrestrial static tests.

ASEIT has recently purchased a state-of-the-art Spirent GNSS simulator of model GSS9000 (as shown in Fig. 3, with HiSGR connection) which is a multiple constellation/ frequency GNSS simulator (GPS/Galileo/BeiDou/GLONASS/SBAS/QZSS). It has up to 16 parallel channels per carrier through each RF output. The GNSS simulator allows users to control virtually all kinds of the simulated GNSS signal properties and modeled error sources.

The LEO scenario of CHAMP (Challenging Minisatellite Payload) is used in this simulation, as orbit elements given as: Eccentricity 0.00398, Inclination 87.272°, Perigee altitude 429 km, Apogee altitude 476 km, Longitude of ascending node 150.089°, Argument of perigee 323.038°.

The receiver antenna is configured to be fixed in the mass center of spacecraft, with orientations of zenith direction. The receiving antenna gain patterns can be configured manually using antenna attenuation pattern editor in GSS9000 simulator and the default gain patterns is used during in this simulation (default_v1-0.ant_pat file in setting GUI).

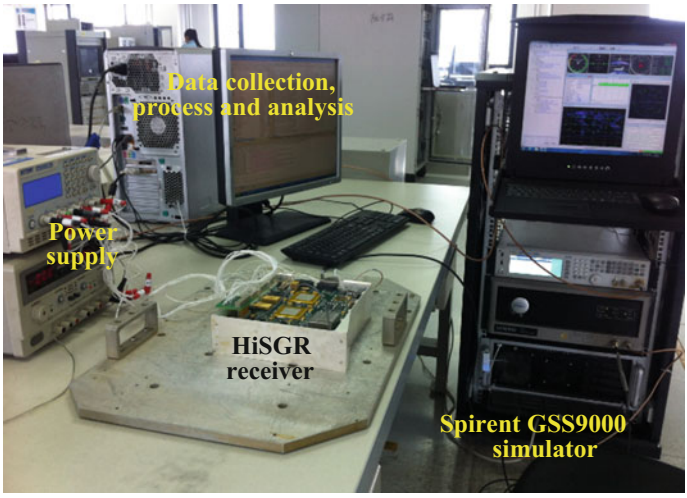


Fig. 3 HiSGR receiver under HIL test

Only GPS constellation is used in the test. The signal levels in the simulator are default to −130 dBm that comply with the minimum guaranteed signal strength from GPS ICD-200 [12, 13].

The recommended GPS reference gain pattern is also provided in the simulator setting GUI (default_v1-1.ant_pat file). However, the default setting file failed to provide the accuracy GPS gain pattern for signal transmitter. Several references have mentioned the actual measurement of GPS antenna gain pattern for deep research [14, 15]. Here we use the recent published results for testing [16]. The actual antenna gain data of Block IIF were adopted, as shown in Fig. 4, and transformed to the ant_pat file, importing to simulator for testing.

Simulation time is set from 14:15:00 to 20:15:00 08/Apr/2016 UTC, with sampling step of 1 s. The GPS reference orbits are based on precise ephemeris which download as version 3.01 renix files from IGS web site [17]. It makes the simulated GPS orbits and clock parameters to closely match the actual GPS constellation at the time of the simulation. Software version RTPODs-LEO is burned into the hardware components which is a GNSS only version without inertial assistant. The signal acquisition and tracking algorithm choose 1 ms correlations for this test.

HiSGR performed excellent during the whole simulation time. Figure 5 illustrated the PDOP values and histogram of satellite geometry visible versus satellites get tracked data from the output from RTPODs-LEO software. The POD performance after filter convergence is given in Fig. 6 (using 15 s sample interval points for concise). The position errors in radial/in-track/cross-track directions are just about a few meters in real-time process, which is suitable for most missions in LEO with real-time requirements.

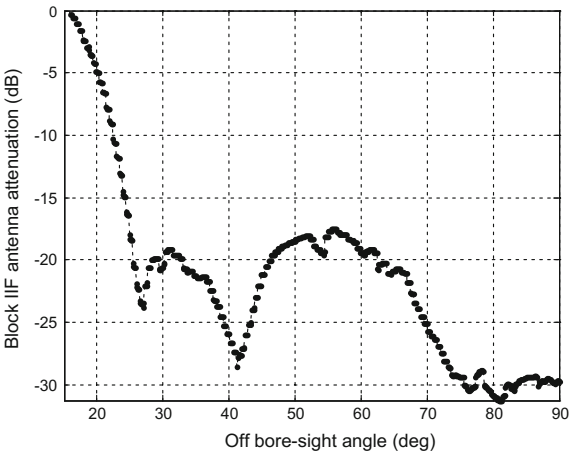


Fig. 4 GPS block IIF antenna attenuation

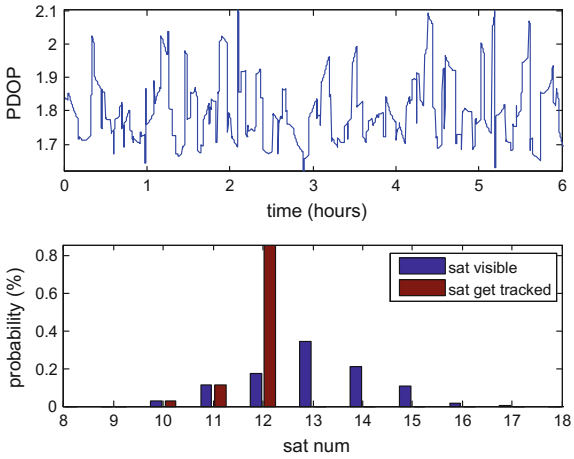


Fig. 5 PDOP and GPS visible analysis

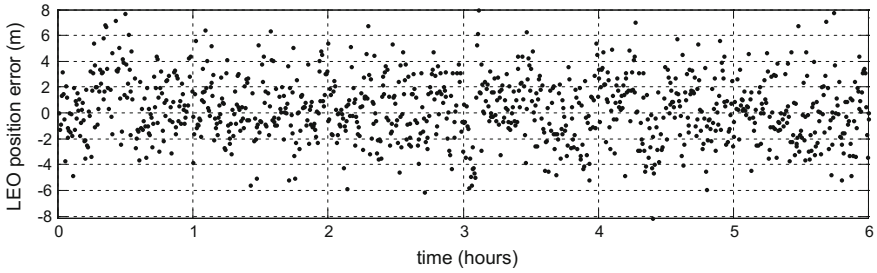


Fig. 6 HiSGR POD position error from RTPODs-LEO software

3.2 Vehicle Test on Ground

In this section, here we provide some brief introduction of the testing results using moving vehicle both on open field and underground, as shown in Fig. 7.

Software version RTPODs-T is used to verify algorithm of ultra-tight coupled integrated navigation design for weak signal acquisition and tracking under varies conditions, currently is under development with limited functions. However, the signal process section is achieved with inertial assistants, and the positioning computation is conducted offline using sampled data during test. Figure 7a demonstrated the assembling of HiSGR in platform with NovAtel's SPAN@ integrated navigation system, which is used for the comparison and evaluation of final solutions. Figure 7b, c shown the basic instruments configurations during running test, and Fig. 7d demonstrated the moving vehicle underground, which

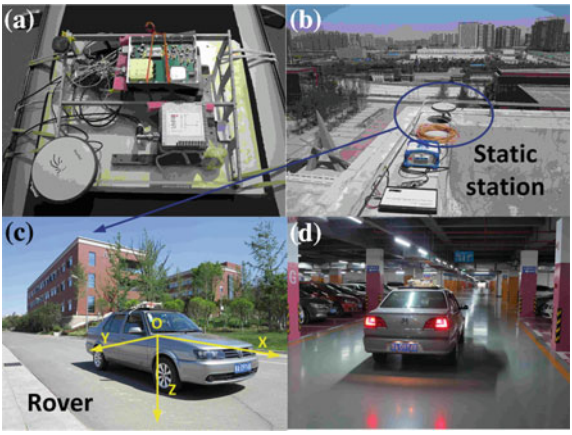


Fig. 7 Vehicle tests on open field and underground

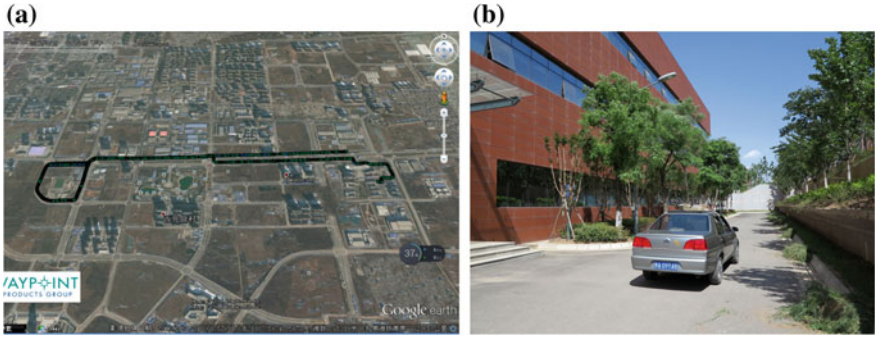


Fig. 8 Vehicle moving trajectory (**a**) with signal occulted routing (**b**)

aiming the assessment of receiver performance when no visible satellites available in space and fast signal reacquisition.

HiSGR performed stable during the whole simulation time. The receiver achieved quick signal reacquisition when out of underground garage. Figure 8 demonstrated the HiSGR test during open field. The reference trajectory is provided by the NovAtel's SPAN@ integrated navigation system.

By using GNSS standalone mode, The HiSGR positioning error is less than 10 m without occlusion, but diverged gradually when running under occulted buildings, as in Fig. 8b. However, the positioning error reduced rapidly, about less than 15 m under occlusion, when HiSGR switched to GNSS/INS ultra-tight coupled mode.

4 Conclusions

This article provided detail introduction of design and development of a novel compact spaceborne GNSS receiver "High Sensitive GNSS Receiver (HiSGR)", which conducted by ASEIT. HiSGR is a space qualified multifunctional GNSS receiver that can operate effectively in the full range of Earth orbiting missions from Low Earth Orbit to geostationary and beyond. Improved signal detection algorithms have been used in signal process section of HiSGR that speed up the acquisition process and provide improved tracking performance for weaker GPS signals. A compact and reconfigurable software package named Real-Time Precise Orbit Determination software (RTPODs) is used in HiSGR that optimized designed for use on embedded systems. HiSGR has experienced extensive tests by using Hardware In-the-Loop simulation and real GNSS signal receive from moving vehicle in open field and underground. The real time POD positioning error is about a few meters in LEO and less than 10 meters on ground test. Simulation results surely demonstrated the promising performance that suitable for future space applications.

Acknowledgements This article and the HiSGR project are supported by National Natural Science Fund under contract No. 91438107. The authors greatly thank Dr. Zhe Su, Dr. Bo Hao, Dr. Wenying Lei, Dr. Ying Wang, Dr. Tao Yan, Dr. Zhaoping Wu, Dr. Zhao Tong, Prof. Lang Bian for the helpful comments to this article and the assistant during the ground vehicle test of HiSGR receiver.

The authors also would like thank to all the other individuals that provide endeavor efforts to HiSGR project: Prof. Xingwang Zhong, Prof. Xianan Zheng, Prof. Xiaoxia Tao, Dengfeng Wang, Wei Xing, Kang Zhang.

References

1. Bauer FH, Hartman K, Lightsey EG, Spaceborne GPS: Current status and future visions. In: Proceedings of the Institute of Navigation GPS 98 conference, Nashville, TN, Sept 1998, pp 1493–1508
2. Montenbruck O, Kroes R (2003) In-flight performance analysis of the CHAMP BlackJack GPS receiver. GPS Solut 7:74–86. doi:10.1007/s10291-003-0055-5
3. Reichert A, Meehan T, Munson T (2002) Toward decimeter-level real-time orbit determination: a demonstration using the SAC-C and CHAMP spacecraft. In: ION-GPS 2002, Portland, OR, Sept 24–27, 2002
4. Wu B-H, Chu V, Chen P, Ting TH (2005) FORMOSAT-3/COSMIC science mission update. GPS Solut 9:111–121
5. Langley RB, Montenbruck O, Markgraf M, Kang CS, Kim D (2004) Qualification of a commercial dual-frequency GPS receiver for the e-POP platform onboard the Canadian CASSIOPE spacecraft. In: Second ESA workshop on satellite navigation user equipment technologies, NAVITEC'2004, Dec 8–10, 2004, Noordwijk, The Netherlands
6. Leyssens J, Markgraf M (2005) Evaluation of a commercial-off-the-shelf dual-frequency GPS receiver for use on LEO satellites. In: ION GNSS 2005 conference, Sept 13–16, 2005, Long Beach, California

7. Wang X et al (2015) Analysis of GNSS signals with application to lunar navigation. In: China satellite navigation conference (CSNC) 2015 proceedings: vol I, Chap 18, May 2015, Xi'an, China, pp 1–8
8. Van Dierendonck A (1996) Global positioning system: theory and applications. In: GPS receivers, vol 1, Chap 8, The American Institute of Aeronautics and Astronautics, Reston, VA, pp 329–407
9. Wang Y, Tian J, Leclère J, Botteron C, Capuano V (2014) An efficient time-frequency algorithm for weak signal acquisition of modernized GNSS signals. ION GNSS 7(6):715–722
10. Berger C, Biancale R, Ill M, Barlier F (1998) Improvement of the empirical thermospheric model DTM: DTM94—a comparative review of various temporal variations and prospects in space geodesy applications. J Geod 72(3):161–178
11. Wu S-C, Yunck TP, Thornton CL (1991) Reduced-dynamic technique for precise orbit determination of low earth satellites. J Guid Control Dyn 14(1):24–30
12. Anon (1993) GPS interface control document. ICD-GPS-200, Revision C, U.S. Air Force, 10 Oct 1993
13. Ward P (1996) GPS satellite signal characteristics, Chap. 4. In: Kaplan E (ed) Understanding GPS: principles and applications, Artech House Publishers
14. Bauer FH, Moreau MC, Dahle-Melsaether ME, Petrofski WP, Stanton BJ, Thomason S, Harris GA, Sena RP, Parker Temple L III (2006) The GPS space service volume. In: ION-GNSS-2006, Fort Worth, 26–29 Sept 2006
15. Force Dale A (2013) Individual global navigation satellite systems in the space service volume. In: Proceedings of the 2013 international technical meeting of the Institute of Navigation, San Diego, California, Jan 2013, pp 604–716
16. Martzen PD, Highsmith DE GPS antenna characterization experiment (ACE): receiver design and initial results. NASA Technical Reports Server (NTRS)
17. IGS web site: http://igscb.jpl.nasa.gov/. Accessed 18 Jan 2017

Multi-period PMF + FFT Acquisition Method Based on Symbol Estimation

Wen Liu, Tian-tong Gao, Zhong-liang Deng and Di Zhu

Abstract Due to the relative motion between the satellite and the target, there is a large Doppler shift in the high dynamic navigation and positioning system. Using a set of partial correlators to combine the fast Fourier transform for PN code acquisition, the two-dimensional search of the code phase-Doppler frequency can be reduced to a one-dimensional search, which reduces the computational burden. Aiming at the problem that the traditional PMF + FFT method is affected by the bit-flipping only for single-cycle acquisition, this paper presents a multi-period PMF + FFT interoperability capture method based on symbol estimation. In the case of bit flipping, the ratio of the main peak to the secondary peak in the acquisition result is improved by 15%, which has better anti-bit transition ability and make multi-period capture more feasible, enhancing the capture effect.

Keywords Matched filter · FFT · Symbol estimation · Multi-period

1 Introduction

The primary task of GNSS receivers is to capture and track satellite signals. The design and implementation of GNSS signal acquisition strategy is one of the core contents of receiver development. The use of spread spectrum ranging code GNSS signal reception power is low, so the receiver must first signal pseudo-code-related despreading. At the same time, due to the uncertainty of the Doppler shift, the two-dimensional search must be performed at fixed intervals on the code phase and frequency. The acquisition time of different search strategies is not the same. The speed of the whole parallel mode is the fastest and the speed of the whole serial mode is the slowest. At present, the PMF + FFT signal acquisition method has been

W. Liu · T. Gao (✉) · Z. Deng · D. Zhu
Beijing University of Posts and Telecommunication, Beijing, China
e-mail: 513503813@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_24

widely used in GNSS, and this method has certain advantages under high dynamic conditions [1]. The classical PMF + FFT method is only applicable to single-cycle acquisition, which limits the capture sensitivity of the inter-operable signal, because bit-flipping affects the multi-period coherent integration results. In this paper, we introduce a symbol-based multi-period PMF + FFT estimation method, and compared with other existing acquisition strategy.

2 PMF + FFT Acquisition Algorithm

The traditional PMF + FFT acquisition method is described in detail in the related literature [2, 3]. It consists of two steps: The first step is to segmentally correlate the signal using a correlator that contains P in-phase and quadrature branches based on a partially matched filter. Suppose that the length of one correlator is X, the product of the number of correlators and the correlator length is XP = M. The first correlator correlates with the first X chips of the pseudocode sequence, the second correlator correlates with the second set of X chips using the pseudocode sequence, and so on. The output of P correlators is the partial correlation result. If the P results are coherently combined, they are equivalent to the result of a perfectly matched filter with standard length M.

In the second step, K-point FFT is performed on the P correlation results. The PMF + FFT-based acquisition method does not coherently combine the P results, but uses them as an input of the N-point FFT ($N \geq P$), and performs Doppler frequency estimation by performing FFT output maximum detection. Where the maximum value of the FFT output will be closest to the actual Doppler frequency.

In the case of multi-period acquisition, the conventional method can cause the coherent accumulation of the matched filter to be inaccurate due to bit jump. In addition, by the time limit of integration, it is difficult to achieve the desired effect in the case of weak signal.

3 Multi-period PMF + FFT Acquisition Method Based on Symbol Estimation

The IF signal is mixed to obtain a baseband signal. The baseband signal first enters the partial matching module. P correlation results can be obtained for every symbol period. The correlation results are entered into a symbol estimation module, which performs bit inversion detection on the P correlation results of the adjacent two symbol periods one by one, and counts the bit flip. The bit count information is used to accumulate the correlation result corresponding to each period. The

correlation result after polarity accumulation is input to the FFT module and transformed into the frequency domain. The maximum value obtained is compared with the local capture threshold in the capture decision to obtain the final capture result.

3.1 Partially Matched Modules

Similar to the traditional PMF-FFT method, the received baseband signal is segmented and correlated, and P correlation values can be obtained for each received code period. The resulting value is denoted by i_m, where i represents the code period number and m represents the corresponding partial match result number. Under the effect of the pseudo-code clock, the code shift controller controls the local code and the received code to partially accumulate. Each cycle can be P partial correlation matching results (Fig. 1).

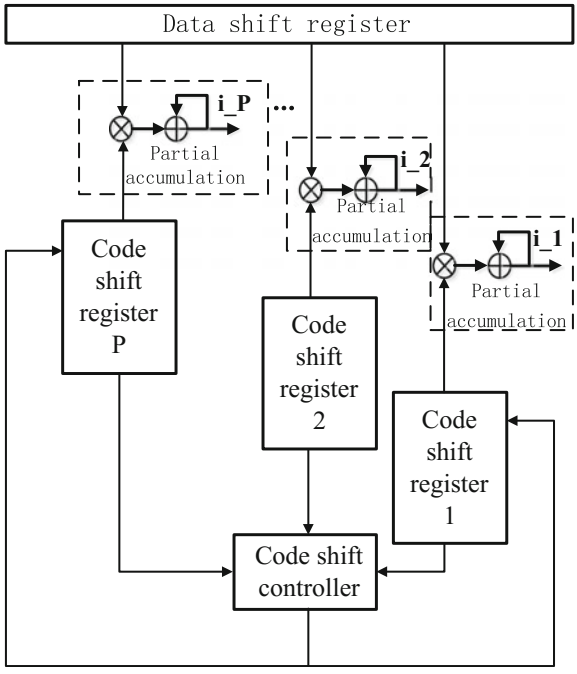


Fig. 1 PMF schematic

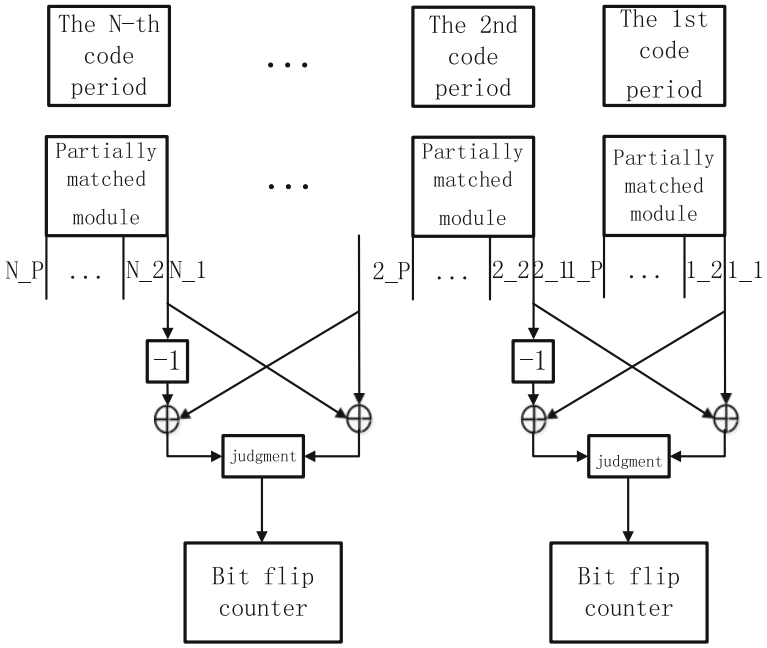


Fig. 2 Partial matching module based on symbol estimation

3.2 Symbol Estimation Module

Similar to the traditional PMF-FFT method, the received baseband signal is segmented and correlated, and P correlation values can be obtained for each received code period. The resulting value is denoted by i_m, where i represents the code period number and m represents the corresponding partial match result number. Under the effect of the pseudo-code clock, the code shift controller controls the local code and the received code to partially accumulate. Each cycle can be P partial correlation matching results.

Using the result of partial matching I_P, symbol decision is made on the decision result of adjacent two periods. As shown in Fig. 2, the symbol estimation step requires comparison of the results of the co-directional and reverse paths of the adjacent two periods.

The M-th result of each cycle and the in-phase result of the adjacent period are:

$$\begin{aligned} I_m^2 &= 1\_m + 2\_m \\ I_m^3 &= 2\_m + 3\_m \\ &\vdots \\ I_m^N &= (N-1)\_m + N\_m \end{aligned} \tag{1}$$

The same results can be obtained with the same cycle:

$$\begin{aligned} I_m'^2 &= 1\_m - 2\_m \\ I_m'^3 &= 2\_m - 3\_m \\ &\vdots \\ I_m'^N &= (N-1)\_m - N\_m \end{aligned} \tag{2}$$

Each group of results need to be judged, the decision expression is as follows:

$$T_m^i = I_m'^i - I_m'^i \tag{3}$$

$T_m^i < 0$ indicates that a flipping occurs at the m-th partial matching portion of the (i − 1)-th cycle and the i-th cycle. Otherwise, the (i − 1)-cycle and the i-cycle are not flipped. A set of adjacent sets of P-group correlation results are counted. The count result is set to J. When one of the set of correlation results is judged to be flipped, J is incremented by 1, otherwise J is decremented. In the P group of relevant results are judged to be the end of the results of J to make judgments. If J > 1, then the first cycle and the second cycle of the final decision for the occurrence of bit flip, if J < 1, then the first cycle and the second cycle of the final decision for a bit flipped. If the bit is flipped, the bit flipping counter is incremented by 1, otherwise it remains unchanged. After each successive cycle decision, the result of the current bit transition counter needs to be saved as X_i ($X_1 = 0$).

When the symbol estimation for all cycles is completed, the correlation result for each period is processed according to the bit counter: when the bit counter is even, the correlation result is not changed. When the bit counter is an odd number, the correlation result is inverted. And then the polarity of the correlation result corresponding to each cycle is accumulated, and P accumulation results are obtained:

$$I^m = \sum_{i=1}^{i=N} i\_m * (-1)^{X_i} \tag{4}$$

3.3 FFT Module

Fast Fourier transform (Fast-Fourier-Transform) is a fast algorithm for discrete Fourier transform. The number of fast Fourier transforms needs to be an integer power of 2. This method performs K-point FFT transform on the P-valued coherent accumulative values obtained by the symbol estimation module. If P < K, the sum of zero padding and P should be exactly equal to K. Finally, the FFT results are output to the capture decider (Fig. 3).

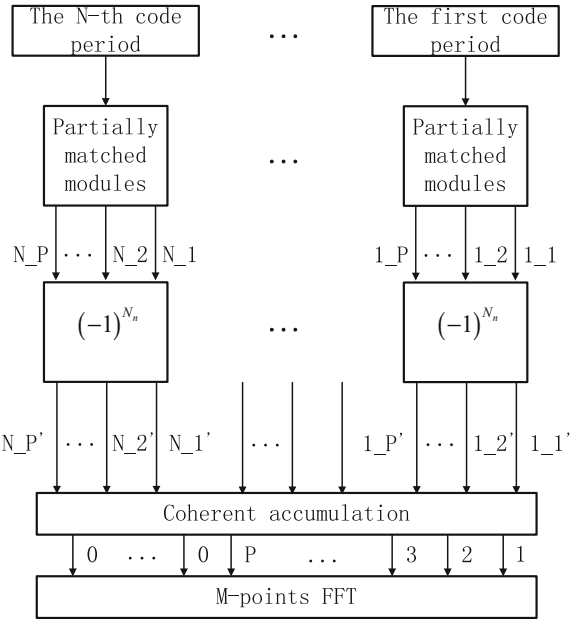


Fig. 3 Multi-symbol periodic FFT transform

4 Capture Performance Simulation

The pseudo-code rate of GPS signal is 1.023 MHz. Assuming the width of AD is 4 bits, the signal sampling rate is 62 MHz, the actual GPS signal is received by the antenna, the sampling time is 2 ms. Using MATLAB to deal with the data under down-conversion and AD sampling. The following Figs. 4 and 5 show the capture result of bit flipping situation and the absence of bit flipping situation respectively.

In the absence of bit flipping, 1 ms of data is captured using the traditional PMF + FFT method, with a ratio of peak to sub-peak of 1.48; 2 ms of data is captured using a symbol-based multi-period PMF + FFT capture method, The ratio of the highest peak to the second peak was 1.51, and the latter increased by 2% compared with the former. As a result of multiple cycles of signal capture, the effect of the capture has improved (Figs. 6 and 7).

In the presence of bit flipping, both the conventional PMF + FFT method and the symbol-based multi-period PMF + FFT acquisition method are used to sample 2 ms data. Traditional PMF + FFT method of the highest peak and the ratio of sub-peak of 1.18, 80% of the original, a larger decline. The ratio of the peak to the sub-peak of the multi-period PMF + FFT method based on symbol estimation is 1.39, which is 92%, and the latter is 15% higher than the former.

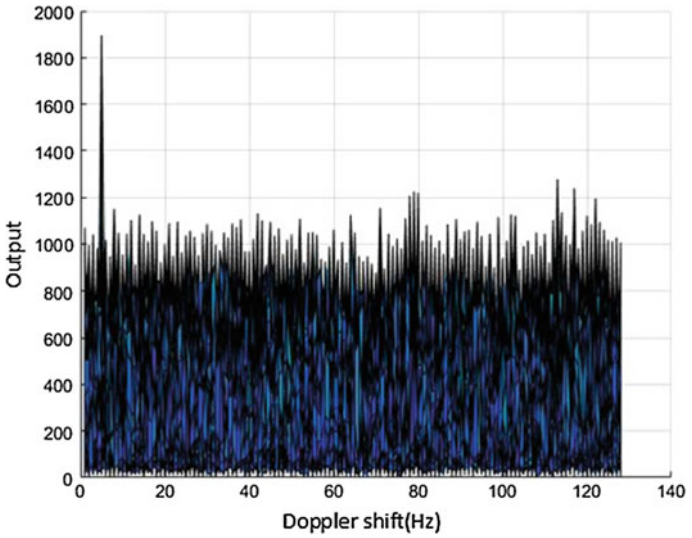


Fig. 4 The conventional PMF + FFT acquisition result (Doppler shift) in the case of no bit flipping

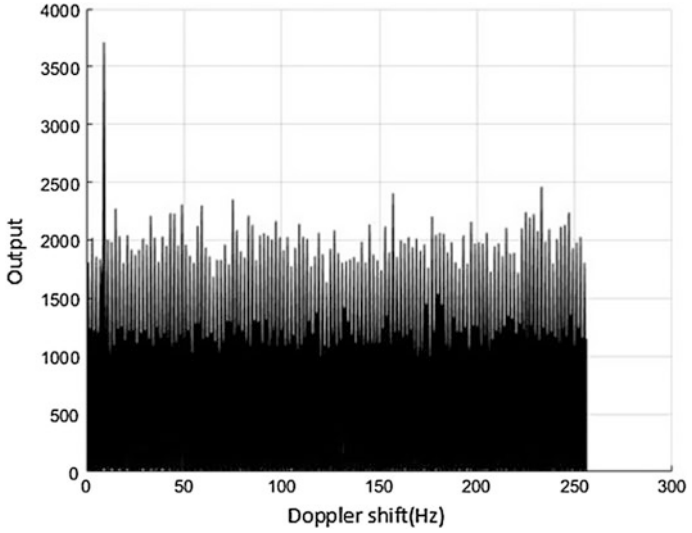


Fig. 5 Symbol-based estimation without bit-flipping PMF + FFT acquisition results (Doppler shift)

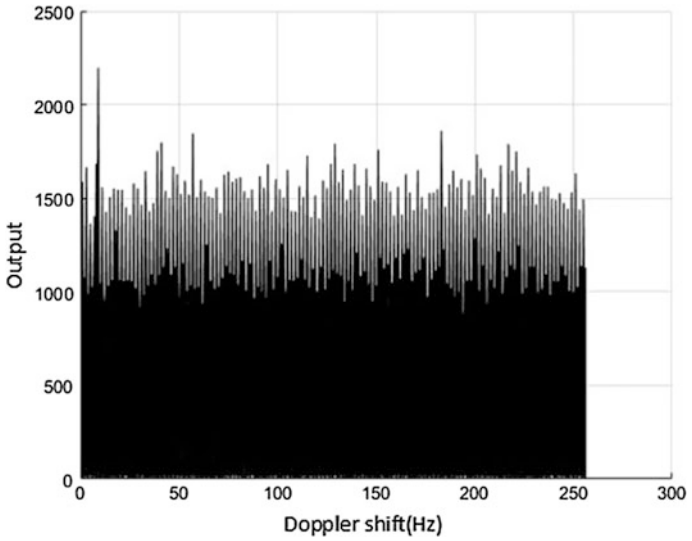


Fig. 6 The conventional PMF + FFT acquisition result (Doppler shift) in the case of bit flipping

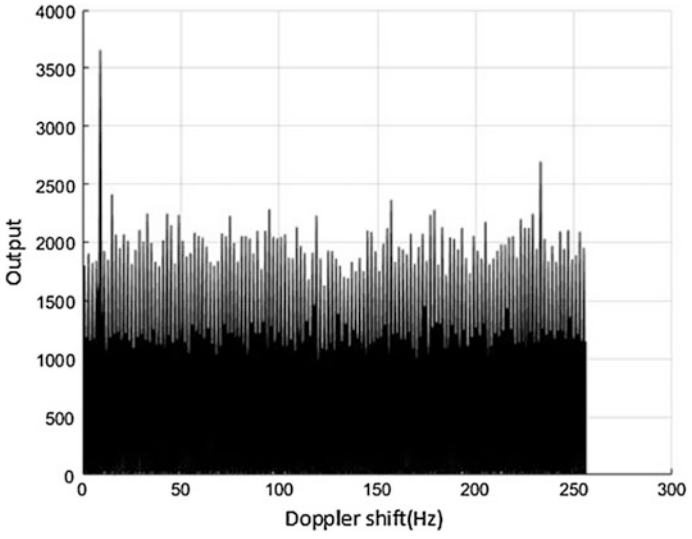


Fig. 7 Symbol-based PMF + FFT acquisition results in case of bit flipping (Doppler shift)

5 Conclusion

In the presence of bit flipping, it can be seen that the multi-period PMF + FFT acquisition method based on symbol estimation has a certain anti-bit-flipping capability compared with the conventional PMF + FFT acquisition method, which makes it possible to capture multi-period signals using the PMF + FFT method. Multi-period coherent accumulation of the gain will bring a certain increase in the capture effect.

Acknowledgements This work is supported by the National Key Research and Development Program (NO. 2016YFB0502001), National Natural Science Foundation of China (NO. 61372110), National High Technology Research and Development Program of China (No. 2014AA123103)

References

1. Deng Z, Zhu D, Yin L (2015) Research on pseudo-random code acquisition based on dimensionality reduction PMF–FFT. In: China satellite navigation conference
2. Ding S, Liu H, Liang X (2012) Analysis and improvement of detecting performance of PMF + FFT acquisition algorithm in spread spectrum communication system. High-tech Commun
3. Xu D, Li M, Wang S (2013) A fast acquisition method of spreading PN code based on PMF + FFT 32(6):P51–54

The Cubature Kalman Filter and Its Application in the Satellite Star-Sensor/Gyro Attitude Determination System

Xiaobo Yuan, Chao Zhang and Chunling Shi

Abstract As the satellite attitude determination system is a typical non-linear system, the application of non-linear Kalman filter in the satellite attitude determination system is necessary. Aiming at the satellite Star-sensor/Gyro attitude determination system, the application of Cubature Kalman filter in satellite Star-sensor/Gyro attitude determination system is discussed in this paper. Firstly, 2n cubature point are generated basing on the cubature principle; then transform the points into new ones by non-linear system equations; finally, apply the new points to the update of system state. Simulation results show that CKF is better than EKF and UKF in accuracy and stability.

Keywords The satellite Star-sensor/Gyro attitude determination system · Non-linear Kalman filters · CKF

1 Introduction

The accurate measurement of satellite attitude is a prerequisite for realizing precise attitude control. It is commonly used in the satellite attitude measurement that measure the attitude with star sensor and gyro, then the filtering algorithm is adopted to information fusion processing, which gain a higher accuracy than single measurement method. When the attitude sensors are certain, choosing the appropriate filtering algorithm is the key factor to improve the attitude accuracy of the satellite.

Due to the nonlinearity of the satellite attitude measurement system, we need a nonlinear filtering algorithm. The Extended Kalman filter [1] (EKF) was firstly proposed to solve non-linear problem. The principle is that linearize the nonlinear

X. Yuan (✉) · C. Zhang · C. Shi
Institute of Navigation and Aerospace Engineering,
Information Engineering University, No. 62 Kexue Street,
Zhengzhou 450001, Henan, China
e-mail: 675050583@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_25

systems by Taylor expansion to solve the non-linear problem which Kalman filter cannot do. By now, EKF has been one of the most widely used algorithm in engineering. EKF is simple to calculate and easy to implement, which are its main advantages, while the calculation of Jacobian matrix and instability problem caused by the truncation error in strong nonlinear problems are disadvantages. Basing on the idea that nonlinear estimation algorithm is more approximate to linear system than linear estimation, Julier and other scholars put forward the unscented Kalman filter [2] (UKF) in 1995. According to deterministic sampling strategy, basing on UT exchange and taking the Kalman filter as the framework, the UKF algorithm approach the posteriori distribution of the state directly. Its calculation amount is equivalent to EKF, and its performance is better than EKF [3]. But it is still not suitable for dealing with strong nonlinear problems. In 2009, a Canadian scholar named Arasaratnam proposed a new nonlinear filtering algorithm—the Cubature Kalman filter [4] (CKF). By using spherical radial Cubature criteria for the corresponding integral calculation, CKF is currently the closest approximation algorithm to Bayesian filtering. It not only overcomes the application limitation of EKF and UKF in strong nonlinear system, but also has higher filtering accuracy than the central difference Kalman filter (Central Difference Kalman filter, CDKF) and the Particle filter [5, 6]. It is a new powerful tool to solve the state estimation problem of nonlinear systems, widely used in navigation and positioning [7], target tracking [8], robot [9], aerospace attitude determination [10] and other fields.

Aiming at the selection of filtering algorithm for satellite integrated attitude determination system based on star sensor/gyro, in this paper, the measurement equation and equation of state are derived, and new integrated filtering based on CKF principle is designed, then we compared CKF with EKF and UKF. The simulation results show that the accuracy of CKF is the highest in the three kinds of algorithms.

2 Satellite Integrated Attitude Determination System

2.1 The Theory of Integrated Attitude Determination

As shown in Fig. 1, the working principle of the satellite integrated attitude system based on star sensor/gyro is as below:

(1) The gyro is used as a short-term attitude reference to estimate the attitude of the satellite;
(2) The star sensors are used as a long-term attitude reference and export the attitude of the three axis of the satellite;
(3) Put the measurement information of star-sensors and gyros into the nonlinear filter. Then estimate the attitude error, gyro drift error and other error basing on measurement information of star-sensors. Finally, output the attitude information and revise the system.

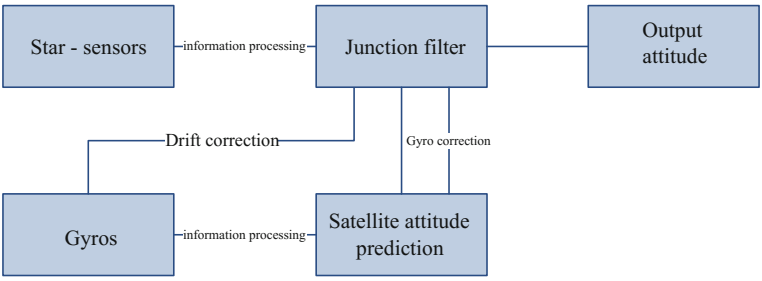


Fig. 1 The principle of satellite Star-sensor/Gyro attitude determination system

2.2 System Measurement Equation

(1) System measurement equation

Taking the typical double-star-sensors system as an example, the measurement equation of the system is showed as [11]:

$$\boldsymbol{Z}_k = \boldsymbol{H}_k \boldsymbol{X}_k + \boldsymbol{V}_k \tag{1}$$

where, $\boldsymbol{H}_k = \begin{bmatrix} -2\boldsymbol{C}_{io} \cdot \boldsymbol{C}_{ob}(\hat{\boldsymbol{q}}) \cdot [\boldsymbol{l}_{b1}\times] & \boldsymbol{0}_{3\times 6} \\ -2\boldsymbol{C}_{io} \cdot \boldsymbol{C}_{ob}(\hat{\boldsymbol{q}}) \cdot [\boldsymbol{l}_{b2}\times] & \boldsymbol{0}_{3\times 6} \end{bmatrix}$, and $\boldsymbol{X}_k = \begin{bmatrix} \Delta\bar{\boldsymbol{q}}^T & \Delta\bar{\boldsymbol{d}}^T & \Delta\bar{\boldsymbol{b}}^T \end{bmatrix}^T$ is the state estimator of the system. $\Delta\bar{\boldsymbol{q}}$, $\Delta\bar{\boldsymbol{d}}$ and $\Delta\bar{\boldsymbol{b}}$ are the vector part of the attitude error of quaternions, the constant drift error of the gyro and the relative drift error of the gyro respectively; $\boldsymbol{V}_k$ is the observation noise sequence, and it satisfies the condition that $E(\boldsymbol{V}_k) = 0$, $\mathrm{E}(\boldsymbol{V}_k \boldsymbol{V}_k^T) = \boldsymbol{R}_k \delta_{kj}$; where, $\boldsymbol{R}_k$ is the observation noise covariance matrix. The measurement accuracy of each star sensor is supposed to be same with each other. By that way, $\boldsymbol{R}_k = vk^2 \boldsymbol{I}_{6\times 6}$, and vk is the measured accuracy. $\boldsymbol{C}_{io}$ is the transformation matrix from orbit coordinate system to inertial coordinate system, which is related to the orbit state of the satellite; $\boldsymbol{C}_{ob}(\hat{\boldsymbol{q}})$ is current attitude matrix. $\hat{\boldsymbol{q}}$ are the current quaternions; $\boldsymbol{l}_{b1}$ and $\boldsymbol{l}_{b2}$ are the installation position of the two star sensor in the satellite system.

(2) The gyro measurement model

Taking fiber optic gyroscope as an example, the measuring model of gyroscope is as below [12]

$$\boldsymbol{\omega}_g = \boldsymbol{\omega} + \boldsymbol{d} + \boldsymbol{b} + \boldsymbol{n}_g \tag{2}$$

Among them, $\boldsymbol{\omega}_g$ it is the actual gyro output; $\boldsymbol{\omega}$ is the true rotation angular velocity of the satellite's three axis, which is relative to the inertial space. $\boldsymbol{d}$ is related to the amount of drift, and satisfy the condition that $\dot{\boldsymbol{d}} = -\boldsymbol{D}_\tau \boldsymbol{d} + \boldsymbol{n}_d$, and $\boldsymbol{D}_\tau = diag(\frac{1}{\tau_x}, \frac{1}{\tau_y}, \frac{1}{\tau_z})$ is a diagonal matrix composed of time correlation constant.

$\boldsymbol{n}_d$ is white noise which is drive by First-order Markov Model. $\boldsymbol{b}$ is the constant drift of gyro, and satisfy the condition that $\dot{\boldsymbol{b}} = \boldsymbol{n}_b$, $\boldsymbol{n}_g$ is the measurement noise, which usually is Gauss white noise.

2.3 System State Equation

According to the different parameters, the state equation is also different. In this paper, the error of the variable part and the gyro drift error of attitude four elements are estimate. The equation of state of the system is showed as below [13]:

$$\dot{\boldsymbol{X}}_k = \boldsymbol{F}_k \boldsymbol{X}_k + \boldsymbol{G}_k \boldsymbol{W}_k \tag{3}$$

where, $\boldsymbol{F}_k = \begin{bmatrix} -[\hat{\boldsymbol{\omega}}\times] & -\frac{1}{2}\boldsymbol{I}_{3\times3} & -\frac{1}{2}\boldsymbol{I}_{3\times3} \\ \boldsymbol{0}_{3\times3} & -\boldsymbol{D}_\tau & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \end{bmatrix}$ is the Jacobian matrix, $[\hat{\boldsymbol{\omega}}\times] = \begin{bmatrix} 0 & -\hat{\omega}_z & \hat{\omega}_y \\ \hat{\omega}_z & 0 & -\hat{\omega}_x \\ -\hat{\omega}_y & \hat{\omega}_x & 0 \end{bmatrix}$, $\boldsymbol{G}_k = \begin{bmatrix} -\frac{1}{2}\boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} \end{bmatrix}$ is the noise driven matrix, $\boldsymbol{W}_k = \begin{bmatrix} \boldsymbol{n}_g \\ \boldsymbol{n}_d \\ \boldsymbol{n}_b \end{bmatrix}$ is the noise matrix.

3 Filter Design and Precision Analysis

3.1 Filter Design Based on CKF

The core of CKF is the Cubature criterion, which is derived from a strict mathematical deduction. By constructing 2n equal-weighted Cubature points and spread them by nonlinear equation, some new points are generated to predict the state of the next moment. Similar to UKF, CKF does not need linearization, and amount of the sampling point is less than UKF, the derivation process is more rigorous. Besides, the accuracy of CKF is higher than that of EKF and UKF. For the satellite attitude measurement system composed of formulas (1) and (3), we discrete the model with fourth-order Runge-Kutta method, then we get a discrete model as below:

$$\begin{cases} \boldsymbol{X}_{k+1} = f_k(\boldsymbol{X}_k) + \boldsymbol{W}_k \\ \boldsymbol{Z}_{k+1} = h_{k+1}(\boldsymbol{X}_{k+1}) + \boldsymbol{V}_{k+1} \end{cases} \tag{4}$$

where, $\boldsymbol{X}_k \in \boldsymbol{R}_n$, $\boldsymbol{Z}_k \in \boldsymbol{R}_m$, and $\boldsymbol{X}_k$ is state estimate, $\boldsymbol{Z}_k$ is updated measurement. $\boldsymbol{W}_k$, $\boldsymbol{V}_k$ are the system noise and measurement noise separately, both of them are independent Gauss white noise, and their covariance matrix are $\boldsymbol{Q}_k$, $\boldsymbol{R}_k$ separately.

Basing on the satellite attitude measurement system expressed by formula (4), we adopt CKF to solve attitude estimation, and the detail is showed as below:

(1) Initialization parameter

$$\begin{cases} \widehat{\boldsymbol{X}}_0 = E\left(\overline{\boldsymbol{X}}_0\right) \\ \boldsymbol{P}_0 = E\left[\left(\boldsymbol{X}_0 - \widehat{\boldsymbol{X}}_0\right)\left(\boldsymbol{X}_0 - \widehat{\boldsymbol{X}}_0\right)^T\right] \end{cases} \tag{5}$$

(2) Decompose the state estimation covariance matrix with Cholesky method

$$\boldsymbol{P}_{k-1} = \boldsymbol{S}_{k-1}\boldsymbol{S}_{k-1}^T \tag{6}$$

(3) Generate Cubature points and spread them with nonlinear equation

$$\begin{cases} \boldsymbol{\chi}_{i,k-1} = \widehat{\boldsymbol{X}}_{k-1} + \boldsymbol{S}_{k-1}\boldsymbol{\xi}_i \\ \boldsymbol{x}_{i,k|k-1} = f\left(\boldsymbol{\chi}_{i,k|k-1}\right) \end{cases}, \quad i = 1, 2, \ldots, 2n \tag{7}$$

where $\boldsymbol{\xi}_i = \sqrt{n}\left[\dot{\boldsymbol{I}}: -\boldsymbol{I}\right]_i$, which represent the ith and jth column value of $\left[\dot{\boldsymbol{I}}: -\boldsymbol{I}\right]$.

(4) Calculated state the one-step predicted value $\boldsymbol{X}_{k|k-1}$ and the predicted covariance matrix $\boldsymbol{P}_{k|k-1}$

$$\begin{cases} \boldsymbol{X}_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n}\boldsymbol{x}_{i,k|k-1} \\ \boldsymbol{P}_{k|k-1} = \frac{1}{2n}\sum_{i=0}^{2n}\left(\boldsymbol{\chi}_{i,k|k-1} - \boldsymbol{X}_{i,k|k-1}\right)\left(\boldsymbol{\chi}_{i,k|k-1} - \boldsymbol{X}_{i,k|k-1}\right)^T + \boldsymbol{Q}_{k-1} \end{cases} \tag{8}$$

(5) Generated Cubature points again and spread them with nonlinear equation

$$\begin{cases} \boldsymbol{P}_{k|k-1} = \boldsymbol{S}_{k-1}\boldsymbol{S}_{k-1}^T \\ \boldsymbol{\chi}_{i,k-1} = \boldsymbol{X}_{k|k-1} + \boldsymbol{S}_{k-1}\boldsymbol{\xi}_i \\ \boldsymbol{z}_{i,k|k-1} = h\left(\boldsymbol{\chi}_{i,k|k-1}\right) \end{cases}, \quad i = 1, 2, \ldots, 2n \tag{9}$$

(6) Calculate the predicted measurement value $\boldsymbol{Z}_{k|k-1}$, the covariance matrix of innovation information $\boldsymbol{P}_k^{zz}$ and the state covariance matrix $\boldsymbol{P}_k^{xz}$

$$\begin{cases} \boldsymbol{Z}_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} \boldsymbol{z}_{i,k|k-1} \\ \boldsymbol{P}_k^{zz} = \frac{1}{2n}\sum_{i=0}^{2n} \left(\boldsymbol{z}_{i,k|k-1} - \boldsymbol{Z}_{k|k-1}\right)\left(\boldsymbol{z}_{i,k|k-1} - \boldsymbol{Z}_{k|k-1}\right)^T + \boldsymbol{R}_{k-1} \\ \boldsymbol{P}_k^{xz} = \frac{1}{2n}\sum_{i=0}^{2n} \left(\boldsymbol{x}_{i,k|k-1} - \boldsymbol{X}_{k|k-1}\right)\left(\boldsymbol{z}_{i,k|k-1} - \boldsymbol{Z}_{k|k-1}\right)^T \end{cases} \tag{10}$$

(7) Status updates

$$\begin{cases} \boldsymbol{K}_k = \boldsymbol{P}_k^{xz}\left(\boldsymbol{P}_k^{zz}\right)^{-1} \\ \widehat{\boldsymbol{X}}_k = \boldsymbol{X}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{Z}_k - \boldsymbol{Z}_{k|k-1}\right) \\ \boldsymbol{P}_k = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}_k^{zz}\boldsymbol{K}_k^T \end{cases} \tag{11}$$

When $\widehat{\boldsymbol{X}}_k$ is obtained, revise the quaternions and gyro drift as below:

$$\begin{cases} \hat{\boldsymbol{q}}_k = \hat{\boldsymbol{q}}_{k|k-1} \otimes \Delta\hat{\boldsymbol{q}}_k \\ \Delta\hat{\boldsymbol{q}}_k = \frac{\Delta\hat{\boldsymbol{q}}_k}{\sqrt{1-\Delta\hat{\tilde{\boldsymbol{q}}}_k\left(\Delta\hat{\tilde{\boldsymbol{q}}}_k\right)^T}} \\ \hat{\boldsymbol{d}}_k = \hat{\boldsymbol{d}}_{k|k-1} + \Delta\hat{\boldsymbol{d}}_k \\ \hat{\boldsymbol{b}}_k = \hat{\boldsymbol{b}}_{k|k-1} + \Delta\hat{\boldsymbol{b}}_k \end{cases} \tag{12}$$

When the update calculation is completed, $\widehat{\boldsymbol{X}}_k$ is reset to zero.

3.2 Precision Analysis

Every filtering algorithm is based on the deduction of Eq. (13), and the precision of different filtering algorithm mainly depend on the integral precision.

$$\boldsymbol{I}(f) = \int_{R^n} f(x)\,\omega(x)\,dx \tag{13}$$

As many literatures have proved the conclusion that precision of UKF is higher than that of EKF, here don't discuss it anymore. In this paper, we mainly focus on the comparison between CKF and UKF.

During the deduction of Eq. (13), both CKF and UKF introduce a series of discrete points with some weight to impend the real integration, and the difference are the selection of discrete points and corresponding weight. CKF adopt Cubature criterion as below:

$$\begin{cases} \xi_i = \sqrt{n}\left[\boldsymbol{I}\vdots - \boldsymbol{I}\right]_i & i = 1, 2\ldots, 2n \\ \omega_i = \frac{1}{2n} \end{cases} \tag{14}$$

UKF adopt UT criterion as below:

$$\begin{cases} \xi_i = 0, \quad \omega_i = \frac{\kappa}{n+\kappa} & i = 1 \\ \xi_i = \sqrt{n+\kappa}[\boldsymbol{I}]_i, \quad \omega_i = \frac{1}{2n} & i = 1, 2\ldots, 2n \end{cases} \tag{15}$$

where, n is dimension of $\boldsymbol{X}$, and variable κ is adjustment factor, usually they satisfy the equation $n + \kappa = 3$. By integrating Eq. (13), we can get the mean value of $f(x)$ as below [14]:

$$\begin{aligned} \bar{f}(x) = & f\left(\widehat{X}_{k-1}\right) + \left(\frac{\nabla^{\mathrm{T}} \boldsymbol{P}_{xx} \nabla}{2!}\right) f + \frac{1 \times 3 \times \cdots \times (2k-1)}{(2k)!} \\ & \times \sum_{i=1}^{n} \left(a_{i1} \frac{\partial}{\partial x_{i1}} + \cdots + a_{in} \frac{\partial}{\partial x_{in}}\right)^{2k} f(x)|_{x=\bar{x}} + \sigma \end{aligned} \tag{16}$$

$$\begin{aligned} \bar{f}(x)_{CKF} = & f\left(\widehat{X}_{k-1}\right) + \left(\frac{\nabla^{\mathrm{T}} \boldsymbol{P}_{xx} \nabla}{2!}\right) f \\ & + \frac{n^{k-1}}{(2k)!} \sum_{i=1}^{n} \left(a_{i1} \frac{\partial}{\partial x_{i1}} + \cdots + a_{in} \frac{\partial}{\partial x_{in}}\right)^{2k} f(x)|_{x=\bar{x}} \end{aligned} \tag{17}$$

$$\begin{aligned} \bar{f}(x)_{UKF} = & f\left(\widehat{X}_{k-1}\right) + \left(\frac{\nabla^{\mathrm{T}} \boldsymbol{P}_{xx} \nabla}{2!}\right) f \\ & + \frac{(n+\kappa)^{k-1}}{(2k)!} \sum_{i=1}^{n} \left(a_{i1} \frac{\partial}{\partial x_{i1}} + \cdots + a_{in} \frac{\partial}{\partial x_{in}}\right)^{2k} f(x)|_{x=\bar{x}} \end{aligned} \tag{18}$$

The variable a_{ij} is the ith and jth column value of the matrix $\boldsymbol{S}_{k-1}$.
By comparing Eq. (16) to (18), we can draw conclusions as below:

(1) If $n < 3$, $\bar{f}(x)_{CKF} < \bar{f}(x)_{UKF} < \bar{f}(x)$, UKF impends the $\bar{f}(x)$ more closely, so the precision of UKF is higher than that of CKF.
(2) If $n = 3$, $\bar{f}(x)_{CKF} = \bar{f}(x)_{UKF} < \bar{f}(x)$, CKF impends the $\bar{f}(x)$ more closely, so the precision of CKF is higher than that of UKF.
(3) If $n > 3$, we compare the performance of the two algorithms on numerical stability. Integral stability factor is defined as below:

$$I = \sum_i |\omega_i| \Big/ \sum_i \omega_i \tag{19}$$

So if $n > 3$, $\omega_0 = 1 - \frac{n}{3} < 0$, $\omega_i = \frac{1}{2n} < 0$ $(i = 1, 2, \ldots, 2n)$, and $I_{UKF} = \frac{2}{3}n - 1 > 1$. According to the conclusion of literature [15], large truncation error will occur and the filtering precision will descend. Considering the fact that $I_{CKF} = 1$, we can easily deduce that $\bar{f}(x)_{UKF} < \bar{f}(x)_{CKF} < \bar{f}(x)$ when $n > 3$, and naturally, the precision of CKF is higher than that of UKF.

4 Simulation and Analysis

Basing on STK software, we chose a low orbit satellite's attitude and three axis angular rate data as the reference data Data1, and the simulation time is 1000 s. Then, according to Table 1, we add the measurement noise and the gyro error to the reference data (The experiment assuming that the performance of the two star sensor is same) to generate experimental data with errors—DATA2.

For convenience of analysis of the performance of the algorithm, we use EKF, UKF and CKF to process Data2 separately. The experimental environment is MATLAB 2014a, and some initial parameters are set as below:

$$\boldsymbol{X}_0 = \boldsymbol{0}_{1\times 9},\ \boldsymbol{P}_0 = \boldsymbol{I}_{9\times 9} \times 10^{-14},\ \boldsymbol{Q}_0 = \boldsymbol{I}_{9\times 9} \times 10^{-14},\ \boldsymbol{R}_0 = \boldsymbol{I}_{6\times 6} \times vk^2$$

The experimental results are shown in Table 2 and Fig. 2. As is showed in Table 2, when the star sensor's measure error is 5arc, each of the three algorithms can achieve a high accuracy, and the accuracy of CKF is highest, followed by UKF,

Table 1 Parameters of attitude sensors

Parameters	Value
The direction of star sensor in satellite system	$\boldsymbol{l}_{b1} = [0 \quad 1 \quad 0]^{\mathrm{T}}$ $\boldsymbol{l}_{b1} = [0 \quad 0 \quad 1]^{\mathrm{T}}$
Star sensor measurement error (3σ)	$vk = 5''$
Star sensor output frequency	5 Hz
Gyro output frequency	50 Hz
Gyro drift	0.1 °/h
Gyro constant drift	0.03 °/h
Gyro measurement error	0.5 °/h

Table 2 The comparison of the attitude accuracy basing on EKF and UKF and CKF

Filtering algorithm	Performance index						
	Error mean (″)			Mean square error (″)			Processing time (s)
	Roll	Pitch	Yaw	Roll	Pitch	Yaw	
EKF	−0.0371	0.1874	0.2668	0.4264	0.9470	0.8336	5.416
UKF	−0.0381	0.1880	0.2434	0.3129	0.7130	0.6329	7.509
CKF	−0.0415	0.1712	0.1910	0.3114	0.6695	0.6024	6.710

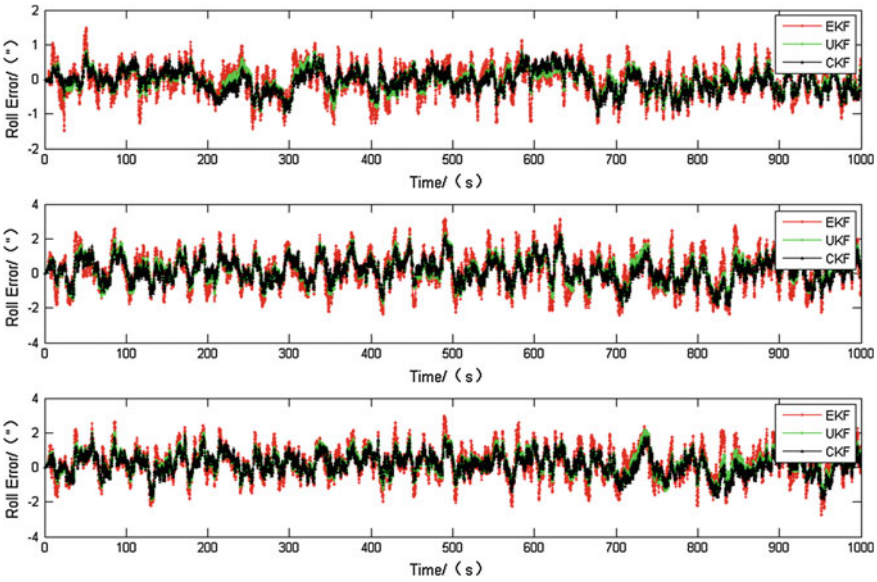


Fig. 2 The principle of satellite Star-sensor/Gyro attitude determination system

EKF is the worse, which is corresponding to the result of Sect. 3.2; And the length of process time: EKF < CKF < UKF; The stability: UKF > CKF > EKF. Such result shows that CKF has a higher precision than UKF with the strict mathematical derivation. Meanwhile, as the amount of the sampling point is less than the UKF, CKF's efficiency is higher than UKF. Last but not least, even though EKF has characteristics of simple calculation, its precision and stability both are the worst among the three algorithms.

5 Conclusions

Aiming at the combination filter of satellite attitude determination system, this paper proposed an idea that adopting the Cubature Kalman filter algorithm to star sensor/gyro integrated attitude determination system, and designed integrated attitude determination filter based on CKF principle. Simulation experiment was also designed to compare the performance of EKF, UKF and the new algorithm, and simulation results showed that the filtering accuracy and stability of CKF both are better than EKF and UKF. The research results have some reference value for engineering time.

References

1. Athans M, Wishner RP, Bertolini A (1968) Suboptimal estimation for continuous time nonlinear system from discrete noisy measurement. IEEE Trans Autom Control 5(13):504–514
2. Julier SJ, Uhlman JK (1997) A new extension of the Kalman filter to nonlinear system. In: Proceedings of the Society of Photo-Optical Instrumentation Engineers 3968(1):182–193
3. Pan Q, Yang F, Ye L, Cheng Y (2005) Survey of a kind of nonlinear filters—UKF. Control Decis 20(5):481–489
4. Arasaratnam I, Hakin S (2009) Cubature Kalman filters. IEEE Trans Autom Control 54 (6):1254–1269
5. Ge QB, Xu DX, Wen CL (2009) Cubature information Kalman filters. IEEE Trans Autom Control 54(6):1254–1269
6. Liu Z, Xie W, Wang P, Yu Y (2013) A Gaussian mixture PHD filter with the capability of information hold. Acta Electron Sin 41(8):1603–1608
7. Duan J, Shi H, Liu D, Yu H () Square root cubature Kalman filter—Kalman filter algorithm for intelligent vehicle position estimate. Proc Eng 137:267–276
8. Cardona-Morales O, Castellanos-Dominguez G (2016) Order tracking by square-root cubature Kalman filter with constraints. Springer International Publishing, Berlin
9. Wang H, Guixia Fu, Bian X, Li J (2013) SRCKF based on simultaneous localization and mapping of mobile robots. Jiqiren/Robot 35(2):200–207
10. Tang X, Liu L, Zhang J (2012) Square-root quaternion cubature Kalman filtering for spacecraft attitude estimation. Acta Astronaut 76(4):84–94
11. Jiao Y (2007) Research on methods of satellite attitude determination based on measurement system with star-sensor/gyroscope. National University of Defense Technology, pp 21–24
12. Mann JA III, Pascal JC (1992) Location noise sources on an industrial air compressor using broadband acoustical holography from intensity measurements. Noise Control Eng J (S0736–2501) 39(1):9–13
13. Yang F, Zhou Z, Liu S (2006) Satellite attitude determination algorithm based on star-sensor and FOG. Control Eng China 13(4):374–376
14. Feng SUN, Li-jun TANG (2013) Estimation precision comparison of cubature Kalman filter and Unscented Kalman filter. Control Dec 28(2):303–308
15. Jia J, Dyer JS (1996) A standard measure of risk and risk-value models. Manag Sci 42 (12):1691–1705

TC-OFDM Receiver Code Tracking Method Based on Extended Kalman Filter

Jun Mo, Zhongliang Deng, Buyun Jia, Xinmei Bian and Jichao Jiao

Abstract Accurate code phase measurement is an important research focus in high sensitivity tracking. Due to the existing time and code division-orthogonal frequency division multiplexing (TC-OFDM) receiver code tracking algorithms based on delay-locked loop (DLL) and the extended Kalman filter (EKF) without amplitude estimation are prone to loosing the loop in dynamic environments, a TC-OFDM receiver code tracking method is proposed. To improve the estimation accuracy of EKF in the environment with drastic change of signal amplitude, the algorithm estimates the phase error of PN code and signal amplitude as the state vector of EKF. Theoretical analysis and simulation results show that this algorithm takes advantage of the statistical properties of signal amplitude, overcomes the shortcomings of the traditional DLL and EKF without amplitude estimation, and effectively improves the convergence speed and tracking quality of the code loop. Thereby, the proposed algorithm can improve the tracking sensitivity of the existing TC-OFDM receiver by 2–5 dB.

Keywords Code tracking · EKF · Signal amplitude · TC-OFDM

1 Introduction

In the spread spectrum receiver, signal tracking includes two parts: carrier tracking and code tracking. High-sensitivity research is an important direction of its development. By improving the convergence speed of code loop when the signal strength greatly changes, the spread spectrum receiver can reduce the unlocked probability of the code loop and improve the tracking sensitivity.

In the Signal Noise Ratio (SNR) dynamic environment, the fluctuation of the signal amplitude may cause the code loop to be unlocked. The traditional code

J. Mo (✉) · Z. Deng · B. Jia · X. Bian · J. Jiao
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) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_26

tracking algorithm is based on the delay-locked loop (DLL) technique [1, 2], the characteristic frequency and damping coefficient to determine the performance of code loop are generally fixed or simply adjusted. It cannot be stably tracking the positioning signal whose amplitude changes over a wide range. Ref. [3] uses extended Kalman filter (EKF) to eliminate the effect of multipath effect on signal tracking. The tracking model is simple and the loop is easy to be unlocked at low SNR. In Ref. [4], a method of combining improved code phase detector with EKF suppression loop noise is proposed to achieve weak signal tracking under multipath interference. In Ref. [5], the EKF algorithm is combined with Bayesian estimation, and can track the carrier and pseudo-code phase of the GPS weak signal. However, neither of the two methods analysed the SNR dynamic change. Refs. [6] and [7] propose a signal tracking algorithm based on adaptive Kalman filter. The noise covariance matrix is estimated by historical information and the system model is tracked. Because the dynamic covariance matrix has uncertainty, low SNR can easily cause the problem of filter divergence.

The time and code division-orthogonal frequency division multiplexing (TC-OFDM) positioning system based on terrestrial mobile broadcast network adopts CDMA. In the indoor environments with frequent occlusion and intense signal strength, the dynamic range of signal amplitude is often large, and the signal strength decreases 30 dB in a short time. Therefore, the TC-OFDM receiver also has the above-mentioned tracking problem. In this paper, an extended Kalman code tracking algorithm including signal amplitude estimation for SNR dynamic change is proposed. The code phase error and the signal amplitude are jointly estimated as EKF state vectors to improve the tracking loop convergence rate, then the tracking sensitivity of the TC-OFDM receiver can be improved.

2 Code Tracking Model

TC-OFDM signal is spread-modulated with the pseudo code firstly, then the pseudo code and data code through the two-phase shift keying mechanism (BPSK) to modulate the carrier. Received by the receiver antenna to the ith base station signal, the digital IF signal converted by the analog-digital conversion (ADC) module can be written as:

$$r^{(i)}(nT_s) = A_i s^{(i)}(nT_s - \zeta_i) e^{j2\pi(f_{IF} + f_{d,i})nT_s + \varphi_{0,i}} + \omega(n) \tag{1}$$

where T_s denotes the sampling clock, A_i is the amplitude of the IF signal, $s^{(i)}$ is the spreading code for modulating the data information, ζ_i is the signal propagation delay, f_{IF} denotes the signal intermediate frequency, $f_{d,i}$ is the residual carrier frequency, $\varphi_{0,i}$ is the initial phase of the carrier and $\omega(n)$ is the additive white Gaussian noise (AWGN).

In the signal tracking stage, the signal channel obtained the current base station signal carrier frequency and code phase through the tracking loop. Since the bit

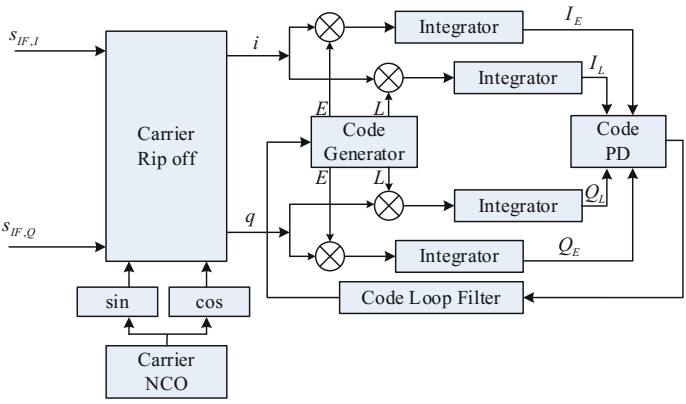


Fig. 1 Traditional code tracking loop structure

synchronization has been completed by the TC-OFDM positioning receiver in the tracking loop, the modulated data bit boundary is known. Subsequent analysis of this paper is based on the bit synchronization has been completed, then the carrier loop has obtained detailed carrier frequency. Traditional code tracking loop using DLL technology shown in Fig. 1.

The delay between the received signal and the local signal is estimated by the code loop. In traditional code tracking loops, the common code loop discriminator is:

$$\tau = (1-d)\frac{E-L}{E+L} \tag{2}$$

where d is the interval between the Early and the Immediate two sets of correlators, E and L are the integral value of the Early and the Later.

$$\begin{cases} E = \sqrt{I_E(\tau)^2 + Q_E(\tau)^2} = A|\sin\, c(f_d T_{coh})|R(\tau - d) + n_E \\ L = \sqrt{I_L(\tau)^2 + Q_L(\tau)^2} = A|\sin\, c(f_d T_{coh})|R(\tau + d) + n_L \end{cases} \tag{3}$$

where A denotes the signal amplitude of the output correlator, T_{coh} is the coherent integration time, $R(\cdot)$ is the autocorrelation function, n_E and n_L are the noise of the output of the corresponding correlator branch. Note $A_c = (A|\sin c(f_d T_{coh})|)^2$, then Eq. (3) can be expressed as:

$$\begin{cases} E = \sqrt{A_c}R(\tau - d) + n_E \\ L = \sqrt{A_c}R(\tau + d) + n_L \end{cases} \tag{4}$$

When calculates τ by Eq. (2), A_c is normalized and eliminated, τ is not sensitive to A_c changes. But in the signal tracking, A_c can reflect the current received signal

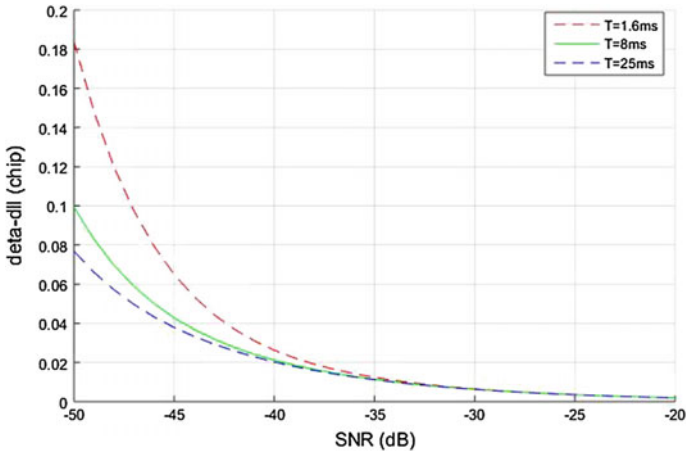


Fig. 2 Code phase error under different SNR

power and carrier tracking. In the case of accurate carrier tracking, SNR of the received signal can indirectly characterized by A_c, and the SNR directly affects the error accuracy of the tracking loop, as shown in Fig. 2.

When the SNR is low, the base station signal received by the TC-OFDM receiver is easily affected by internal thermal noise and dynamic stress error, causes comparatively large signal parameter estimation error in the TC-OFDM receiver tracking loop. There are nonlinear factors when the discriminator works, so the receiver is prone to appear the code loop unlocked phenomenon. Because the traditional code tracking loop algorithm is difficult to deal with the dramatic changes of the base station signal, τ and A_c are used together as the state vector of EKF in this paper, and EKF method is used to replace traditional code tracking discriminator.

3 Improved Extended Kalman Code Tracking Algorithm

3.1 Improved Code Loop Model

Figure 3 is the proposed extended Kalman code tracking algorithm in this paper. The core idea is to replace the code loop discriminator with EKF in the TC-OFDM receiver tracking loop. The EKF algorithm is an estimation method based on the minimum mean square error estimation criterion. The nonlinear function is expanded by Taylor polynomials near the state estimation value. So the nonlinear filtering problem is transformed to a linear filtering problem. In practice applications, the form of EKF is closely related to the state equation and the observation equation.

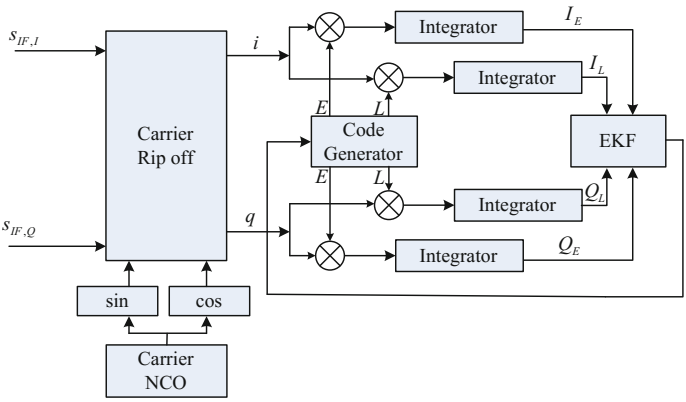


Fig. 3 EKF code tracking loop structure

3.2 State Equation

State equation describes the state vector. In order to accurately track the phase of the input signal, the state of the filter is select as:

$$\mathbf{X} = [\tau, A_c]^T \tag{5}$$

Assuming that signal energy A_c changes slowly over a short period of time, it can be modeled as a stationary first-order Markov model unit matrix with one-step state transition matrix:

$$A_{c,k} = A_{c,k-1} + n_{A_c} \tag{6}$$

where n_{A_c} is the change value of the signal energy at two moments. Then the EKF linear state equation can be expressed as:

$$\mathbf{X}_k = \mathbf{\Phi}\mathbf{X}_{k-1} + \mathbf{W}_{k-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{X}_{k-1} + \begin{bmatrix} n_\tau \\ n_{A_c} \end{bmatrix} \tag{7}$$

where $\mathbf{\Phi}$ is the state transition matrix, $\mathbf{W}_{k-1}$ is the system process error. n_τ, n_{A_c} are the Gaussian white noise with mean 0 and variance of δ_τ^2 and $\delta_{A_c}^2$. Therefore, the covariance matrix is:

$$\mathbf{Q} = \begin{bmatrix} \delta_\tau^2 & 0 \\ 0 & \delta_{A_c}^2 \end{bmatrix} \tag{8}$$

3.3 Observation Equation

The observation equation describes the relationship between the observed value and the state vector. The integral clear E and L branch outputs in the code tracking loop are given by Eq. (3). For the convenience of analysis, consider only a non-coherent integration code loop tracking model, observation values can be abbreviated as follows vector form:

$$\mathbf{Y} = \begin{bmatrix} E^2 \\ L^2 \end{bmatrix} = \begin{bmatrix} A_c^2 R(\tau - d)^2 \\ A_c^2 R(\tau + d)^2 \end{bmatrix} + \begin{bmatrix} n_E^2 \\ n_L^2 \end{bmatrix} \tag{9}$$

Equation (9) also can be written as

$$\mathbf{Y}_k = h(\mathbf{X}_k) + \mathbf{V}_k \tag{10}$$

where $\mathbf{V}_k$ is the random observation noise, the observation matrix corresponds to $\mathbf{Y}_k$ is:

$$\mathbf{H}_k = \left.\frac{\partial h(\mathbf{X}_k)}{\partial \mathbf{X}_k}\right|_{\mathbf{X}_k = \hat{\mathbf{X}}_{k,k-1}} = \begin{bmatrix} A_{c,k-1} \frac{\partial (R(\tau - d)^2)}{\partial \tau} & R(\tau_{k-1} - d)^2 \\ A_{c,k-1} \frac{\partial (R(\tau + d)^2)}{\partial \tau} & R(\tau_{k-1} + d)^2 \end{bmatrix} \tag{11}$$

Due to the complexity of PN code auto-correlation function with limited bandwidth, PN code autocorrelation function of infinite bandwidth is used instead:

$$R(\tau) = \begin{cases} 1 - |\tau|/T_{coh} & |\tau| < T_{coh} \\ 0 & others \end{cases} \tag{12}$$

by taking $R(\tau)$ into Eq. (11):

$$\mathbf{H}_k = \begin{bmatrix} 2A_{c,k-1}\left(\frac{1}{T_{coh}} - \frac{\tau_{k-1} - d}{T_{coh}^2}\right) & 1 - 2\frac{|\tau_{k-1} - d|}{T_{coh}} + \frac{(\tau - d)^2}{T_{coh}^2} \\ 2A_{c,k-1}\left(\frac{1}{T_{coh}} - \frac{\tau_{k-1} + d}{T_{coh}^2}\right) & 1 - 2\frac{|\tau_{k-1} + d|}{T_{coh}} + \frac{(\tau + d)^2}{T_{coh}^2} \end{bmatrix} \tag{13}$$

The I/Q output noise of the correlator follows a Gaussian distribution with mean 0 and variance δ_N^2, then:

$$E(n_{E,I}^2) = E(n_{E,Q}^2) = E(n_{L,I}^2) = E(n_{L,Q}^2) = \delta_N^2 \tag{14}$$

At the same time, the noise is independent in I/Q two ways of the same correlator, so

$$\mathrm{cov}(n_{E,I}, n_{E,Q}) = \mathrm{cov}(n_{L,I}, n_{L,Q}) = 0 \tag{15}$$

$$\mathrm{cov}(n_{E,I}^2, n_{E,Q}^2) = \mathrm{cov}(n_{L,I}^2, n_{L,Q}^2) = 0 \tag{16}$$

The corresponding by (3):

$$n_E^2 = n_{E,I}^2 + n_{E,Q}^2 + 2\sqrt{A_c}R(\tau - d)\,cos(\varphi)n_{E,I} + 2\sqrt{A_c}R(\tau - d)\ \sin(\varphi)n_{E,Q} \tag{17}$$

$$n_L^2 = n_{L,I}^2 + n_{L,Q}^2 + 2\sqrt{A_c}R(\tau + d)cos(\varphi)n_{L,I} + 2\sqrt{A_c}R(\tau + d)\,\sin(\varphi)n_{L,Q} \tag{18}$$

The variance of n_L^2 and n_E^2 is:

$$Var(n_E^2) = E(n_E^4) - E(n_E^2)^2 = 4\delta_N^4 + 4A_cR(\tau - d)^2\delta_N^2 \tag{19}$$

$$Var(n_L^2) = E(n_L^4) - E(n_L^2)^2 = 4\delta_N^4 + 4A_cR(\tau + d)^2\delta_N^2 \tag{20}$$

and

$$\mathrm{cov}(n_E^2, n_L^2) = 0 \tag{21}$$

where the covariance matrix of $\mathbf{V}_k$ is

$$\mathbf{R} = \begin{bmatrix} 4\delta_N^4 + 4A_cR(\tau - d)^2\delta_N^2 & 0 \\ 0 & 4\delta_N^4 + 4A_cR(\tau + d)^2\delta_N^2 \end{bmatrix} \tag{22}$$

3.4 EKF Algorithm

The EKF algorithm obtains the filtering value of the current time according to the estimated value of the state at the previous time and the observation value of the current time. Therefore, the filtering process of the entire code loop can be divided into two parts: state estimation and state prediction.

(1) Prediction
state one-step prediction:

$$\mathbf{X}_{k,k-1} = \mathbf{\Phi}\,\mathbf{X}_{k-1} \tag{23}$$

one step prediction error matrix:

$$\mathbf{P}_{k,k-1} = \mathbf{\Phi}\,\mathbf{P}_{k-1}\mathbf{\Phi}^T + \mathbf{Q}_{k-1} \tag{24}$$

(2) Update
Filter gain:

$$\mathbf{K}_k = \mathbf{P}_{k,k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k,k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \tag{25}$$

State estimation:

$$\mathbf{X}_k = \mathbf{X}_{k,k-1} + \mathbf{K}_k(\mathbf{Y}_k - h(\mathbf{X}_{k,k-1})) \tag{26}$$

Estimated error square:

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k,k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \tag{27}$$

From Eq. (25), when $\mathbf{R}_k$ is large, the corresponding $\mathbf{K}_k$ will be very small, resulting in the state estimation value calculated by Eq. (26) is very small; when q is small, the one-step prediction error equation $\mathbf{P}_{k,k-1}$ calculated by Eq. (24) will be small, resulting in a smaller state estimate $\mathbf{X}_k$. From the above analysis, the updating of EKF system state is a compromise between the current system state uncertainty and the observed uncertainty. In this paper, $\mathbf{R}_k$ and $\mathbf{Q}_k$ are determined by comparing the historical observations with the current observations to enhance the adaptability of the EKF filter to noise.

4 Simulation and Analysis

In order to verify the performance and correctness of the improved extended Kalman code tracking algorithm, the real TC-OFDM signal is collected by the signal collector, and the noise is added artificially. To test the performance of the code loop, the carrier tracking method proposed in Ref. [4] is used to accurately track the carrier without noise. Simulation mainly tests the tracking quality of the code loop under different SNR and the loop convergence speed under the signal strength fluctuation. The simulation results are compared with the traditional second-order DLL with 25 ms coherent integration time, the EKF without signal amplitude estimation and the improved EKF tracking algorithm with amplitude estimation proposed in this paper.

Under different SNR, the code rate fluctuation of the three code tracking algorithms is shown in Fig. 4. Figure 4a is the code loop stable tracking situation when SNR is −25 dB, and Fig. 4b is when SNR is −35 dB. It can be seen from Fig. 4a that the traditional second-order DLL can track the TC-OFDM signal better in the case of strong signal, the code rate fluctuation is within 0.4 Hz; but the proposed EKF code tracking algorithm with signal amplitude estimation has better signal tracking performance, smaller code rate fluctuation range and more accurate tracking accuracy. From Fig. 4b, it can be seen that the tracking performance of

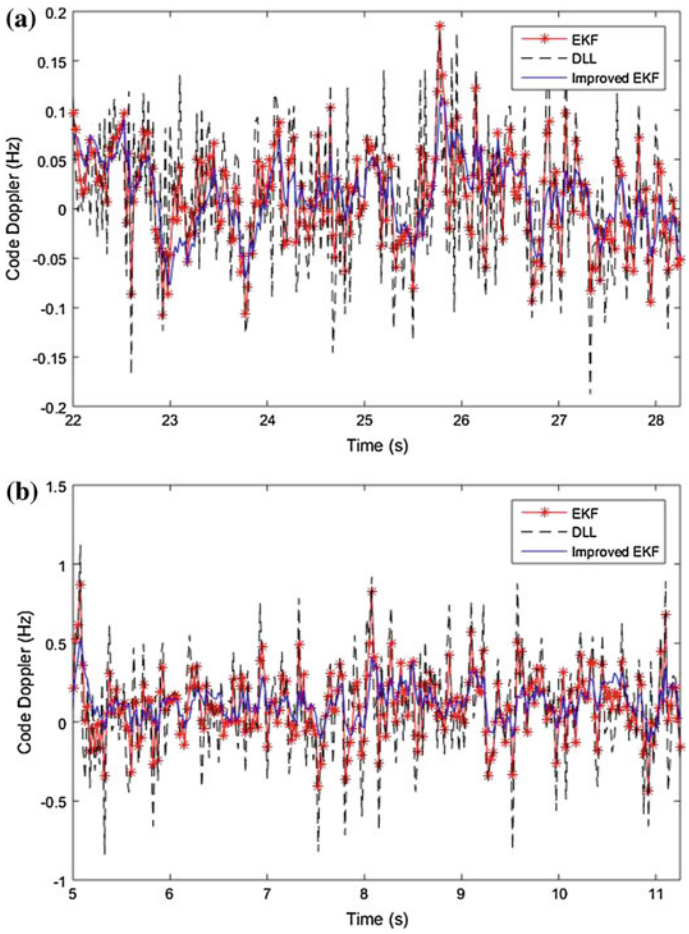


Fig. 4 Comparison of code rate fluctuation under different SNR

traditional DLL is degraded, and the fluctuation of steady-state code rate is extended to 2 Hz. The range of code rate fluctuation of the proposed algorithm is increased, but it is smaller than that of traditional second-order DLL algorithm and EKF algorithm without signal amplitude estimation. So the tracking precision of the proposed is the best among the three algorithms.

To test the code loop convergence rate, a time to increase the original data on the 10 dB attenuation and maintain 2 s, and code rate fluctuations of three tracking algorithms are shown in Fig. 5. It can be seen that the code rate fluctuations of traditional DLL and EKF algorithm are fluctuating greatly, but EKF with amplitude estimation are small and convergent rapidly. At the same time, the EKF algorithm with amplitude estimation has small fluctuation of code rate in the region of signal attenuation.

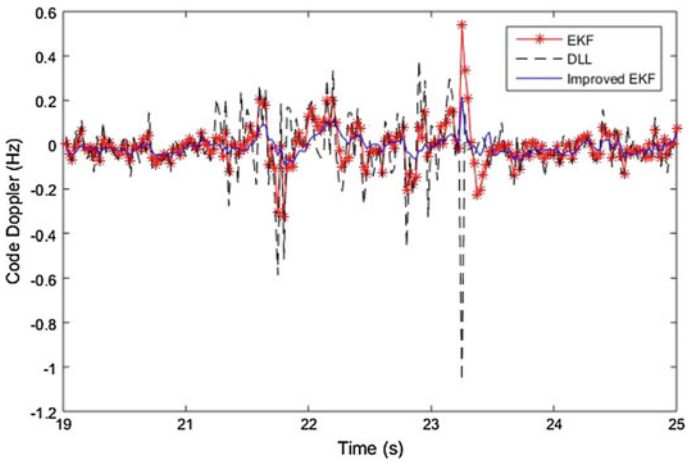


Fig. 5 Comparison of code rate fluctuation when the SNR is abrupt

5 Conclusion

Based on the analysis of traditional DLL and EKF code tracking algorithm without amplitude estimation, an EKF code loop model including signal amplitude estimation is proposed and simulated. Compared with these three algorithms, the EKF code tracking algorithm with signal amplitude estimation is more efficient than traditional DLL and EKF code tracking algorithm. The proposed algorithm has the advantages of small fluctuation range of code rate and fast convergence rate when the signal strength is large and dynamic changing, and tracks the TC-OFDM signal better in the low SNR environment. At the same time, this algorithm is implemented on FPGA, and can improve the tracking ability of the existing TC-OFDM receiver by 2–5 dB.

Acknowledgements Foundation project: The National Key Project of Research and Development Plan of China (2016YFB0502001).

References

1. Tsui JBY (2005) Fundamentals of global positioning system receivers: a software approach, 2nd edn. Wiley, pp 203–225
2. Xie G (2009) Principles of GPS and receiver design. Publishing House of Electronics Industry, Beijing, pp 304–323
3. Jee G, Kim HS, Lee YJA (2002) GPS C/A code tracking loop based on extended Kalman filter with multipath mitigation. In: Proceedings of the Institute of Navigation GPS, Portland, USA, pp 446–451
4. Wei J, Fan SL, Zhou C, Liu JY (2014) Research on the tracking algorithm of GPS weak signal under multipath interference. J Yunnan Natl Univ (Natural Sciences Edition) 2014.05:373–378

5. Psiaki ML, Jung H (2002) Extended Kalman filter methods for tracking weak GPS signals. In: Proceedings of the Institute of Navigation GPS 2002, Portland, 2002, pp 2539–2553
6. Shi E (2012) An improved real-time adaptive Kalman filter for low-cost integrated GPS/INS navigation. In: Proceedings of the 2012 international conference on measurement, information and control (MIC), Harbin, China, pp 1093–1098
7. 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
8. Feng S, Wei-dong L, Qiang L (2015) GPS receiver signal tracking method based on I/Q branch coherent integration measurements filter. J Electron Inf Technol 2015.01: 37–42.

A RTK Float Ambiguity Estimation Method Based on GNSS/IMU Integration

Di Zhu, Zhong-liang Deng, Kai-qin Lin and Jun Lu

Abstract Recently, unmanned aerial vehicles (UAVs) have been widely used in personal entertainment, aerial photography, geographical mapping, and precision agriculture. Accurate position of the UAV is necessary for intelligent and autonomous UAV. RTK technology can provide centimeter level positioning accuracy making it one of the development trends of UAV precise navigation. However, due to the high dynamic of vibration and motion of UAV, the accuracy of float ambiguity estimation for a low cost single frequency RTK-GNSS receiver is limited. In this paper, a single frequency RTK-GNSS/IMU integration method is proposed to improve the performance of Real time kinematic (RTK). Testified by a low cost single frequency GNSS receiver, the proposed method can decrease the time to fix ambiguity by more than 14% and increase 20% of fix rate.

Keywords UAV · Precise navigation · RTK-GNSS/IMU integration

1 Introduction

Recently, unmanned aerial vehicles (UAVs) have been widely used for a variety of applications. According to KPCB, the largest venture capital fund in the United States, global consumer unmanned aerial vehicle shipments reached 4.3 million units in 2015, which are widely used in personal entertainment, aerial photography, geography and precision agriculture. With the increasing demand for intelligent and autonomous UAVs, the traditional standard single point positioning (SPP) accuracy has been difficult to provide sufficient position accuracy.

Real Time Kinematic (RTK) can provide real-time centimeter level position accuracy [1]. Current consumer-grade UAVs typically use low-cost single-frequency RTK-GNSS receiver. However, the vibration and motion of the

D. Zhu (✉) · Z. Deng · K. Lin · J. Lu
School of Eletronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
e-mail: zhudi_bupt@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_27

UAV is high dynamic, much longer time ambiguity converge time is therefore needed, decreasing the ambiguity fix rate. To resolve integer ambiguity, first, we need to estimate the float ambiguity, and then use the Least-Squares Ambiguity Decorrelation Algorithm (LAMBDA) to search the integer ambiguity [5]. The accuracy of the float ambiguity determines whether the ambiguity search is successful, which affects the ambiguity convergence time and fix rate. At present, float ambiguity is typically estimated by the Kalman filter with a constant acceleration model [3]. However, the model does not take the high dynamic characteristics of the UAV into consideration, which enlarges the estimation error, leading to a longer converge time.

To improve the accuracy of float ambiguity, we introduce an improved Kalman filter method based on RTK-GNSS/IMU Integration. By using the attitude data provided by the UAV controller's MEMS acceleration sensor, we could improve the positioning accuracy and shorten the ambiguity converge time.

2 Float Solution Estimation Based on Kalman Filter

The carrier-phase, pseudo-range and Doppler measurements can be expressed as:

$$\begin{cases} P_u^{s_i} = \rho_u^{s_i} + c(\delta t_u - \delta t^s) + I_u^{s_i} + T_u^{s_i} \\ \phi_u^{s_i} = \lambda^{-1}(\rho_u^{s_i} + c(\delta t_u - \delta t^s) - I_u^{s_i} + T_u^{s_i}) + N_u^{s_i} \\ D_u^{s_i} = -\lambda^{-1}(\dot{\rho}_u^{s_i} + c(\delta \dot{t}_u - \delta \dot{t}^s) + \dot{I}_u^{s_i} + \dot{T}_u^{s_i}) \end{cases} \tag{1}$$

where $P_u^{s_i}$ is pseudo-range measurement; $\phi_u^{s_i}$ is carrier-phase measurement and $D_u^{s_i}$ is Doppler measurement.

RTK is based on the double-differencing carrier phase measurement. The satellite clock biases δt^s and receiver clock biases δt_u can be completely eliminated by using double-differencing technique. Under short baseline, most of the ionospheric errors $I_u^{s_i}$ and tropospheric errors $T_u^{s_i}$ can also be eliminated. The double-difference measurement equation at epoch k is expressed as:

$$\begin{cases} P_{k,ub}^{s_2 s_1} = \rho_{k,ub}^{s_2 s_1} \\ \phi_{k,ub}^{s_2 s_1} = \lambda^{-1}\rho_{k,ub}^{s_2 s_1} + N_{k,ub}^{s_2 s_1} \\ D_{k,ub}^{s_2 s_1} = -\lambda^{-1}\dot{\rho}_{k,ub}^{s_2 s_1} \\ \vdots \\ P_{k,ub}^{s_n s_1} = \rho_{k,ub}^{s_n s_1} \\ \phi_{k,ub}^{s_n s_1} = \lambda^{-1}\rho_{k,ub}^{s_n s_1} + N_{k,ub}^{s_n s_1} \\ D_{k,ub}^{s_n s_1} = -\lambda^{-1}\dot{\rho}_{k,ub}^{s_n s_1} \end{cases} \tag{2}$$

where $P_{k,rb}^{s_2 s_1}$ is double-differencing pseudo-range measurement; $\phi_{k,rb}^{s_2 s_1}$ is double-differencing carrier-phase measurement and $D_{k,rb}^{s_2 s_1}$ is double-differencing

Doppler measurement. The components of Kalman filter state vector are position, velocity, acceleration and double-differencing integer ambiguity which is defined as:

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{r} & \boldsymbol{v} & \boldsymbol{a} & N_{ub}^{s_2 s_1} & \cdots & N_{ub}^{s_n s_1} \end{bmatrix}^T \tag{3}$$

The partial derivatives matrix of Eq. (2) can be written as:

$$\boldsymbol{H}_k = \begin{bmatrix} \frac{\partial \rho_u^{s_2 s_1}}{\partial \boldsymbol{r}} & \boldsymbol{0}_{1\times 3} & \boldsymbol{0}_{1\times 3} & 0 & \cdots & 0 \\ \lambda^{-1}\frac{\partial \rho_u^{s_2 s_1}}{\partial \boldsymbol{r}} & \boldsymbol{0}_{1\times 3} & \boldsymbol{0}_{1\times 3} & 1 & \cdots & 0 \\ -\lambda^{-1}\frac{\partial \dot{\rho}_u^{s_2 s_1}}{\partial \boldsymbol{r}} & -\lambda^{-1}\frac{\partial \dot{\rho}_u^{s_2 s_1}}{\partial \boldsymbol{v}} & \boldsymbol{0}_{1\times 3} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \rho_u^{s_n s_1}}{\partial \boldsymbol{r}} & \boldsymbol{0}_{1\times 3} & \boldsymbol{0}_{1\times 3} & 0 & \cdots & 0 \\ \lambda^{-1}\frac{\partial \rho_u^{s_n s_1}}{\partial \boldsymbol{r}} & \boldsymbol{0}_{1\times 3} & \boldsymbol{0}_{1\times 3} & 0 & \cdots & 1 \\ -\lambda^{-1}\frac{\partial \dot{\rho}_u^{s_n s_1}}{\partial \boldsymbol{r}} & -\lambda^{-1}\frac{\partial \dot{\rho}_u^{s_n s_1}}{\partial \boldsymbol{v}} & \boldsymbol{0}_{1\times 3} & 0 & \cdots & 0 \end{bmatrix} \tag{4}$$

where $\frac{\partial \rho_u^{s_i}}{\partial \boldsymbol{r}} = \frac{\partial \dot{\rho}_u^{s_i}}{\partial \boldsymbol{v}} = -\frac{\boldsymbol{r}^{s_i} - \boldsymbol{r}}{\|\boldsymbol{r}^{s_i} - \boldsymbol{r}\|}$, $\frac{\partial \dot{\rho}_u^{s_i}}{\partial \boldsymbol{r}} = \left(\frac{\boldsymbol{v}^{s_i} - \boldsymbol{v}}{\|\boldsymbol{r}^{s_i} - \boldsymbol{r}\|} \cdot \frac{\boldsymbol{r}^{s_i} - \boldsymbol{r}}{\|\boldsymbol{r}^{s_i} - \boldsymbol{r}\|}\right)\frac{\boldsymbol{r}^{s_i} - \boldsymbol{r}}{\|\boldsymbol{r}^{s_i} - \boldsymbol{r}\|} - \frac{\boldsymbol{v}^{s_i} - \boldsymbol{v}}{\|\boldsymbol{r}^{s_i} - \boldsymbol{r}\|}$.

(1) **Kalman filter update phase**

By using Eq. (1), the measurement vector can be written as:

$$\boldsymbol{y} = \begin{bmatrix} P_{k,ub}^{s_2 s_1} & \phi_{k,ub}^{s_2 s_1} & D_{k,ub}^{s_2 s_1} & \cdots & P_{k,ub}^{s_n s_1} & \phi_{k,ub}^{s_n s_1} & D_{k,ub}^{s_n s_1} \end{bmatrix}^T \tag{5}$$

The Kalman filter observation equation at epoch k is defined as:

$$\boldsymbol{y}_k = \boldsymbol{H}_k \tilde{\boldsymbol{x}}_k + \boldsymbol{v}_k \tag{6}$$

Where $\boldsymbol{v}_k$ is the observation noise vector which is assumed to be zero mean Gaussian white noise with the normal distribution $\boldsymbol{v}_k \sim N(0, \boldsymbol{R}_k)$. The Kalman filter update phase uses the following equation:

$$\begin{aligned} \boldsymbol{K}_k &= \tilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T (\boldsymbol{H}_k \tilde{\boldsymbol{P}}_k \boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1} \\ \hat{\boldsymbol{x}}_k &= \tilde{\boldsymbol{x}}_k + \boldsymbol{K}_k (\boldsymbol{y}_k - \boldsymbol{H}_k \tilde{\boldsymbol{x}}_k) \\ \hat{\boldsymbol{P}}_k &= (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k) \tilde{\boldsymbol{P}}_k \end{aligned} \tag{7}$$

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 $\hat{\boldsymbol{x}}_k$.

(2) **Kalman filter predict phase**

The Kalman filter predict phase between the two epochs can be expressed as:

$$\tilde{\boldsymbol{x}}_k = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{w}_k \tag{8}$$

$$\tilde{\boldsymbol{P}}_k = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{P}}_{k-1}\boldsymbol{F}_{k,k-1}^T + \boldsymbol{Q}_k \tag{9}$$

where $\boldsymbol{F}_{k,k-1} = \begin{bmatrix} \boldsymbol{I}_{3\times3} & \Delta t \cdot \boldsymbol{I}_{3\times3} & \Delta t^2/2 \cdot \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times n} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \Delta t \cdot \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times n} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times n} \\ \boldsymbol{0}_{n\times3} & \boldsymbol{0}_{n\times3} & \boldsymbol{0}_{n\times3} & \boldsymbol{I}_{n\times n} \end{bmatrix}$ is state transition matrix from epoch $k-1$ to epoch k. $\boldsymbol{w}_k$ is the process noise vector which is assumed to be zero mean Gaussian white noise with the normal distribution $\boldsymbol{w}_k \sim N(0, \boldsymbol{Q}_k)$ and the $\tilde{\boldsymbol{x}}_k$ is the predicted state vector at epoch k. Assuming that the carrier is at constant acceleration between the two epochs and that the acceleration is a white noise process with constant spectral density, the covariance matrix of the process noise can be expressed as:

$$\boldsymbol{Q}_k = \begin{bmatrix} \frac{\Delta t^4}{20}\boldsymbol{I}_{3\times3} & \frac{\Delta t^3}{8}\boldsymbol{I}_{3\times3} & \frac{\Delta t^2}{6}\boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times n} \\ \frac{\Delta t^3}{8}\boldsymbol{I}_{3\times3} & \frac{\Delta t^2}{3}\boldsymbol{I}_{3\times3} & \frac{\Delta t}{2}\boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times n} \\ \frac{\Delta t^2}{6}\boldsymbol{I}_{3\times3} & \frac{\Delta t}{2}\boldsymbol{I}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times n} \\ \boldsymbol{0}_{n\times3} & \boldsymbol{0}_{n\times3} & \boldsymbol{0}_{n\times3} & \boldsymbol{0}_{n\times n} \end{bmatrix} \sigma_a^2 \tag{10}$$

where σ_a^2 is the acceleration constant variance, determined by the actual movement model of the carrier. However, to determine the size of σ_a^2 is difficult. A large σ_a^2 will cause poor filter precision, and if σ_a^2 is too small, the filter position will not reflect the actual movement of the carrier in time. Since the vibration and motion of the UAV is high dynamic and the observation rate output by low cost RTK-GNSSs receiver is low, the constant acceleration assumption cannot be assured. Thus, the accuracy of float ambiguity based on the constant acceleration model is not suitable for UAV application.

3 Float Solution Estimation Based RTK-GNSS/IMU Integration

Current UAVs typically use IMU to estimate the attitude of the UAV. We apply the IMU gyroscopes and accelerometers measurement to Kalman filter predict phase to improve the accuracy of float ambiguity. The Kalman filter state vector can be rewritten as:

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{r} & \boldsymbol{v} & \boldsymbol{b}^b & N_{ub}^{s_2 s_1} & \ldots & N_{ub}^{s_n s_1} \end{bmatrix}^T \tag{11}$$

where $\boldsymbol{b}^b$ is the accelerometers bias estimate expressed in the body-fixed-frame. The Kalman filter input vector $\boldsymbol{u}_k$ at epoch k can be added to the state transition equation (Eq. 8)

$$\boldsymbol{u}_k = \boldsymbol{C}_{k,b}^e \cdot \boldsymbol{f}^b + \boldsymbol{g}^e + \boldsymbol{m}_k \tag{12}$$

where $\boldsymbol{f}^b$ is accelerometers measurement output by IMU. $\boldsymbol{C}_{k,b}^e$ is the rotation matrix from body-fixed-frame to ECEF frame. $\boldsymbol{g}^e$ is gravitational acceleration vector expressed in ECEF frame and $\boldsymbol{m}_k$ is accelerometers input noise. Using the input vector $\boldsymbol{u}_k$, (Eq. 8) can be rewritten as:

$$\boldsymbol{x}_k = \boldsymbol{F}_{k,k-1}\boldsymbol{x}_{k-1} + \boldsymbol{B}_{k,k-1}\boldsymbol{u}_k + \boldsymbol{w}_k \tag{13}$$

where $\boldsymbol{F}_{k,k-1} = \begin{bmatrix} \boldsymbol{I}_{3\times3} & \Delta t \cdot \boldsymbol{I}_{3\times3} & \Delta t^2/2 \cdot \boldsymbol{C}_{k,b}^e & \boldsymbol{0}_{3\times n} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \Delta t \cdot \boldsymbol{C}_{k,b}^e & \boldsymbol{0}_{3\times n} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times n} \\ \boldsymbol{0}_{n\times3} & \boldsymbol{0}_{n\times3} & \boldsymbol{0}_{n\times3} & \boldsymbol{I}_{n\times n} \end{bmatrix}$ is the state transition matrix and $\boldsymbol{B}_{k,k-1} = \begin{bmatrix} \Delta t^2/2 \cdot \boldsymbol{I}_{3\times3} \\ \Delta t \cdot \boldsymbol{I}_{3\times3} \\ \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{n\times3} \end{bmatrix}$ is control-input matrix. The covariance matrix of process noise $\boldsymbol{Q}_k$ is:

$$\tilde{\boldsymbol{Q}}_k = \boldsymbol{B}_{k,k-1}\boldsymbol{M}_k\boldsymbol{B}_{k,k-1}^T + \boldsymbol{Q}_k \tag{14}$$

where $\boldsymbol{M}_k$ is the covariance matrix of $\boldsymbol{m}_k$. Since the accelerometer measurement output frequency is typically up to 100 Hz, the acceleration can be considered constant between adjacent measurements time and the constant variance of the acceleration σ_a^2 can be approximated to zero.

$$\tilde{\boldsymbol{Q}}_k \approx \boldsymbol{B}_{k,k-1}\boldsymbol{M}_k\boldsymbol{B}_{k,k-1}^T \tag{15}$$

By combining the accelerometer measurements, the RTK-GNSS/IMU Integration can more accurately predict the next epoch double-differencing carrier phase, pseudo-range and Doppler, and these measurements are subtracted from the actual measurements of the RTK-GNSS receiver to obtain the measurement residuals, which are subjected to Kalman filter to obtain the current epoch optimal state estimate $\hat{\boldsymbol{x}}$. The components $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{v}}$ of $\hat{\boldsymbol{x}}$ are the best estimated position and velocity, $\hat{\boldsymbol{b}}^b$ is the accelerometer bias estimate, and $N_{ub}^{s_i s_1}$ is double-differencing float ambiguity estimate.

The Rotation matrix $\boldsymbol{C}_{k,b}^{e}$ is determined by the attitude of the UAV. In general, the attitude of the UAV is expressed as a unit quaternion $\boldsymbol{q}$. The Rotation matrix $\boldsymbol{C}_{k,b}^{e}$ expressed as a unit quaternion is defined as:

$$\boldsymbol{C}_{k,b}^{e}(\boldsymbol{q}) = \begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2+q_0q_3) & 2(q_1q_3-q_0q_2) \\ 2(q_1q_2-q_0q_3) & q_0^2-q_1^2+q_2^2-q_3^2 & 2(q_2q_3+q_0q_1) \\ 2(q_1q_3+q_0q_2) & 2(q_2q_3-q_0q_1) & q_0^2-q_1^2-q_2^2+q_3^2 \end{bmatrix} \tag{16}$$

$$\boldsymbol{q} = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}^T \tag{17}$$

The implementation of RTK-GNSS/IMU Integration is shown in Fig. 1. The IMU acceleration measurement equations expressed in body-fixed-frame is defined as:

$$\boldsymbol{g}^b = -(\boldsymbol{f}^b - \boldsymbol{a}^b - \boldsymbol{b}^b) \tag{18}$$

where $\boldsymbol{g}^b$ is gravitational acceleration, $\boldsymbol{a}^b$ is motion acceleration of the UAV and can be approximated as:

$$\boldsymbol{a}^b \approx \boldsymbol{C}_{k,e}^{b} \frac{(\hat{\boldsymbol{v}}_k - \hat{\boldsymbol{v}}_{k-1})}{\Delta t} \tag{19}$$

The best estimate of the gravitational acceleration is expressed as:

$$\hat{\boldsymbol{g}}^b = \boldsymbol{C}_{k,e}^{b} \cdot \bar{\boldsymbol{g}}^e \tag{20}$$

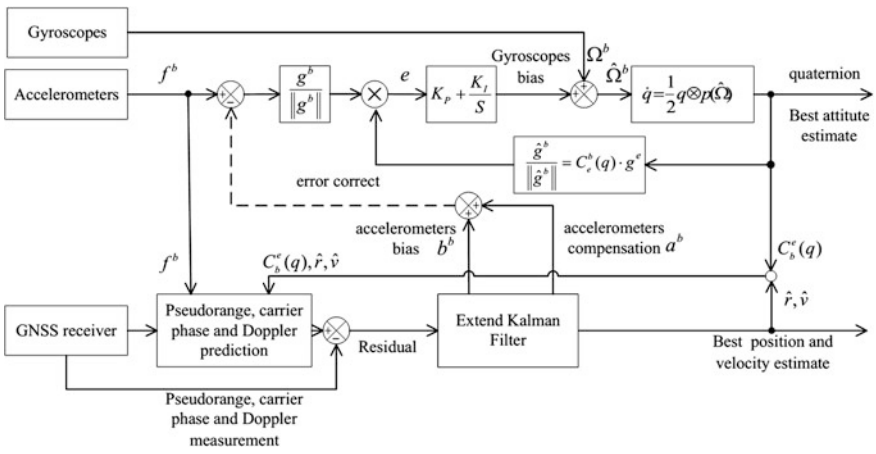


Fig. 1 Diagram of RTK-GNSS/IMU integration method

where $\bar{\boldsymbol{g}}^e$ is gravitational acceleration constant expressed in ECEF frame. The error vector between $\boldsymbol{g}^b$ and $\hat{\boldsymbol{g}}^b$ is there unit cross product:

$$\boldsymbol{e} = \frac{\boldsymbol{g}^b}{\|\boldsymbol{g}^b\|} \times \frac{\hat{\boldsymbol{g}}^b}{\|\hat{\boldsymbol{g}}^b\|} \tag{21}$$

The gyro angular rate is corrected by a proportional-integral block:

$$\hat{\Omega}^b = \Omega^b + K_p \boldsymbol{e} + K_I \int e \tag{22}$$

The proportional-integral term governs the frequency cross-over between accelerometer based attitude estimates and integrated gyro estimates [2, 4]. Using the corrected gyro angular rate $\hat{\Omega}^b$ to update quaternion:

$$\dot{\boldsymbol{q}}_{k-1} = \frac{1}{2}\boldsymbol{q}_{k-1} \otimes \boldsymbol{p}(\hat{\Omega}^b) \tag{23}$$

$$\boldsymbol{q}_k = \dot{\boldsymbol{q}}_{k-1} \cdot \Delta t \tag{24}$$

where $\boldsymbol{p}(\hat{\Omega}^b) = \begin{bmatrix} 0 & \hat{\Omega}^b \end{bmatrix}^T$. The attitude of the UAV and rotation matrix $\boldsymbol{C}^e_{k,b}(\boldsymbol{q}_k)$ can be obtained by Eq. (16).

4 Field Experiment

4.1 Experiment Approach

To verify the performance of the proposed RTK-GNSS/IMU integration, we conducted an experiment of the small UAV. The experiment setup was as follows:

(1) UAV platform: wheelbase 250 mm unmanned aerial vehicle.
(2) Base station: CTI N71, antenna CHCC220GR.
(3) Reference receiver: ComNav K505 triple frequency RTK receiver.
(4) Airborne Antenna: triple helical antenna.

The experiment was carried out at the playground of Beijing University of Posts and Telecommunications. The base station was erected in the new scientific

Fig. 2 Experiment UAV system

research building. The baseline length was about 0.5 km. The airborne antenna was mounted on the top of the UAV and connected to the reference and test receiver through a power divider (Fig. 2).

4.2 Experiment Result

Test result is shown in Fig. 3. There were some buildings at the south of playground, which causes lots of cycle slips. The green points denote fixed integer ambiguity, the yellow points denote float ambiguity and the red points denote standard single point positioning. By using RTK-GNSS/IMU the fix rate of ambiguity is increased about by 20% and the time to re-fix ambiguity is decreased by more than 14%.

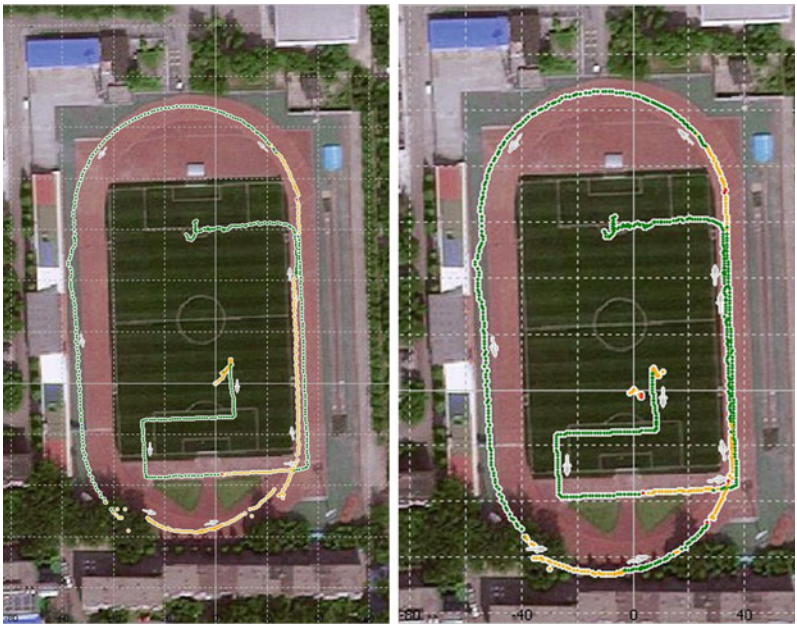


Fig. 3 Experiment track and fix rate

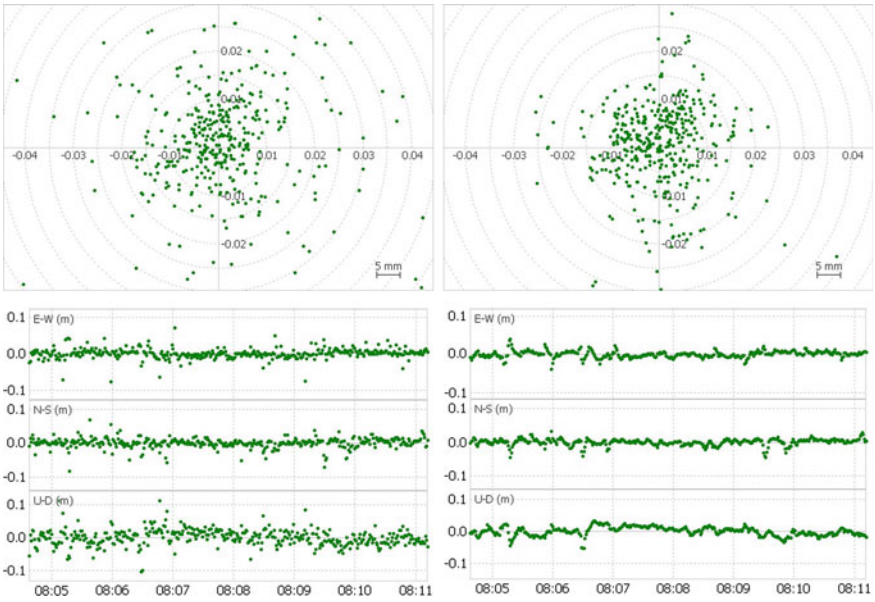


Fig. 4 Positioning error

Table 1 Experiment result between RTK-GNSS and RTK-GNSS/IMU

	RTK-GNSS	RTK-GNSS/IMU
Fix rate	51.2%	71.8%
RMS	0.021 m	0.013 m
Re-fix time	43 s	37 s

By using RTK-GNSS/IMU Integration, the positioning error is decreased from 0.021 to 0.013 m (Fig. 4, Table 1).

According to the results, the proposed RTK-GNSS/IMU integration used in Kalman filter can effectively improve the fix rate of ambiguity and the re-fix time.

5 Conclusion

The proposed RTK-GNSS/IMU integration improve the accuracy of float ambiguity to increase the fix rate and re-fix time and the proposed integration work perfectly in high dynamic and high vibration especially in the UAV.

It is predicted that by 2020 the output value of China's unmanned aircraft will reach billions of dollars. With the implementation of China's "civil space infrastructure, long-term development plan" and the maturing national differential network stations are maturing, Low-cost RTK receiver system in the field of consumer UAV has broad prospects. The method proposed in this paper is of great practical significance to enhance the usability of single-frequency RTK-GNSS receivers, and can be used for the integration and development of the UAV and location services.

Acknowledgements This work is supported by the National Key Research and Development Program (No. 2016yfb0502001), National Natural Science Foundation of China (No. 61372110) and National High Technology Research and Development Program of China (No. 2014AA123103).

References

1. Kaplan ED (2007) Understanding GPS: principle and applications. Publishing House of Eletronics Industry, Beijing
2. Euston M, Coote P, Mahony R (2008) A complementary filter for attitude estimation of a fixedwing UAV. In: International conference on intelligent robots and systems. IEEE, pp 340–345
3. Kalman RE (1960) A new approach to liner filtering and prediction problems. J Basic Eng 82(1):35–45
4. Mahony R, Hamel T, Pflimlin JM (2005) Complementary filter design on the special orthogonal group. In: Proceedings of the 44th IEEE conference on decision and control. IEEE
5. Teunissen P (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geodesy 70(1–2):65–82

Accuracy Analysis of GNSS/INS Deeply-Coupled Receiver for Strong Earthquake Motion

Hengrong Liu, Tisheng Zhang, Penghui Zhang, Farui Qi and Zhuo Li

Abstract With the development of high-frequency GNSS receivers and precision positioning technology, GNSS has been one of the main observation means to obtain high precision crustal deformation information in seismic monitoring. However, the measurement accuracy of carrier phase of GNSS receiver is obviously decreased under the motion condition, which affects the performance of GNSS. The accelerograph and the GNSS are highly complementary, and the combination of the two has been extensively studied in seismic monitoring in recent years. The loose couple and the tight-couple are integrated at the data level, which has no help to improve GNSS carrier phase accuracy. GNSS/INS deep integration implements the signal level aiding of GNSS receiver by INS (Inertial Navigation System), which can improve the carrier phase accuracy. In this paper, we focus on analyzing carrier phase accuracy of the GNSS/INS deeply-coupled receiver under strong earthquake conditions. First, we discuss the spectrum characteristics of strong earthquake signals, and analyze the tracking error of PLL (Phase-locked Loops) before and after IMU (Inertial Measurement Unit) aiding under strong earthquake based on the error model. Then, the carrier tracking performance of the loop before and after IMU aiding is tested on our self-developed software deep integration receiver, with strong earthquake motion generated by a GPS/INS hardware signal simulator. The analysis and testing results show that the tracking error of PLL with INS aiding is significantly smaller than traditional PLL. The RMS of phase discriminator error of INS-aided carrier PLL with the optimized bandwidth is smaller than 6° under the 2G single frequency simulation conditions, which is consistent with the analysis.

Keywords Strong earthquake motion · Inertial navigation system · GNSS/INS deeply-coupled receiver · Carrier phase

H. Liu (✉) · T. Zhang · P. Zhang · F. Qi · Z. Li
GNSS Reasearch Center, Wuhan University, Wuhan, China
e-mail: 594647662@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_28

1 Introduction

Strong earthquakes means earthquakes with a magnitude of not less than 6. When the strong earthquake occurs, real-time reliable seismic monitoring can not only help people take steps in time, which can reduce casualties and economic losses, but also can provide data sources for seismological research. The accelerograph records the seismic information by inertial measurement unit, but the integral will amplify the low frequency noise in the data processing, which results in the low measurement accuracy for the low frequency components in seismic signal [1]. Conversely, GNSS technology, as one of the main observation means to obtain high precision crustal deformation information in seismic monitoring, can accurately measure low-frequency components under the motion condition. However, the measurement accuracy of the carrier phase of the GNSS receiver is limited by the tracking loop, which affects the performance of GNSS receiver under the complex dynamic conditions [2].

In order to solve the problem of seismic measurement accuracy and adapt to complex earthquake scene, the combination of the accelerograph and the GNSS has been extensively studied in recent years. The researchers of GFZ in Germany used the loosely-coupled Kalman filter to process the results of GPS single frequency receiver and MEMS (Micro-Electro-Mechanical System) inertial sensor under simulated strong earthquake motion, and found that the precision was obviously better than that of accelerograph [3]. In the United States, Geng J compared the results of tightly-coupled Kalman filter and the loosely-coupled Kalman filter to the accelerograph and GNSS, and concluded that the performance of the tightly-coupled filter is better. In conclusion, most previous studies have shown that the combination of the accelerograph and the GNSS can improve the measurement accuracy under the strong earthquake condition [4]. However, the loose couple and the tight couple are integrated at the data level, which has no help to improve GNSS measurement accuracy, so that they can not solve the problem of the declined performance of GNSS receivers essentially.

The GNSS/INS deeply-couple technology can implement the signal level aiding of GNSS receiver by INS. Therefore, in this paper, we apply this technology to the research of strong-earthquake measurement, and focus on analyzing and testing the carrier phase measurement accuracy of it under the strong earthquake motion.

2 Analysis of Strong Seismic Signal

In this chapter, based on its acceleration sequence and spectral distribution, the signal characteristics of the actual seismic wave are studied. On June 14, 2008, Iwate-Miyagi Inland Earthquake occurred in Japan. The ground-wave magnitude was 7.2 ms. The accelerograph at IWTH25KiK-net recorded 100 Hz acceleration

sequence in three axes. Figure 1 shows the 60-s acceleration sequence in the north–south (N–S) direction.

As shown in Fig. 1, the maximum acceleration of the seismic wave reaches 1036 gal, so it is a strong seismic wave. Then based on the amplitude-frequency characteristics analysis, the power spectrum can be obtained by using SeismoSignal Version 5.1.0 (seismic wave processing software) to process the acceleration sequence, as shown in Fig. 2, it reflects the vibration energy of each frequency component of seismic wave.

As is seen in Fig. 2, the frequency spectrum of the strong seismic signal is very complex. With the frequency increases, the dynamic energy increases first and then decreases, and mainly concentrates in 1–10 Hz.

Therefore, based on the characteristics of the acceleration sequence of strong seismic signal, the GNSS signal simulator is used to design a 2G single-frequency

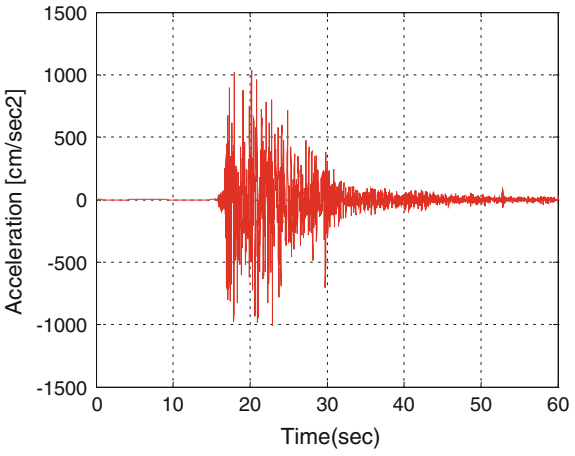


Fig. 1 Acceleration sequence of strong seismic signal

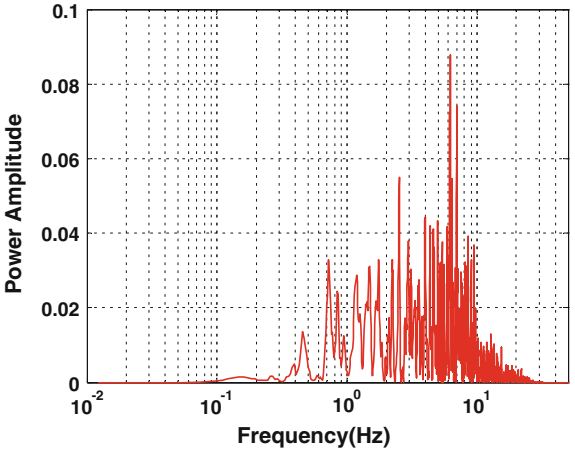


Fig. 2 Power spectrum of acceleration sequence

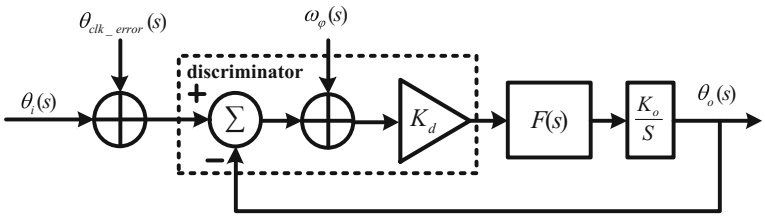


Fig. 3 Error propagation model of traditional PLL

sine-wave motion scene to test the tracking error of traditional loop with different frequency signals.

3 Analysis of Error Model

In this chapter, based on the tracking error model, the tracking errors of the traditional PLL and the deeply-coupled PLL are analysed.

3.1 Error Analysis of Traditional PLL

The error propagation model of traditional PLL is shown in Fig. 3.

Where $\theta_i(s)$ and $\theta_0(s)$ is the phase of the PLL input signal and output signal, $\theta_{clk\_error}(s)$ is the local oscillator error, brought by $\theta_i(s)$ during down conversion processing, $\omega_\varphi(s)$ is thermal noise. In addition, the dynamic change can also lead to tracking error. Therefore, the phase error of GNSS receiver PLL is composed of thermal noise, vibration-induced oscillator jitter, Allan variance-induced oscillator jitter and dynamic stress error, which denoted as σ_{tPLL}, σ_{rv}, σ_{rA}, θ_e respectively. Different error are not related to each other, so the mean squared error of total error can be written as follows:

$$\sigma_{PLL} = \sqrt{\sigma_{tPLL}^2 + \sigma_{rv}^2 + \sigma_{tA}^2} + \theta_e/3 \tag{1}$$

For the second-order traditional PLL, the formula (1) can be expanded as follows [5]:

$$\sigma_{tPLL}^2 = (\frac{180}{\pi})^2 \cdot \frac{B_n}{C/N_0} \cdot \left(1 + \frac{1}{2T_{coh} \cdot C/N_0}\right)(^\circ) \tag{2}$$

$$\sigma_{rv}^2 = \left(\frac{180}{\pi}\right)^2 \cdot \frac{\pi^2 f_0^2 K_g^2 G_g}{2.67 B_n}(^\circ) \tag{3}$$

$$\sigma_{tA}^2 = \left(\frac{180}{\pi}\right)^2 \cdot 2\pi^2 f_0^2 \cdot \left(\frac{\pi^2 h_{-2}}{\sqrt{2}(1.89B_n)} + \frac{\pi h_{-1}}{4(1.89B_n)^2} + \frac{h_0}{4\sqrt{2}(1.89B_n)}\right)(^\circ) \quad (4)$$

$$\theta_e = \frac{\Delta\Re}{(1.89B_n)^2}(^\circ) \quad (5)$$

where B_n is noise bandwidth, C/N_0 is carrier-to-noise ratio, T_{coh} is coherent integration time, f_0 is carrier frequency, K_g is oscillator sensitivity, G_g single-sided vibration spectral density, h_i is a constant, $\Delta\Re$ is frequency ramp-up amplitude.

Most of the receivers use the Costas PLL, which is insensitive to 180° phase inversion. So when the σ_{tPLL} is three times larger than one quarter of the traction range of discriminator, in other words, when the σ_{tPLL} is more than 15°, the PLL is considered to be lost.

Figures 4 and 5 shows the all errors of second-order PLL with different parameter when the carrier acceleration is 2G. Figure 4 shows the all errors with different bandwidths when the carrier-to-noise ratio is 40 dB-Hz and the coherent integration time is 1 ms. As is seen in this figure, the main error of the loop is dynamic stress error, and when the bandwidth is less than 18 Hz, the total error is greater than the threshold, the loop is lost. Figure 5 shows the all errors with different signal strength when the bandwidth is 20 Hz and coherent integration time also is 1 ms. As is seen in this figure, the dynamic stress error is independent of the signal strength, and due to the thermal noise, when the carrier-to-noise ratio is lower than 34 dB-Hz, the loop is lost. In short, both the dynamic performance and anti-noise performance should be considered in selecting bandwidth. However, the bandwidth can be reduced by IMU aiding, which can reduce the tracking error of the PLL under dynamic conditions.

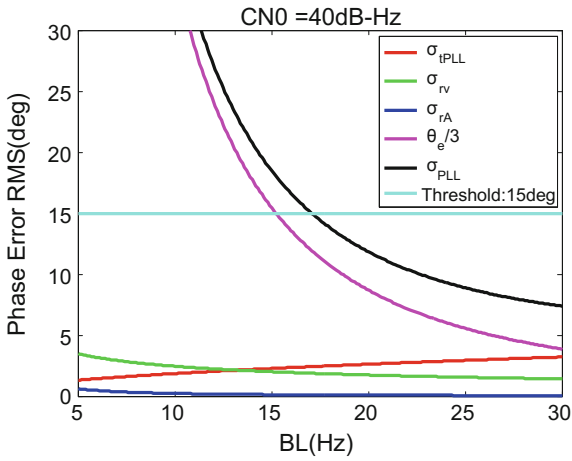


Fig. 4 All errors of second-order traditional PLL with different bandwidth

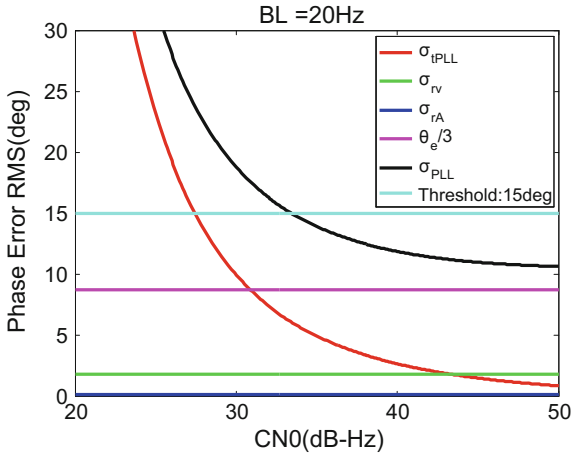


Fig. 5 All errors of second-order traditional PLL with different signal strength

3.2 *Error Analysis of Deeply-Coupled PLL*

According to the author's previous research results [6], the error propagation model of deeply-coupled carrier PLL is shown in Fig. 6.

Where $\Delta f_{IMU}(s)$ is the Doppler estimation error of auxiliary information caused by the IMU error, and $f_{aid}(s)$ is the Doppler shift calculated from auxiliary information. The feed-forward branch of the deeply-coupled PLL provides real-time carrier motion information for the tracking loop, which eliminates the tracking error caused by the motion of the carrier, so that this loop only needs to track the Doppler estimation error of auxiliary information. So Doppler estimation error is analysed below.

According to the previous research, the initial velocity error of GNSS revised is the main error source of the feed-forward branch in IMU errors, which including initial velocity error, initial attitude error, sensor bias, sensor scaling factor, etc. Therefore, in order to simplify the analysis, only the Doppler estimation error caused by the initial velocity error is considered in this paper, denoted as σ_{V0}, and the formula is as follows:

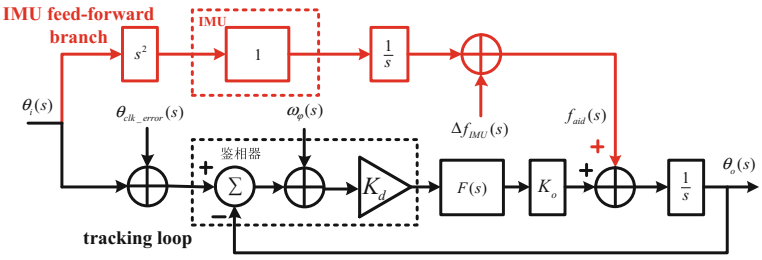


Fig. 6 Error propagation model of deep-coupled carrier PLL

$$\sigma_{V0}^2 = \left(\frac{180}{\pi}\right)^2 \cdot \frac{2\pi \cdot \delta V_N(0)}{\lambda \cdot e \cdot \omega_n}(^\circ) \tag{6}$$

where $\delta V_N(0)$ is the initial velocity error, λ is the carrier wavelength, and ω_n is the characteristic frequency. Although $\delta V_N(0)$ will change with time and converge to zero, the general update cycle of integration is 1 s, so that the cycle of $\delta V_N(0)$ is 1 s. Therefore, $\delta V_N(0)$ should be considered as an integral part of the loop steady-state errors. So the formula (1) becomes:

$$\sigma_{PLL} = \sqrt{\sigma_{tPLL}^2 + \sigma_{rv}^2 + \sigma_{tA}^2 + \sigma_{V0}^2} \tag{7}$$

According to the above analysis, taking the MEMS inertial measurement unit MTI-G shown in Table 1 as an example, the all errors of second-order deeply-coupled PLL with MTI-G aiding under can be drawn with different signal strength and bandwidth, as shown in Figs. 7 and 8.

Table 1 Error parameter of typical IMU

	Gyroscope	Accelerometer 加速度计
MTI-G (MEMS)	Bias: 15°/h	Bias: 800 mGal
	Bias instability: 100°/h	Bias instability: 2000 mGal
	Correlation time: 600 s	Correlation time: 600 s
	Scaling factor: 1000 ppm	Scaling factor: 1000 ppm
	ARW: 3°/h	VRW: 0.12 m/s/$\sqrt{h}$
Initial velocity error	$\delta_{VN}(0)$: 0.04 m/s	

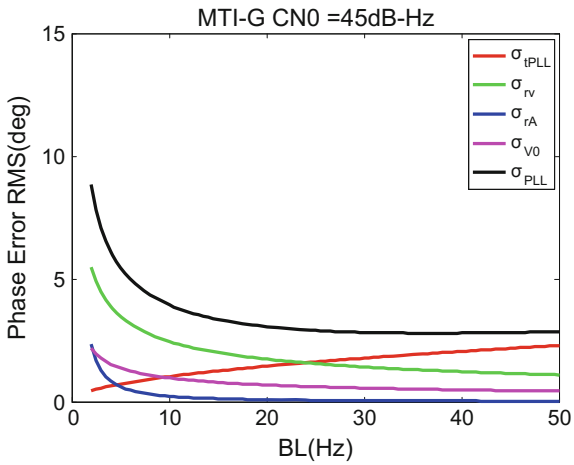


Fig. 7 All errors of second-order deep-coupled PLL with 45 dB-Hz carrier-to-noise ratio

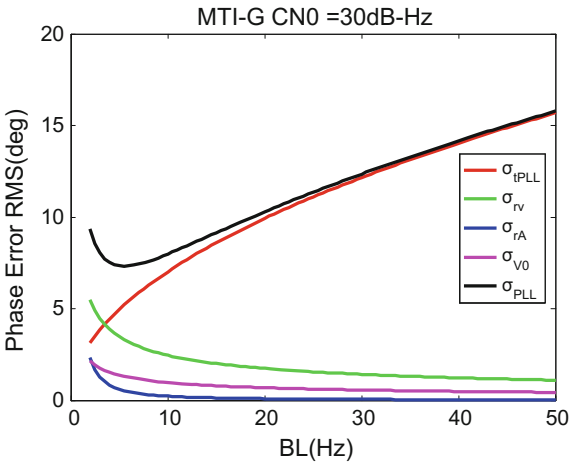


Fig. 8 All errors of second-order deeply-coupled PLL with 30 dB-Hz carrier-to-noise ratio

It can be seen from the figures that the influence of different signal strength on thermal noise error is very large, and on other errors is very small. When the signal strength is strong, the lower the bandwidth is, the larger the influence of Doppler estimation error is under the dynamic conditions, and the optimal bandwidth of deeply-coupled loop is not less than 15 Hz, which the corresponding phase error RMS is 3.3°. When the signal strength is weak, thermal noise is always the main error under the dynamic conditions, and the optimal bandwidth of deeply-coupled loop is about 5 Hz, which the corresponding RMS error is 7.3°.

4 Test Verification and Analysis

In this paper, the GNSS/INS signal simulator is used to simulate strong seismic signal, the self-developed GNSS/INS signal recorder is used to collect the signal, and the self-developed deeply-coupled software system is used to process the data. In addition, this paper only simulates the circular motion with the same attitude on the horizontal plane. The scenes that set on the GNSS signal simulator are single-frequency sine-wave motion with the frequency of 1 and 2 Hz, the maximum acceleration of 2G and 100-second dynamic duration. During the simulation period, the satellite G28 is located at the zenith, while G06 and G10 are in the east and south directions respectively, which the elevation both are the lowest.

Figure 9 shows the angular velocity and acceleration of the carrier under 1 Hz simulation scene. As is seen in this Figure, there are only horizontal motion in this scene, and the waveform acceleration in the north and east directions are sinusoidal and cosine wave with the amplitude of 2G and the frequency of 1 Hz, respectively.

Figures 10 and 11 show the discriminator output of the second-order traditional PLL and the second-order deeply-coupled PLL MTI-G aiding with while tracking low-elevation G10 satellite signal in the 1 Hz simulation scene, and the coherent

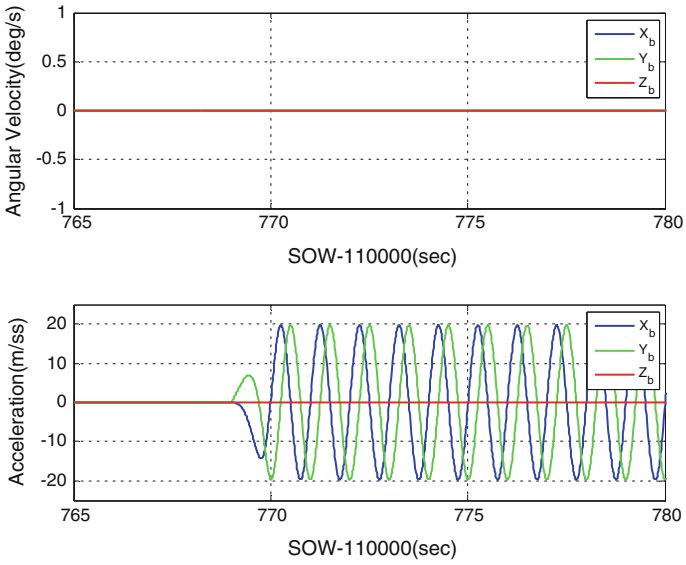


Fig. 9 Motion information of 1 Hz scene

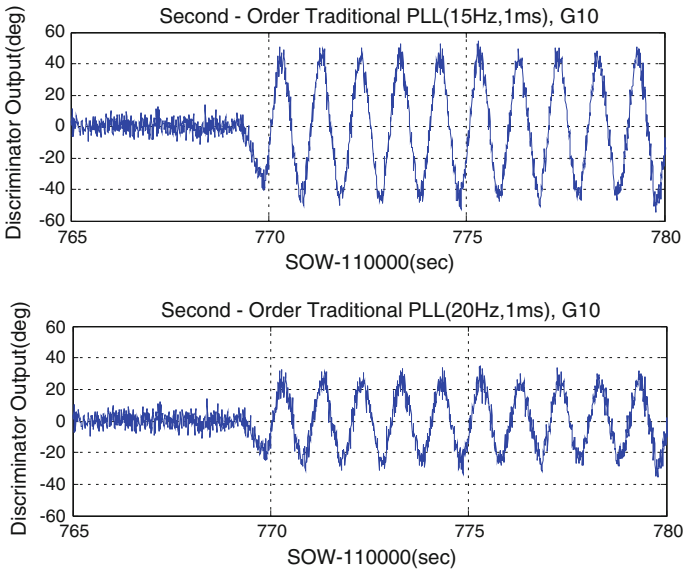


Fig. 10 Discriminator output of second-order traditional PLL

integration time is 1 ms, the bandwidths are 15 and 20 Hz. It can be seen that in the high dynamic case, the phase error of the deeply-coupled PLL is much smaller than that of the traditional PLL.

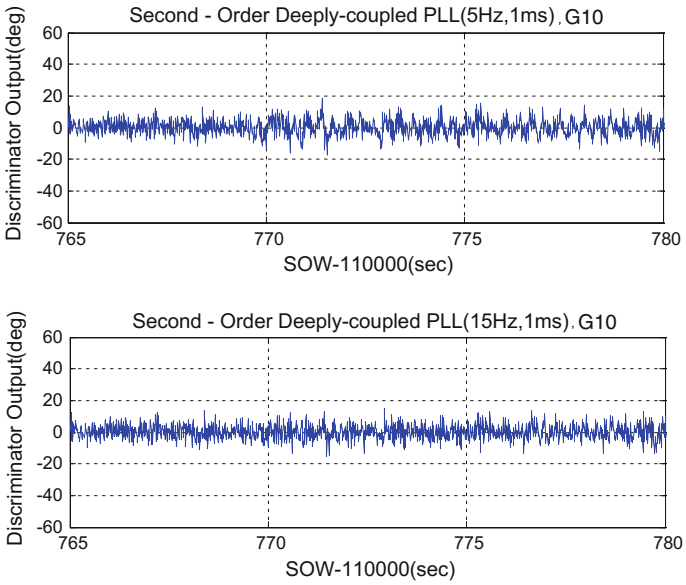


Fig. 11 Discriminator output of deeply-coupled traditional PLL

The following is a statistical analysis. The output of each phase error source in the discriminator is different, and the thermal noise variance has the following relationship with the discriminator output:

$$\sigma_{\delta\vartheta} \cdot \sqrt{2B_L T} = \sigma_{\mathrm{tPLL}} \tag{8}$$

Referring to Eqs. (1)–(7), Table 2 shows that the discriminator error statistics for the 1 and 2 Hz scenes, where the integration time is 1 ms and regardless of the oscillator error. It can be seen from the table that the traditional PLL can reduce the discriminator error by increasing the bandwidth, but the error is still large, which will lead to loop instability. Contrarily, the discriminator error of deeply-coupled PLL is obviously reduced, and the RMS of discriminator error is less than 6° with the optimal bandwidth of 15 Hz, which agrees well with the theory. Meanwhile, there is a greater dynamic frequency in the 2 Hz scene than in the 1 Hz scene, which resulting in a slight increase discriminator error. In conclusion, the carrier tracking accuracy can be significantly improved by MEMS IMU aiding in the strong earthquake scene.

Table 2 Tracking discriminator error statistics of G10 in various scenes (*Unit* Deg)

	Bandwidth (Hz)	RMS of phase error In the static period (1 Hz)	RMS of phase error In the dynamic period (1 Hz)	RMS of phase error In the static period (2 Hz)	RMS of phase error In the dynamic period (2 Hz)
Traditional PLL	15	4.48	30.54	4.52	29.07
	20	4.49	17.60	4.57	17.09
Deeply-coupled PLL	5	4.49	5.20	4.48	6.53
	15	4.49	4.67	4.52	5.57

5 Summary

In this paper, based on the analysis of the spectrum characteristics of strong seismic signals, the tracking errors of PPL before and after IMU aiding are analysed theoretically according to the tracking error model, and tested on our self-developed software deeply-coupled receiver, with simulate strong earthquake motion generated by a GPS/INS hardware signal simulator. Theoretical analysis and experimental tests show that GNSS/INS deeply-coupled technology can greatly improve the measurement accuracy of GNSS receiver under the strong earthquake motion. Next, experimental verification will be carried out, and the algorithm of the deeply-coupled tracking loop will be optimized according to the results.

References

1. Bock Y, Melgar D, Crowell BW (2011) Real-time strong-motion broadband displacements from collocated GPS and accelerometers. Bull Seismol Soc Am 101(6):2904–2925
2. Moschas F, Stiros S (2015) PLL bandwidth and noise in 100 Hz GPS measurements. GPS Solutions 19(2):173–185
3. Tu R, Wang R, Ge M et al (2013) Cost-effective monitoring of ground motion related to earthquakes, landslides, or volcanic activity by joint use of a single-frequency GPS and a MEMS accelerometer. Geophys Res Lett 40(15):3825–3829
4. Geng J, Bock Y, Melgar D et al (2013) A new seismogeodetic approach applied to GPS and accelerometer observations of the 2012 Brawley seismic swarm: implications for earthquake early warning. Geochem Geophys Geosyst 14(7):2124–2142
5. Kaplan E, Hegarty C (2006) Understanding GPS: principles and applications, 2nd edn. Artech House, Norwood, MA, pp 3–30
6. Zhang T, Niu X, Ban Y., Zhang H, Shi C, Liu J (2015) Modeling and development of INS-aided PLLs in a GNSS/INS deeply-coupled hardware prototype for dynamic applications. Sensors 15(1):733–759

Research on MEMS IMU Aided BeiDou Receiver Carrier Loop Technology

Chunyu Liu, Lei Chen, Yangbo Huang, Ling Yong, Shaojie Ni and Feixue Wang

Abstract In dynamic scenarios, the loop bandwidth of BeiDou receiver must be sufficiently wide and the integration time should be short to ensure stable tracking. It will introduce too much noise to the tracking loop with the increase of bandwidth and the reduction of integration time, resulting in the loss of lock of BeiDou receiver carrier tracking loop in high dynamic scenarios. Therefore, high dynamic position is hard to be achieved. In this paper, steady tracking performance with auxiliary has been performed under different C/N_0 (carrier to noise ratio) and dynamic scenarios, and we propose the range of accuracy of the PLL effectively aided by MEMS IMU under different orders, and then analyze the auxiliary optimal bandwidth under different orders. Simulation results show the improvement of bandwidth with auxiliary under different grades of MEMS IMU and different orders of PLL.

Keywords BeiDou · MEMS IMU · PLL · Optimal bandwidth · Tracking error

1 Introduction

In December 27, 2012, BeiDou regional satellite navigation system officially began to provide positioning, navigation and timing services in the Asia-Pacific region. For the pseudo-range and carrier phase measurement of BeiDou and GPS, the accuracy can achieve the same grade [1]. BeiDou satellite signals are easily blocked in some cases, such as indoor, lush trees, urban canyon and other scenarios. Additionally, in high dynamic scenarios, such as various tactical missiles and moving carrier with high speed, the receiver cannot balance the influence of high dynamic and noise. Consequently, it cannot track satellite signals stably and is difficult to achieve high dynamic positioning and navigation.

C. Liu · L. Chen · Y. Huang (✉) · L. Yong · S. Ni · F. Wang
Satellite Navigation Research and Development Center,
National University of Defense Technology, Changsha, China
e-mail: yangbohuang@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_29

Inertial Navigation System (INS) is a kind of autonomous navigation system. It has the advantage of excellent dynamic response, high data updating rate, strong anti-interference and the ability of outputting attitude information, which is complementary to BeiDou. From the perspective of INS error, it can be divided into strategic grade, navigation grade, tactical grade and commercial grade [2]. The corresponding INS sensor precision index is shown in Table 1.

Since the inertial navigation system is expensive, this limits the application of BeiDou/INS integrated navigation system. MEMS is a new product of Micro Electronic Mechanical System device. Inertial measuring element (IMU) is the application of MEMS technology, which can measure the linear acceleration and angular velocity of moving objects. MEMS IMU is a new strap-down inertial navigation system. It has the advantage of small volume, light weight, low consumption, low cost, strong anti vibration and high reliability in a variety of tactical weapons guidance system [3].

The world's leading MEMS IMU research institutes are BAE, BEI, Draper, Honeywell and Analog Device (ADI). The following table lists the major MEMS IMU product models of Honeywell and ADI. The performance of tactical grade MEMS IMU is further improved with volume and weight decreased gradually. The MEMS acceleration index of some models has reached to the navigation grade. At present, the accuracy of commercial MEMS gyro has been up to the tactical grade, and the accuracy of international laboratory for MEMS gyro has been better than 0.01°/h [4].

In order to track high dynamic stress under high dynamic scenarios, the bandwidth of PLL needs to be increased. However, the larger bandwidth, the more noise will be introduced, thus the tracking loop will be easier loss of lock. To solve this problem, this paper studies MEMS IMU aided BeiDou receiver scalar PLL.

2 MEMS IMU Aided PLL Model

In order to achieve a more stable tracking performance, PLL can be assisted by the carrier dynamic information from the MEMS IMU, the same as the code loop supported by the carrier loop, the dynamic stress which the PLL needs to detect and endure will be greatly reduced, and the tracking performance can be improved by compressing the bandwidth and increasing the integration time. Figure 1 shows the MEMS aided PLL model [5].

In Fig. 1, $x_0(s)$ is the input reference signal, $\theta_e(s)$ is the dynamic tracking error, $F(s)$ means the loop filter function, the loop gain is $K_o \cdot K_d$, NCO denotes the numerically controlled oscillator, $1/s$ is the integrator, $G(s)$ is the equivalent MEMS IMU information filtering. For the convenience, the loop gain is set to 1, the following content focus on designing loop filter, and analysis of the performance of PLL assisted by MEMS IMU.

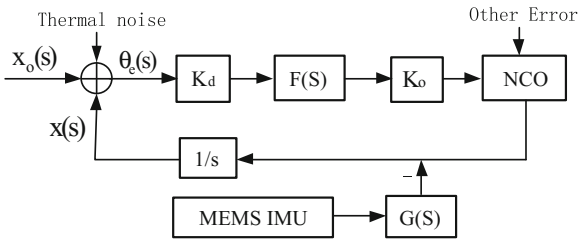


Fig. 1 Structure diagram of MEMS IMU aided PLL

To simplify the MEMS IMU error model, a one-dimensional error model can be obtained by using only the acceleration bias ε_a and the gyro drift ε_g, where $\overrightarrow{a}$ denotes the specific force, x is the position parameter, g stands for the gravitation, R is the earth radius, t is the time and $\sqrt{g/R}$ is the Schuler frequency.

$$\overrightarrow{a} = \ddot{x} + \frac{g}{R} \cdot x - \varepsilon_a - \varepsilon_g \cdot g \cdot t \tag{1}$$

Figure 1 shows the dynamic tracking loop error $\theta_e(s) = x_0(s) - x(s)$, taking it into the loop transfer function we can obtain:

$$\theta_e(s) = \frac{s - F(s) \cdot G(s) \cdot \frac{s^2 + g/R}{s^2}}{s + F(s)} \cdot x_0(s) + \frac{F(s) \cdot G(s)}{s + F(s)} \cdot \left(\frac{\varepsilon_a}{s^3} + \frac{g \cdot \varepsilon_g}{s^4}\right) \tag{2}$$

In order to ensure the independence between the dynamic tracking error and the input signal, the first term of Eq. (2) must be 0, thus we can obtain:

$$\Gamma_0 s = F(s) \cdot G(s) \cdot \frac{s^2 + g/R}{s^2} \tag{3}$$

In Eq. (3), Γ_0 represents the coupling coefficient between BeiDou and MEMS IMU, the value of 1 stands for the coupling degree and 0 denotes that the BeiDou cannot receive the aid of MEMS IMU. From Eq. (3) we can get equation of MEMS IMU information filter, where ω_s is the Schuler frequency. If the loop filter is the ideal second-order loop, the above formula can be substituted into Eq. (2) and converted to the time domain, then we can obtain:

$$\theta_e(t) = \frac{1 - \Gamma_0}{\omega_n^2} \cdot \ddot{x}_0 + \frac{\Gamma_0 \cdot \varepsilon_a}{\omega_n^2} \cdot \cos(\omega_s \cdot t) + \frac{\Gamma_0 \cdot \varepsilon_g \cdot g}{\omega_n^2} \cdot \frac{\sin(\omega_s \cdot t)}{\omega_s} \tag{4}$$

If the loop filter is the ideal third-order loop, the dynamic tracking error can also be obtained:

$$\theta_e(t) = \frac{1-\Gamma_0}{\omega_n^3} \cdot \dddot{x}_0 - \frac{\Gamma_0 \cdot \varepsilon_a}{\omega_n^3} \cdot \frac{\sin(\omega_s \cdot t)}{\omega_s} + \frac{\Gamma_0 \cdot \varepsilon_g \cdot g}{\omega_n^3} \cdot \cos(\omega_s \cdot t) \tag{5}$$

When the coupling coefficient is 1, the steady-state error of the dynamic stress is no longer dependent on the carrier dynamics, rather determined entirely by the system error of MEMS IMU after compensation. Compared with the thermal noise mean square error, the error of Allan phase jitter and PLL phase jitter can be ignored. Combined with the dynamic error of the formula, the loop error expression [6] can be obtained after MEMS IMU support:

$$\sigma_{PLL} = \frac{180^\circ}{\pi} \sqrt{\frac{B_n}{C/N_0}\left(1 + \frac{1}{2T_{coh}C/N_0}\right)} + \frac{1}{3} \cdot \frac{360}{\lambda} \cdot \theta_e(t) \tag{6}$$

It is the same as the conventional PLL for the analysis of optimal bandwidth, if $\frac{\partial \sigma_{PLL}}{\partial B_n} = 0$, the formula can be expressed as:

$$B_n = \sqrt[2N+1]{\frac{\left(\frac{2N}{3}\frac{\theta_e(t)}{(1/\mu)^N}\right)^2}{\left(\frac{\lambda}{2\pi}\right)^2 \frac{1}{C/N_0}\left(1 + \frac{1}{2T_{coh}C/N_0}\right)}}, \quad \begin{cases} \mu = 0.53, & N = 2 \\ \mu = 0.7845, & N = 3 \end{cases} \tag{7}$$

We can obtain the optimal bandwidth if the derivation of time is set to 0, and the optimal bandwidth will be grow with the increasing of the error parameter of the MEMS IMU device. It can be seen from Eq. (6), the PLL supported by high precision MEMS IMU device can obtain low tracking error with high cost.

3 Effects of MEMS IMU Accuracy on PLL Tracking Error

Combining the inertial performance parameters of different grades in Table 1, HG1930 device and ADIS16480 device are taken as examples and the coherent integration time is set to 1 ms. The comparison of auxiliary second-order PLL tracking error curves of ADIS16480 and HG1930 under the 1 and 20 Hz bandwidth is shown in Fig. 2.

Table 2 shows that the gyro performance of HG1930 is better than that of ADIS16480, and the accelerometer performance of ADIS16480 is better than that

Table 1 Accuracy grade of inertial sensor

Performance	Strategic	Navigation	Tactical	Commercial
Gyro drift (°/h)	0.0001	0.01	1	100
Accel bias (mg)	0.001	0.01	0.5	10

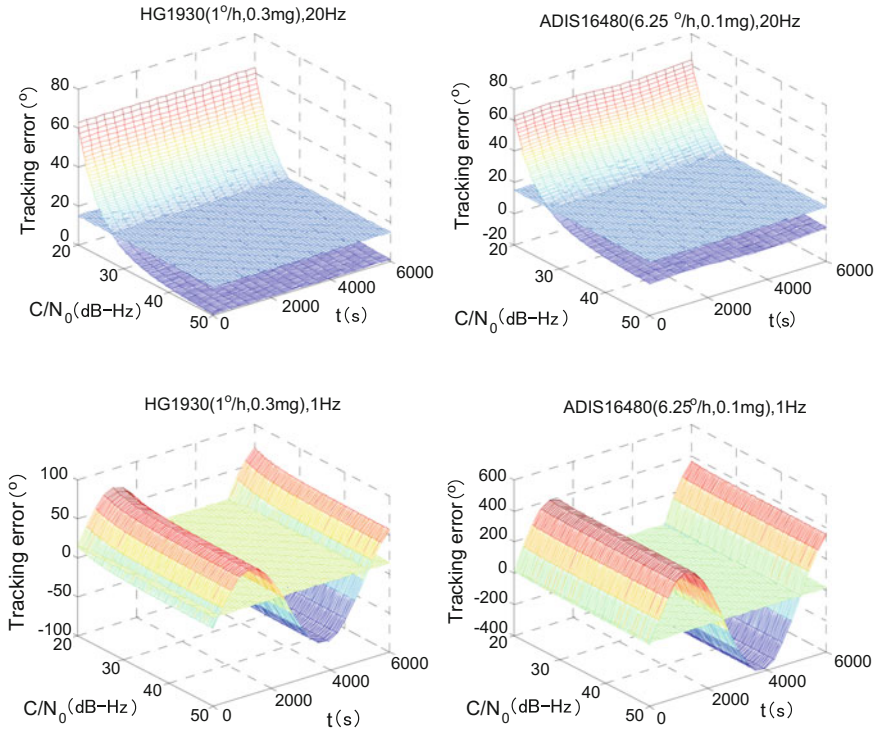


Fig. 2 MEMS IMU aided second-order PLL tracking error curves

Table 2 Product model and performance of MEMS IMU

Institutes	Model	Gyro drift (°/h)	Accel bias (mg)
	HG1900	1	0.3
Honeywell	HG1920	6	0.6
	HG1930	1	0.3
	ADIS16460	8	0.2
ADI	ADIS16480	6.25	0.1
	ADIS16490	1.8	0.0036

of HG1930. Simulation results show that the loop tracking error of the tactical grade MEMS IMU can be kept below the threshold (15°) when the C/N_0 is greater than 27 dB-Hz and the bandwidth is 20 Hz. When the bandwidth is reduced to 1 Hz, there is a large amplitude jitter of loop tracking error due to periodic Schuler jitter. The auxiliary PLL tracking error may exceed the threshold even when C/N_0 is relatively high, thus the auxiliary PLL tracking may not work in this scenario. The auxiliary tracking error of HG1930 under narrow bandwidth is smaller than

ADIS16480's. Hence, the accuracy of gyro is the main factor of MEMS IMU, which affects the auxiliary loop tracking error. When the bandwidth of loop is 20 Hz and the accuracy of MEMS gyro is set to the tactical grade, It can be found that the trend of loop tracking error is almost unchanged with the accuracy of the accelerometer increasing from 0.01 to 10 mg.

Tactical MEMS IMU devices have been relatively mature and the accuracy of commercial MEMS accelerometer can reach the navigation grade. However, the accuracy of commercial MEMS gyro is not up to the navigation grade. If simulations above are supported by devices with higher precision, the tracking error with auxiliary will be under the threshold. The tracking error curves of the navigation grades under 1 and 20 Hz bandwidth are shown in Fig. 3.

From Fig. 3, the threshold value in the case of 20 dB-Hz can be reduced with the decrease in PLL bandwidth. In the conjunction with the simulation results in Fig. 2, it can be seen that the cyclical jitter caused by tactical MEMS gyro can lead to the loss of lock when bandwidth is reduced to 1 Hz. To facilitate the analysis of MEMS gyro accuracy on the tracking error, we maintain the accuracy of MEMS accelerometer at the tactical grade, and set the loop bandwidth and the coherent integration time to 1 Hz and 1 ms, respectively. The trends of tracking accuracy of the gyro under 20 and 25 dB-Hz are shown in Fig. 4.

It can be seen that when the bandwidth is reduced to 1 Hz, the gyro drift value needs to be lower than 0.016°/h if the 20 dB-Hz signal is stably tracked for the second-order PLL assisted by the MEMS IMU. The gyro drift value should be lower than 0.17°/h if it is stably tracked at 25 dB-Hz weak signal. For third-order loop, devices with gyro accuracy lower than 7.47°/h can assure effective assistance under different C/N_0. Devices with gyro accuracy lower than 76.2°/h can be effective when the C/N_0 is greater than 25 dB-Hz. The critical curves of gyro drift, which achieve the stable tracking of second and third-order PLL at different C/N_0, are shown in Fig. 5.

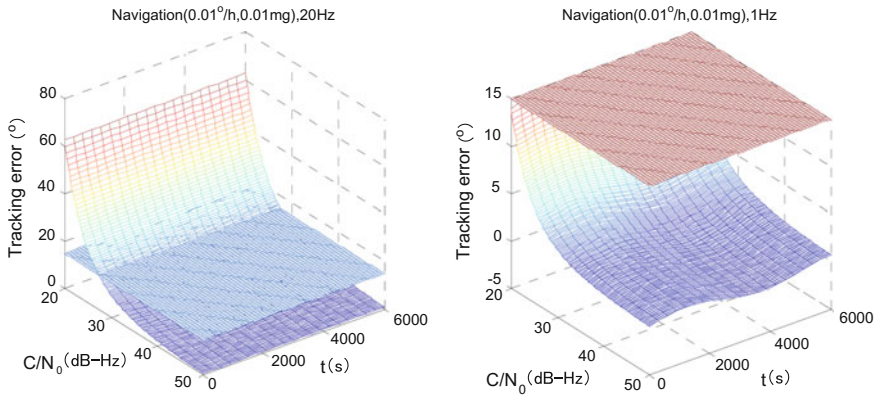


Fig. 3 MEMS IMU aided second-order PLL tracking error

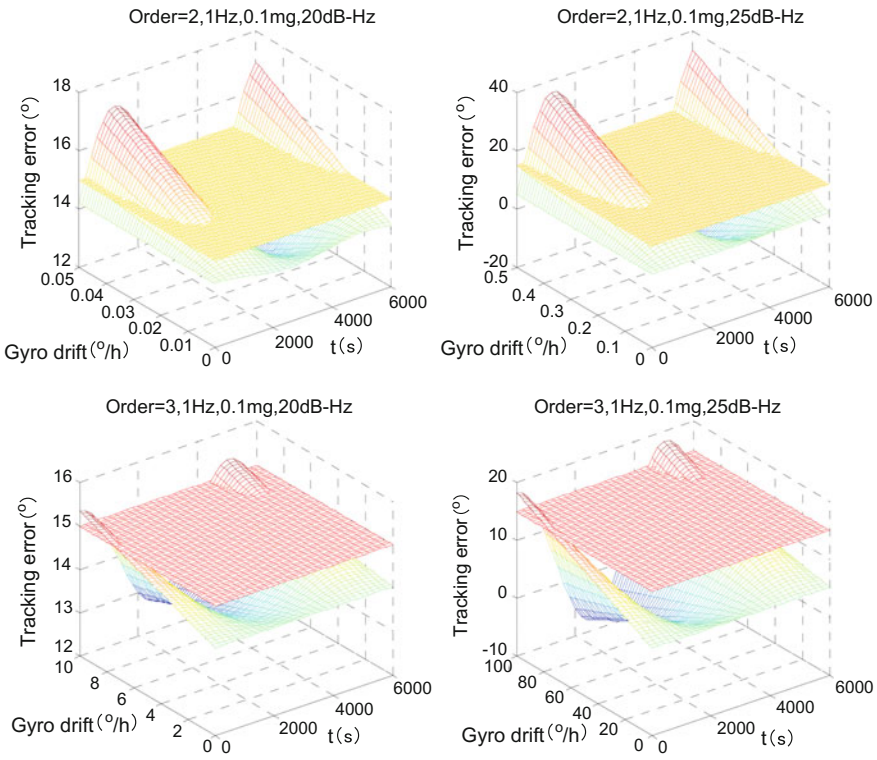


Fig. 4 Relationship between gyro drift and tracking error

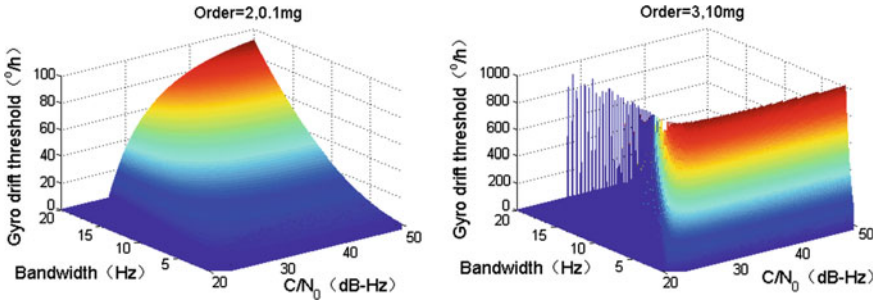


Fig. 5 Critical curves of gyro drift

The simulation results show that navigation grade devices can effectively assist the second-order PLL in 1 Hz bandwidth under weak signal with 20–25 dB-Hz. Tactical grade devices with bandwidth less than 4 Hz cannot be effectively assisted, and commercial grade devices cannot be effectively assisted. When C/N_0 is higher than 30 dB-Hz, Tactical grade devices can effectively assist the loop when the

bandwidth is greater than 2 Hz. Commercial grade devices can achieve assistance when C/N_0 is more than 50 dB-Hz and the bandwidth is more than 20 Hz. For the third-order PLL under 1 Hz bandwidth, navigation and tactical grade devices can effectively assist the 20–25 dB-Hz weak signal, and commercial grade devices can effectively assist the signal when C/N_0 is more than 30 dB-Hz.

4 Optimal Bandwidth Analysis of PLL with MEMS IMU Aiding

From above analysis, we can see that with the decrease of loop bandwidth, high precision MEMS IMU auxiliary PLL is required to ensure the tracking performance. That is, as the accuracy of the MEMS IMU increases, the bandwidth of the loop can be designed to be smaller. Figure 6 verifies the simulation of the optimal bandwidth on the tactical grades HG1930 with auxiliary PLL or not.

From Fig. 6, the decrease of optimal bandwidth with HG1930 aiding is the most obvious at low C/N_0, and the optimal bandwidth reduction of the third-order PLL is much better than that of the second-order PLL. To facilitate the comparative analysis of the degree of bandwidth with the auxiliary or not, C/N_0 is set to 20, 25, 30, 40 and 50 dB-Hz, respectively. Figure 7 shows the change of bandwidth with the auxiliary or not.

From Fig. 7, the improvement in the optimal bandwidth of the third-order PLL is significantly better than that of the second-order PLL, and the improvement of the optimal bandwidth is the same at different C/N_0. Equation (7) can be verified independently of C/N_0. Since the coupling coefficient is set to 1, the optimal bandwidth of the HG1930 auxiliary PLL is independent of the carrier dynamics. That is, the optimal bandwidth is not affected by carrier dynamics. The improvement of the optimal bandwidth is only related to the accuracy of the MEMS IMU device. To facilitate the study, we set the C/N_0 to 20 dB-Hz. Figure 8 shows the improvement of the optimal bandwidth with auxiliary of the inertial navigation devices of different grades.

From Fig. 8, for the auxiliary second-order PLL, because the accuracy of navigation grade and tactical grade devices are relatively high, the gyro drift value is 10°/h or less. It can effectively improve the 1–50 g dynamic scenarios assisted loop optimal bandwidth, in which the optimal bandwidth with navigation grade aiding is improved by 96%. The accuracy grade of commercial device is relatively low, when the gyro drift value is higher than 100°/h. Optimal bandwidth under 0–5 g dynamic has no improvement because of the larger error leading to increased bandwidth. For 5–50 g high dynamic optimization of the bandwidth, the improvement is less than 50%.

For the auxiliary third-order loop, the improvement of the optimal bandwidth of tactical grade MEMS IMU devices is more than 90% in the dynamic scenarios of 0.5–20 g/s. The improvement effect of commercial grade MEMS IMU devices on

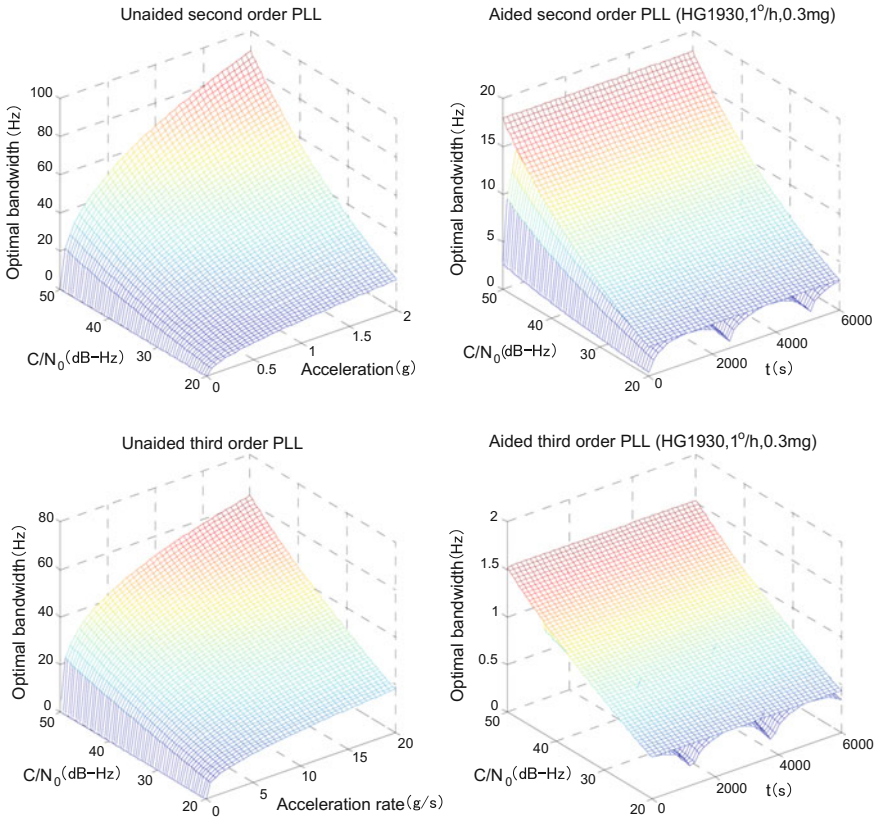


Fig. 6 MEMS IMU aided PLL bandwidth

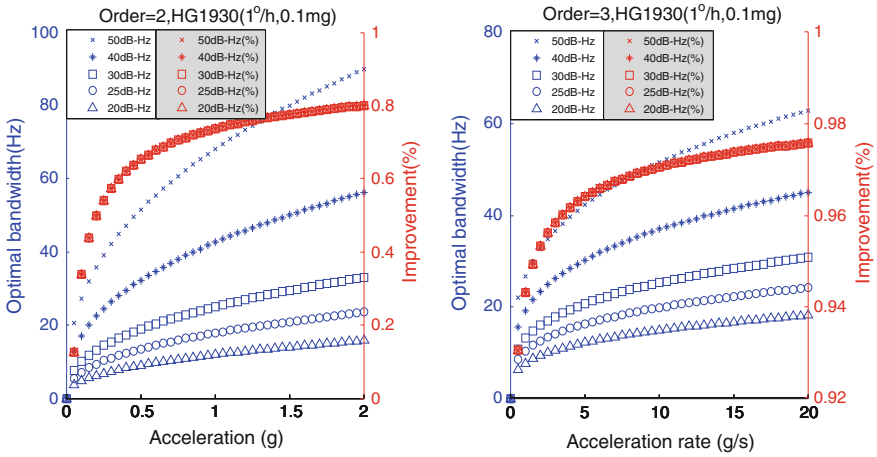


Fig. 7 Contrast of HG1930 aided PLL bandwidth

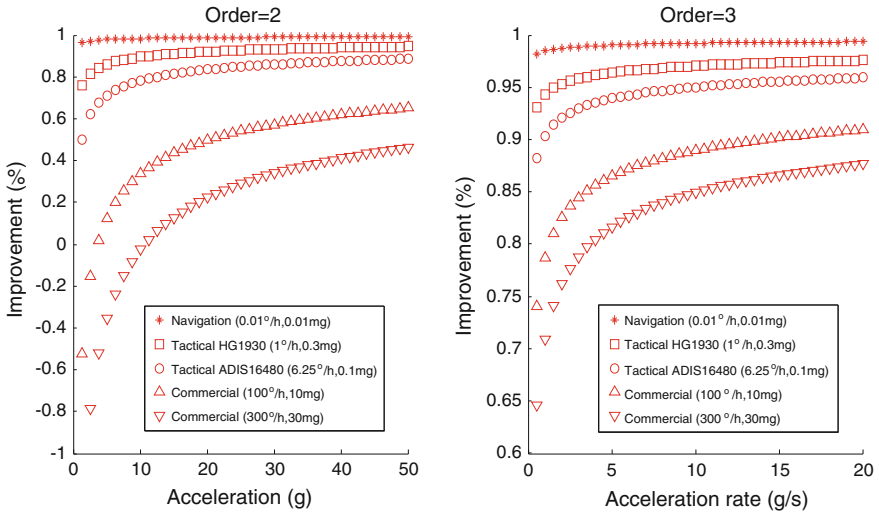


Fig. 8 Contrast of different accuracy grade MEMS IMU aided PLL bandwidth

the optimal bandwidth is more than 60%. When the carrier dynamic is less than 0.1 g/s, commercial grade MEMS IMU devices have less than 60% improvement in the optimal bandwidth. If the coupling coefficient is not 1, the carrier dynamic will affect the optimal bandwidth of the loop with the auxiliary. The improvement ratio of the loop bandwidth will decrease with the increase in the carrier dynamic, which is in accordance with the previous researches.

5 Conclusion

In this paper, the relationship between steady-state tracking error and the combination of C/N_0 and accuracy of the MEMS IMU device is analyzed quantitatively based on the MEMS IMU auxiliary PLL. The theoretical analysis and simulation results show that commercial MEMS IMU devices can effectively assist the third-order PLL for 30 dB-Hz signal. Tactical grade MEMS IMU devices can effectively assist the second-order PLL when the bandwidth is greater than 2 Hz.

The optimal bandwidth with auxiliary is analyzed based on the conventional PLL, which is related to the C/N_0 of signal and the precision of MEMS IMU devices. The improvement of the loop bandwidth is not in dependent of C/N_0. The optimal bandwidth will decrease obviously with the increase in C/N_0. The improvement of bandwidth will be better when the accuracy of the MEMS IMU devices is higher. The auxiliary effect will be better with a higher order.

References

1. Yang Y, Li J, Wang A et al (2014) Study on the basic navigation and positioning performance of the BeiDou regional satellite navigation system. Earth Sci 44(1):72–81
2. Schwarz KP, EI-Sheimy N (2007) Digital mobile mapping systems state of the art and future trends. Taylor and Francis Group
3. Zhang T (2010) GPS/SINS ultra compact integrated navigation system of key technology research. Harbin Engineering University, Harbin
4. Osiander R, Darrin MG, Champion JL (2006) MEMS and microstructures in aerospace application. CRC Press
5. Kreye C, Eissfeller B, Winkel JQ (2000) Improvement of GNSS receiver performance using deeply coupled INS measurements. ION GPS 2000, Salt Lake City, UT, September 2000, pp 844–854
6. Kaplan ED, Hegarty CJ (2006) Understanding GPS: principle and applications, 2nd edn. Artech House, Boston

An Improved Robust Fading Filtering Algorithm for the GPS/INS Integrated Navigation

Chen Jiang, Shu-bi Zhang and Qiu-zhao Zhang

Abstract As an estimator of the state space, Kalman filter has been widely applied in the field of dynamic navigation and positioning. However, the divergence will be likely to happen when significant errors of the model exist. Thus, the fading factor is introduced to control the influences of the state model errors. In order to improve the performance of the filter, the multiple fading factors are adopted to address the problem that a single fading factor fails to control the interferences of all model errors. By minimizing the estimation error in the worst case, the H-infinity filter can be adopted to address the uncertainties under different conditions. Nevertheless, the H-infinity filter cannot resist the influences of outliers. The robust estimation method is thus integrated with the H-infinity filtering algorithm to improve the stability of the filter furtherly. Data of the Global Positioning System (GPS) and the Inertial Navigation System (INS) integrated navigation system are collected with GPS receivers and Inertial Measurement Units (IMU) under actual conditions. Experiments using different filtering algorithms together with the contrastive analysis are performed with the collected data. Results demonstrate that the proposed filtering algorithm shows better stability. Both the filter divergence and the influences of the outliers are controlled effectively with the proposed filtering algorithm, and precision of the filtering results are improved simultaneously.

Keywords Fading filter · H-infinity filter · Integrated navigation · Robust estimation

C. Jiang · S. Zhang · Q. Zhang (✉)
School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China
e-mail: qiuzhao.zhang@cumt.edu.cn

C. Jiang · S. Zhang · Q. Zhang
Collaborative Innovation Center for Resource Utilization and Ecological Restoration of Old Industrial Base, China University of Mining and Technology, Xuzhou 221116, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_30

1 Introduction

The required accuracy of the navigation instrument increases along with the development of the navigation techniques. It has become an important aspect by integrating the navigation techniques to improve the instruments' performance. GPS is characterized by all-whether measurement and high precision, and INS is a self-navigation system which indicates that no signals of the INS is sent and received. The integration of GPS and INS has been widely studied and applied. As an optimal estimation method, the Kalman filter has found extremely wide applications. Kalman filter is proposed based on the H_2 norm, and the performance is closely related to the accuracy of the system model and the statistical information of the noises. It may cause a filter divergence when significant errors of the model exist [1]. Based on the fading memory, an optimal fading filter was proposed to restrain the filter divergence [2]. However, the multiple-dimension system was controlled with only a single fading factor, which indicates that it is difficult to correctly present the change of the system and to guarantee the absolute optimization of the filter. Then, the suboptimal fading filter [3] and the multiple fading filter [4] were proposed based on the conventional fading factor. In theory, the fading filter will show the ability to control the influences of the model deviations. However, the aforementioned algorithms' ability of resisting the outliers should be furtherly improved, and the outliers may even cause a filter divergence. By minimizing the estimation errors in the worst case [5, 6], the H-infinity filter can be adopted to address the system model and uncertain interferences of the noises, and it has been adopted in the GPS/INS integrated navigation system [7, 8]. Nevertheless, it may cause a filter divergence in the presence of outliers [9]. The robust estimation method is a practical method of resisting the influences of abnormal observations [10–12]. Rather than disproportionately pursuing validity and unbiasedness, the robust estimation method focuses on the robustness and the reliability of the estimator. Accordingly, the robust estimation method is introduced to improve the performance of the filter.

In this paper, an improved multiple fading filtering algorithm is proposed. The data were collected under natural conditions with the GPS/INS integrated navigation instruments. The data were calculated with the proposed algorithm, and validity of the improved algorithm were demonstrated with various experiments. Contrastive analysis was performed between the improved algorithms and the other algorithms.

2 The Fading Filter and the H-infinity Filter

2.1 The Fading Filter

For the dynamic model equation and the observation equation

$$\begin{cases} x_k = \Phi_{k,k-1}x_{k-1} + w_k \\ z_k = H_k x_k + v_k \end{cases} \tag{2.1}$$

where x_k is the state vector at epoch k, $\Phi_{k,k-1}$ is the state transition matrix, H_k is the measurement matrix, z_k is the observation vector, w_k and v_k are the system and the observation noises, respectively. The fading filter approach is given by

$$x_{k/k} = x_{k/k-1} + \bar{K}_k(z_k - H_k x_{k/k-1}), \tag{2.2}$$

where $\bar{K}_k$ is the gain matrix and $\bar{K}_k = \bar{P}_{k/k-1}H_k^{\mathrm{T}}(H_k\bar{P}_{k/k-1}H_k^{\mathrm{T}} + R_k)^{-1}$, $\bar{P}_{k/k-1}$ is the covariance matrix of the state vector and $\bar{P}_{k/k-1} = S_k(\Phi_{k,k-1}P_{k-1}\Phi_{k,k-1}^{\mathrm{T}}) + Q_k$, S_k is the fading factor and $S_k \geq 1$, P_{k-1} is the covariance matrix of the state vector at epoch $k-1$, R_k and Q_k are the covariance matrices of the observation and the state noises, respectively. Obviously, in the fading filter, the covariance matrix of the a priori state is inflated for S_k times, and this will degrade the efficiency of the past state information. Compared with the standard Kalman filter, the state errors from the last epoch are well controlled [2].

It is clear that a proper fading factor should be selected. A fading filter was proposed based on the optimization where two forms of fading factors were derived [2], and the concise one is:

$$S_k = \max\{1, \frac{1}{n}tr(N_k M_k^{-1})\}, \tag{2.3}$$

where $M_k = H_k\Phi_{k,k-1}P_{k-1}\Phi_{k,k-1}^{\mathrm{T}}H_k^{\mathrm{T}}, N_k = P_{V_k} - H_kQ_kH_k^{\mathrm{T}} - R_k, P_{V_k} = E(V_kV_k^{\mathrm{T}})$ and $V_k = H_kx_{k/k-1} - z_k, tr(\cdot)$ denotes taking trace of a matrix, V_k denotes the predicted residual vector, and P_{V_k} denotes the covariance matrix of V_k, and $\hat{P}_{V_k} = \frac{1}{k}\sum_{i=1}^{k} V_i V_i^{\mathrm{T}}$. A simplified method of S_k is $S_k = \max\{1, tr(N_k)/tr(M_k)\}$, this fading factor S_k is the optimal in theory [1]. Since a single fading factor can hardly guarantee the optimization, a multiple fading filter was proposed [3], and the fading factor is:

$$S_k = diag(s_1, s_2, \ldots, s_m, 1, \ldots, 1, \ldots, 1), \tag{2.4}$$

$$S_i = \max\left(1, \sqrt{\frac{[v_i(k)]^2}{\lambda_i^2 j_{ii}(k)\varepsilon_i} - \frac{b_{ii}(k)}{j_{ii}(k)}}\right), \quad (i = 1, 2, \ldots, m), \tag{2.5}$$

where $v_i(k)$ is the ith element of V_k, m is the number of the observable variables, λ_i is the ith observable element of H_k, $b_{ii}(k)$ is the ith element of B_k and $B_k = H_k Q_{k-1} H_k^{\mathrm{T}} + R_k$, $j_{ii}(k)$ is the ith element of and $J_k = \Phi_k P_{k-1} \Phi_k^{\mathrm{T}}$, and ε_i is the threshold value fixed according to the Chi-square distribution.

2.2 The H-infinity Filter

Assume that a discrete nonlinear system is given by

$$\begin{cases} x_k = f(x_{k-1}) + w_{k-1} \\ z_k = h(x_k) + v_k \end{cases}, \tag{2.6}$$

where $f(\cdot)$ and $h(\cdot)$ are the nonlinear state transition function and the nonlinear observation function, respectively, x_k is the state vector to be estimated. The cost function J is [13]:

$$J = \frac{\sum_{k=1}^{N} \|x_k - \hat{x}_k\|^2}{\|x_0 - \hat{x}_0\|_{P_0^{-1}}^2 + \sum_{k=1}^{N} \left(\|w_k\|_{Q_k^{-1}}^2 + \|v_k\|_{R_k^{-1}}^2\right)}, \tag{2.7}$$

where N is the total number of filtering time limit, x_0 is the initial value of x with the covariance matrix $P_0, \hat{x}_0$ and $\hat{x}_k$ are the estimated state vectors of x_0 and x_k, respectively.

The estimate of x_k should satisfy $\hat{x}_k = \arg\min \|J\|_\infty$. However, the closed-form approach of an optimal H-infinity filter cannot be achieved under common conditions [14]. Then a suboptimal recursion algorithm was proposed with a threshold value γ which satisfied the Riccati inequality [15]

$$P_k^{-1} + H_k^{\mathrm{T}} H_k - \gamma^2 L_k^{\mathrm{T}} L_k > 0, \tag{2.8}$$

where P_k is the covariance matrix of x_k, and L_k is an unit matrix in this paper. As for the nonlinear system, the recursion formula of $P_{k/k}$ in the H-infinity filter is given by:

$$P_{k/k} = P_{k/k-1} - [P_{xz,k}\, P_{k/k-1}] \begin{bmatrix} P_{zz,k} - R_k + I & P_{xz,k}^{\mathrm{T}} \\ P_{xz,k} & P_{k/k-1} - \gamma^2 I \end{bmatrix}^{-1} \begin{bmatrix} P_{xz,k}^{\mathrm{T}} \\ P_{k/k-1}^{\mathrm{T}} \end{bmatrix}, \tag{2.9}$$

where $P_{xz,k}$ and $P_{zz,k}$ denote the cross-covariance matrix and the innovation covariance matrix, respectively, $P_{xz,k} = P_{k-1/k} H_k^{\mathrm{T}}$ and $P_{zz,k} = H_k P_{k/k-1} H_k^{\mathrm{T}} + R_k$. In general, $P_{xz,k}$ and $P_{zz,k}$ are obtained using a nonlinear filtering algorithm such as the unscented Kalman filter and the cubature Kalman filter. In this paper, the nonlinear problem is addressed with the cubature Kalman filter.

It should be noticed that, the value of γ is closely related to the robustness of the filter, and the filter becomes more robust when γ decreases. However, it may cause a filter divergence if γ is too close to 0 [16]. Thus, a proper value of γ is often set by experience.

3 An Improved Robust Multiple Fading Filter

Since a single fading factor cannot guarantee the optimal filtering, a multiple fading factor was proposed to address this problem [3]. However, the robustness of this filter can still be improved. Accordingly, a new comprehensive filtering algorithm is proposed to furtherly improve the performance of conventional fading filter. In a multiple fading filter, $S_k = \mathrm{diag}(s_1, \ldots, s_m, 1, \ldots, 1)$, where m is the number of observable elements in x_k. $P_{k/k-1}$ can be expressed as $P_{k/k-1} = S_k \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^{\mathrm{T}} S_k + Q_k$ to keep its symmetry.

The robust estimation method is employed to improve the robustness to the abnormal observations. In general, the components of the predicted state vector are correlated, so the equivalent covariance matrix constructed by scaling factor λ_{ij} is adopted, and λ_{ij} is [17]:

$$\lambda_{ij} = \sqrt{\lambda_{ii}}\sqrt{\lambda_{jj}}, \tag{3.1}$$

$$\lambda_{ii} = \begin{cases} 1 & \left|\bar{V}_{\bar{X}_{k_i}}\right| \le c \\ \dfrac{\left|\bar{V}_{\bar{X}_{k_i}}\right|}{c} & \left|\bar{V}_{\bar{X}_{k_i}}\right| > c \end{cases}, \tag{3.2}$$

where $\left|\bar{V}_{\bar{X}_{k_i}}\right|$ is the component of the predicted residual vector, $1.0 \le c \le 1.5$, and λ_{jj} can be fixed in the same way as λ_{ii}. Thus, the equivalent covariance matrix $\bar{R}_k$ is achieved by

$$\bar{R}_k = \lambda_{ij} R_k. \tag{3.3}$$

Then, a robust multiple fading H-infinity filter is constructed, and the whole recursion formulas of the new algorithm are:

$$x_{k/k-1} = \Phi_{k/k-1} x_{k-1}, \tag{3.4}$$

$$\bar{P}_{k/k-1} = S_k \Phi_{k,k-1} P_{k-1/k-1} \Phi_{k,k-1}^{\mathrm{T}} S_k^{\mathrm{T}} + Q_k, \tag{3.5}$$

$$x_{k/k} = x_{k/k-1} + \bar{K}_k(z_k - H_k x_{k/k-1}), \tag{3.6}$$

$$\bar{K}_k = \bar{P}_{k/k-1} H_k^{\mathrm{T}} (H_k \bar{P}_{k/k-1} H_k^{\mathrm{T}} + \bar{R}_k)^{-1}, \tag{3.7}$$

$$\bar{P}_{k/k} = \bar{P}_{k/k-1} - \Phi_{k/k-1} \bar{P}_{k/k-1} [H_k^{\mathrm{T}} L_k^{\mathrm{T}}] \bar{R}_{e,k}^{-1} \begin{bmatrix} H_k \\ L_k \end{bmatrix} \bar{P}_{k/k-1} \Phi_{k/k-1}^{\mathrm{T}}, \tag{3.8}$$

$$\bar{R}_{e,k}^{-1} = \begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix} + \begin{bmatrix} H_k \\ L_k \end{bmatrix} \bar{P}_{k/k-1} [H_k^{\mathrm{T}} \quad L_k^{\mathrm{T}}], \tag{3.9}$$

4 The GPS/INS Integrated Navigation System

In this GPS/INS integrated navigation system, the loosely-coupled navigation is easier to be implemented and the computation process is more concise [18]. In the loosely-coupled GPS/INS integrated navigation system, a 15-dimension state vector is designed, and the state vector $\hat{X}$ is given by:

$$\hat{X} = [\delta x, \delta y, \delta z, \delta v_x, \delta v_y, \delta v_z, \delta \phi_e, \delta \phi_n, \delta \phi_u, \delta g_x, \delta g_y, \delta g_z, \delta a_x, \delta a_y, \delta a_z], \tag{4.1}$$

where the symbols denote the deviations of the position, velocities, attitudes, and the noises of the gyroscope and accelerometer, respectively.

In the loosely-coupled GPS/INS integrated navigation system, ation form of the state equation is achieved by Taylor series expansion [1]:

$$X_k = \Phi_{k,k-1} X_{k-1} + W_k, \tag{4.2}$$

where X_k is the state vector, W_k is the model noise, and $\Phi_{k,k-1}$ is the discretized state transition matrix. The differences of position and velocity between GPS and INS are selected as the external measurement:

$$Z_\rho(t) = \rho_{\mathrm{GPS}} - \rho_{\mathrm{INS}}. \tag{4.3}$$

r_{GPS} and v_{GPS} are the output information of the GPS, r_{INS} and v_{INS} are that of the INS. The measurement equation is given by:

$$Z_k = \begin{bmatrix} r_{\mathrm{GPS}} - r_{\mathrm{INS}} \\ v_{\mathrm{GPS}} - v_{\mathrm{INS}} \end{bmatrix}, \tag{4.4}$$

5 Experiments and Analysis

In these experiments, data were collected by a vehicle mounted GPS/INS integrated navigation system. In this system, two GPS receivers and a low cost inertial measurement unit (IMU) were adopted. GPS data were calculated through the double difference pseudorange. The precise results calculated by the double difference carrier phase were taken as references. Then two cases were designed to examine the stability and the robustness of the new algorithm. Four schemes were performed in each case:

Scheme 1: Kalman Filter (KF);
Scheme 2: H-infinity Filter (HF) (γ was set as 2);
Scheme 3: The multiple fading H-infinity filter (MF) (γ was set as 2);
Scheme 4: The robust multiple fading H-infinity filter (RMF) (γ was set as 2);

5.1 Case 1

Each filtering algorithm was calculated by self-designed program. Position errors of the four schemes are demonstrated in Figs. 1, 2, 3 and 4.

The RMSEs of each scheme are displayed in Table 1:

It is learned from Figs. 1 and 2, the HF algorithm performs much better than the KF algorithm. With the model deviations and the uncertain interferences controlled, the multiple fading filter shows a better performance than the HF algorithm which

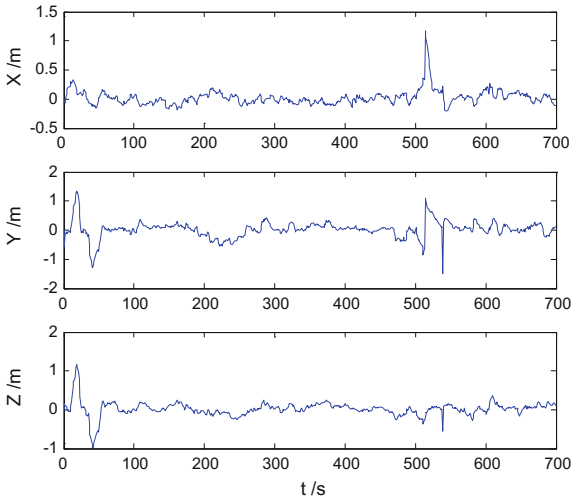


Fig. 1 Position errors of the KF algorithm

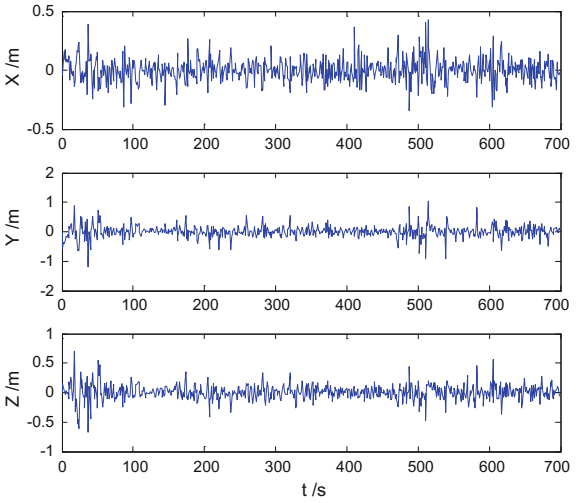


Fig. 2 Position errors of the HF algorithm

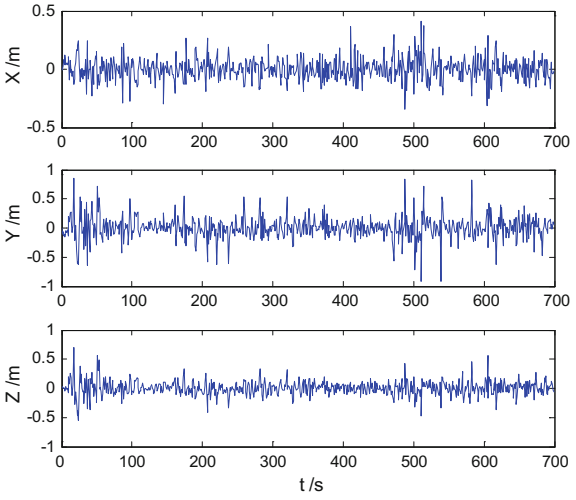


Fig. 3 Position errors of the MF algorithm

can be seen from Figs. 2 and 3. Aided by the robust estimation method, the RMF algorithm performs the best among the four schemes, and the amplitudes of the position errors are furtherly reduced which can be concluded from Fig. 4. It is demonstrated in Table 1 that the improved algorithm shows a higher precision comparing with the other algorithms.

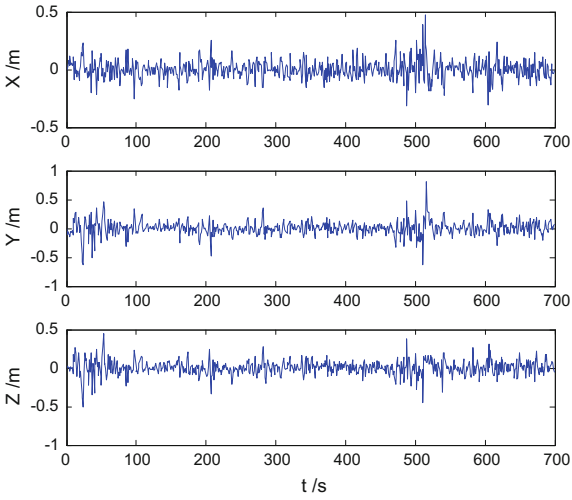


Fig. 4 Position errors of the RMF algorithm

Table 1 RMSEs of each scheme (m)

Axis	KF	HF	MF	RMF
X	0.132	0.098	0.094	0.080
Y	0.291	0.208	0.192	0.124
Z	0.192	0.131	0.124	0.096

5.2 Case 2

In this case, gross errors were added artificially to the GPS initial observations at the 160th, 260th, 360th, 460th, 560th epochs, respectively. Position errors of these schemes are displayed in Figs. 5, 6, 7 and 8.

In this case, performances of these schemes are mainly affected by outliers. Obviously, Figs. 5 and 6 demonstrate that the KF and HF algorithms have little ability to resist the effects of outliers. However, both the conventional and the robust fading filters show a great robustness to outliers which is learned from Figs. 7 and 8. Moreover, error amplitudes of the robust multiple fading filter are smaller than those of the other algorithms.

Then the RMSEs of each scheme are listed in Table 2.

Since the performances of the KF and the HF algorithms are significantly affected, the RMSEs of them become much bigger. Little change is found for the MF and the RMF algorithms in terms of the RMSEs which indicates that they can resist the influences of the outliers. With influences of the dynamic model errors and the uncertain interferences controlled, the new fading filter is more robust than the other three algorithms.

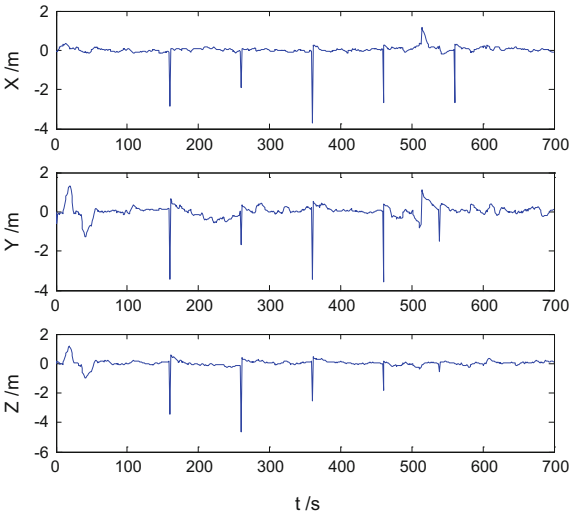


Fig. 5 Position errors of the KF algorithm

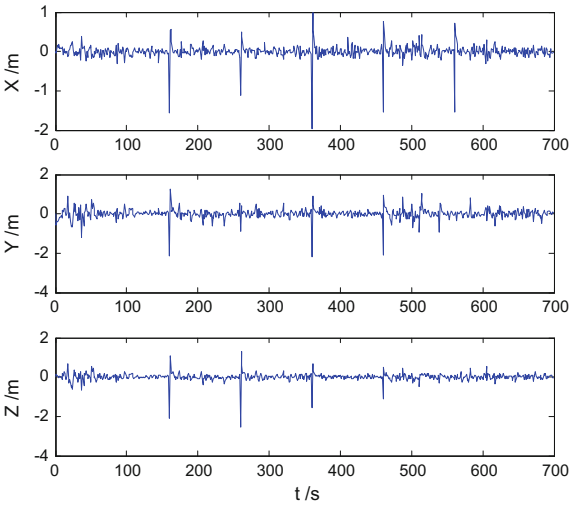


Fig. 6 Position errors of the HF algorithm

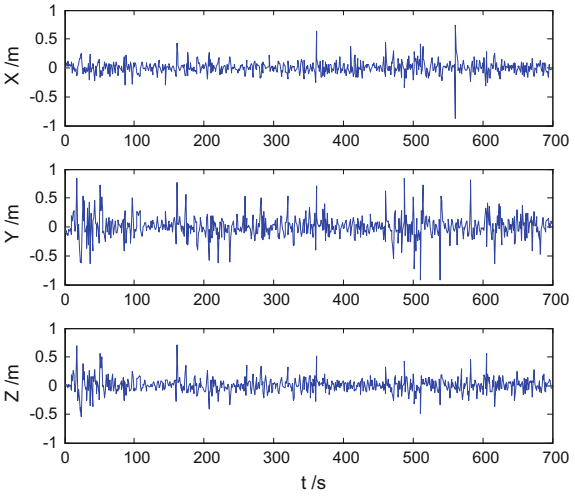


Fig. 7 Position errors of the MF algorithm

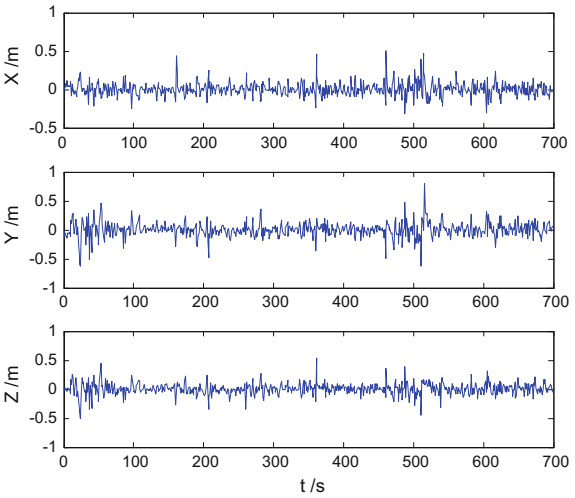


Fig. 8 Position errors of the RMF algorithm

Table 2 RMSEs of each scheme (m)

Axis	KF	HF	MF	RMF
X	0.275	0.182	0.111	0.088
Y	0.383	0.264	0.198	0.127
Z	0.321	0.210	0.130	0.101

6 Conclusion

In this study, an improved robust multiple fading filter is proposed. Aided by the robust estimation method, a comprehensive filtering algorithm is constructed by integrating the multiple fading factor and the H-infinity filter. The improved filtering algorithm is performed in the loosely-coupled GPS/INS integrated navigation system and the detail conclusions are summarized as follows:

(1) Both the Kalman filter and the H-infinity filter will be affected significantly by the outliers. With the model deviations controlled, the multiple fading filter shows some robustness.
(2) Aided by the robust estimation method and the H-infinity filter, performance of the conventional multiple fading filter is improved furtherly. The robust multiple fading filter is implemented in the loosely-coupled GPS/INS integrated navigation system, and the stability and robustness are demonstrated with both the initial and the perturbative data.

Acknowledgements The authors are grateful for the support of the National Natural Science Foundation of China (No. 41504032), the Natural Science Foundation of Jiangsu Province (No. BK20150175) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

1. Yang YX (2006) Adaptive navigation and dynamic positioning. Surveying and Mapping Press
2. Xia QJ, Sun YX, Zhou CH (1990) An optimal adaptive algorithm for fading Kalman filter and its application. Acta Automatica Sinica 16(3):210–216
3. Geng YR, Wang JL (2008) Adaptive estimation of multiple fading factors in Kalman filter for navigation applications. GPS Solutions 12(4):273–279
4. Qian HM, Ge L, Peng Y (2012) Multiple fading factors Kalman filter and its application in SINS initial alignment. J Chin Inertial Technol 20(3):287–291
5. De Souza CE, Palhares RM, Peres PD (2001) Robust H_∞ filter design for uncertain linear systems with multiple time-varying state delays. IEEE Trans Sig Proc 49(3):569–576
6. Nagpal KM, Khargonekar PP (1991) Filtering and smoothing in an H_∞ setting. IEEE Trans Autom Control 36(2):152–166
7. Chen YR, Yuan JP (2009) An improved robust H_∞ multiple fading fault-tolerant algorithm for INS/GPS integrated navigation. J Astronaut 30(3):930–936
8. Zhang QZ, Meng XL, Zhang SB et al (2015) Singular value decomposition-based robust cubature Kalman filtering for an integrated GPS/SINS navigation system. J Navig 68(3): 549–562
9. Zhang QZ, Stephenson S, Meng XL et al (2016) A new robust filtering for a GPS/INS loosely coupled integration system. Surv Rev 48(348):181–187
10. Chang GB (2014) Robust Kalman filtering based on Mahalanobis distance as outlier judging criterion. J Geodesy 88(4):391–401
11. Yang YX (1999) Robust estimation of geodetic datum transformation. J Geodesy 73(5): 268–274

12. Zhou JW (1989) Classical theory of errors and robust estimation. Acta Geodaetica Cartogr Sin 18(2):115–120
13. Simon D (2006) Optimal state estimation: Kalman, H_∞ and nonlinear approaches. Wiley, New Jersey
14. Yue XK, Yuan JP (2001) H-infinity filtering algorithm and its application in GPS/SINS integrated navigation system. Acta Aeron Autica et Astronautica Sinica 22(4):366–368
15. Hassibi C, Sayed AH, Kailath T (1996) Linear estimation in Krein spaces-part II: applications. IEEE Trans Autom Control 41(1):34–49
16. Fagin SL (1964) Recursive linear regression theory: optimal filter theory and error analysis. In: IEEE international convention record, vol 12 (Part 1), pp 216–245
17. Yang YX, Song LJ, Xu TH (2002) Robust parameter estimation for geodetic correlated observations. Acta Geodaetica Cartogr Sin 31(2):95–99
18. Gao WG, Yang YX, Zhang SC (2006) Adaptive robust Kalman filtering based on the current statistical model. Acta Geodaetica Cartogr Sin 35(1):15–18

Rapid Extrinsic Calibration of Seamless Multi-sensor Navigation System Based on Laser Scanning

Yanglin Zhou, Guangyun Li, Fengyang Li, Ming Dong and Shuaixin Li

Abstract With the rapid development of national economy, higher requests have been put forward on PNT (positioning navigation and timing) service. Researches on integrated navigation systems, including fusion technologies of multi-navigation sources and adaptive algorithms for various application scenarios, have become the highlighted topic. The method of laboratory calibration for GNSS/INS/LiDAR integrated navigation system, applied for indoor and outdoor seamless positioning, was studied in this paper. Combined with high-precision Theodolite and Laser Scanning System, this paper proposed a laboratory high-accuracy calibration scenario for integrated navigation system, designed a rapid laboratory calibration model, deduced the corresponding error model with the analysis of differential equations and the method of Monte Carlo, and finished the lever-arm calibration of GNSS and the mounting parameters of laser scanner, and designed a rapid laboratory calibration model. Laboratory experiments had been carried out to verify the feasibility of this millimeter-level calibration method, which could be used for calibrating the initial mounting parameters of the integrated navigation system.

Keywords Indoor and outdoor seamless positioning · Laboratory calibration model · GNSS lever-arm · LS mounting parameters

Y. Zhou (✉) · G. Li · F. Li · S. Li
School of Navigation and Aerospace Engineering,
Information Engineering University, Zhengzhou, China
e-mail: zhouyanglin@126.com

G. Li
e-mail: guangyun_li@sina.com

F. Li
e-mail: freyr_li@outlook.com

S. Li
e-mail: lsx_navigation@sina.com

M. Dong
China Satellite Navigation Engineering Center, Beijing, China
e-mail: 597865268@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_31

Nomenclature

$\overrightarrow{r_{a_i}}$	position vector of point i in coordinate a
$\overrightarrow{r^a_{P_i}}$	position vector of point P_i in coordinate a
$\overrightarrow{r^b_a}$	relative translation from coordinate a to coordinate b
R^b_a	relative rotation from coordinate a to coordinate b
P_{GNSS}	the geometric center of GNSS antenna
P_{ANT}	the geometric center of the lower edge of GNSS antenna
L	laser scanner coordinate system
S	auxiliary measurement coordinate system
I	POS coordinate system
k_1	scale change between coordinate S to coordinate L
k_2	scale change between coordinate I to coordinate S

1 Introduction

With the rapid development and wide popularization of the digital information, the concept of 'Digital earth', 'smart earth' or 'smart city' [1] has been reloaded and becomes common knowledge, and thus higher requests have been put forward on PNT (positioning navigation and timing) service's quantity, quality and renew speed [2]. Traditional navigation techniques could no longer satisfy the increasing demand for LBS (location based service), Integrated with different navigation sensors (e.g. GNSS, Accelerometer, Gyroscope, Magnetometer, 2-D/3-D imager, LiDAR, Bluetooth, Barometer, WLAN etc.), adapted to complex scenarios (e.g. urban canyon, building interior, open field, jungle, forest canopy, suburban, underground, underwater etc.), high-efficiency and robust seamless indoor and outdoor navigation system quickly becomes the new trend of acquiring the real-time geometric information [3].

Seamless navigation and positioning is in fact a new navigation fusion technology, which will support integration of combination of existing navigation techniques [GNSS, WNL (wireless network locating) and SCS (self-contained sensor)], capable of heterogeneous and asynchronous navigation sources and have a rapid adaptation to different scenarios for the creation of robust navigation and positioning technology.

Integrating GNSS, wireless locating with self-contained sensors requires new navigation filtering (i.e., fusion) algorithms, abstraction methods, and an overall navigation system architecture [4], so that it attracts extensive attentions and intensive investigation: SoOPs (signals of opportunity) based navigation technology [5], high-precision wireless network locating technologies integrated system [6], multi-sensor, multi-network integrated navigation technology [7], Camera/IMU/GNSS integrated navigation technology for challenging GNSS

environments [8], vision-aided Pedestrian/GNSS coupled navigation technology for smart terminals [9], etc.

Key technologies of Seamless navigation and positioning can be divided into four parts: (1) multi-sensor calibration; (2) Information synchronization and fusion; (3) Online optimization of navigation sensors; (4) Robust and economical filtering algorithm. This paper focus on optimizing the extrinsic calibration procedure of multi-sensor navigation system. Accurate inter-sensor geometric relationship is of great importance to information synchronization and fusion, which will influence navigation and positioning directly [10]. In the current study, Users depend greatly on the manufacturer, whose calibration procedure is complex and not revealed, lacking of autonomy and innovation. In order to solving these issues, this paper presents an extrinsic calibration approach, utilizing laser scanning based aiding, for: GNSS/INS/LiDAR integrated seamless indoor and outdoor navigation and positioning system (Indoor: point cloud based registration [11] or SLAM [12] (Simultaneous Localization and Mapping; Outdoor: integrated GNSS/INS navigation [13]). This paper tackles a rapid and robust calibration scheme, whose performance is evaluated through experimental results using simulated model. Besides, a comparative analysis between the theodolite measuring system based and laser scanning based results has been carried out to verify the feasibility and efficiency of this calibration method.

2 Laboratory Calibration Scheme

An operational and low-cost seamless navigation system (GIL) is integrated on a wheel-based platform, which consists of position and orientation system (SPAN-CPT [14]), laser sensor (FARO focus3D-X120 [15]) and data processing system (Laptop Think pad T530), as shown in Fig. 1

2.1 Introduction of Measuring Sensors

(1) Laser scanner system
Laser scanner (LS) is the main measuring sensor in GIL, which is current state-of-the-art active imaging modality, capable of measuring detailed dense geometric information of the object space efficiently and quickly. A FARO focus3D-X120 laser scanner is used as a navigation sensor in GIL and a Rigel VZ400 laser scanner [16] is used as a secondly measuring sensor in the calibration procedure. Both of the introduced laser sensors are time-of-flight system, but using different distance measurement principle [17] (pulse-based focus3D-X120 and phase-based VZ400): FARO laser scanner cannot identify the retro-reflective material (RRT) targets and Rigel laser scanner cannot

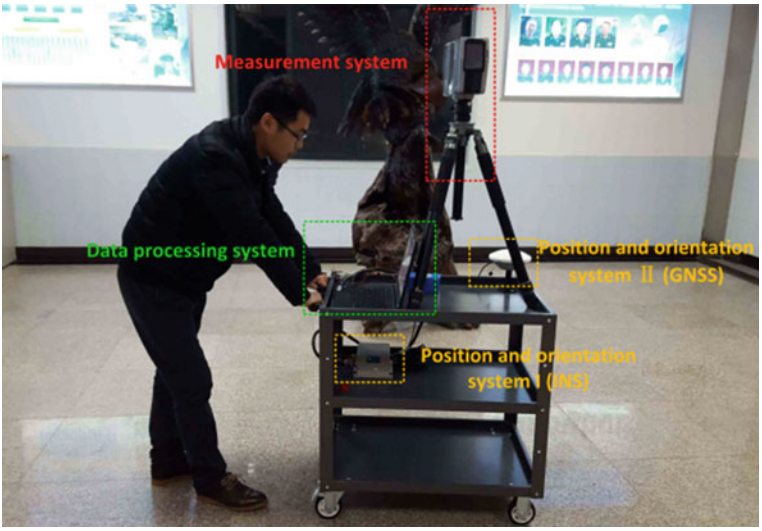
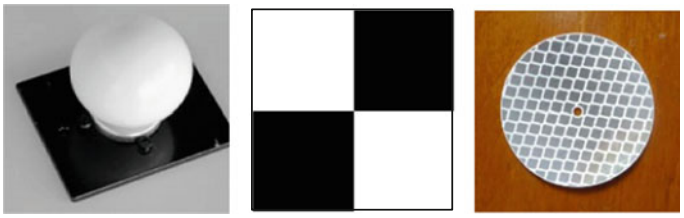


Fig. 1 Sensor configuration of GIL



(a) FARO Spherical target (b) FARO Chessboard target (c) Rigel Retro-reflect target

Fig. 2 Cooperation targets for laboratory calibration experiment

identify the spherical target automatically. The cooperative targets used in this paper are shown in Fig. 2. The measurement specifications of laser scanners are shown in Table 1.

(2) Theodolite measuring system

The theodolite measuring system (TMS), carrying out frontage intersecting measurement according to the fixed reference line, consists of three T3000A theodolites in this paper, with a throughout consideration of precision, reliability, efficiency. The intersection measuring accuracy of the integrated TMS can attain 0.02–0.1 mm within the range of 10 m [18].

When using LS and TMS as an auxiliary measurement system, its accuracy and spatial data type has a huge difference due to different measuring principle and corporative targets. As shown in Fig. 3, TMS attains only singe point spatial

Table 1 Measurement specifications of different measurement systems

	Name	Range	Accuracy
LS I	RigeL	1–500 m	5 mm @ 100 m
	VZ 400		
LS II	FARO	0.6–120 m	2 mm @ 50 m
	Focus3DX120		
TMS	Leica	1–100 m	0.05 mm
	TM3000		

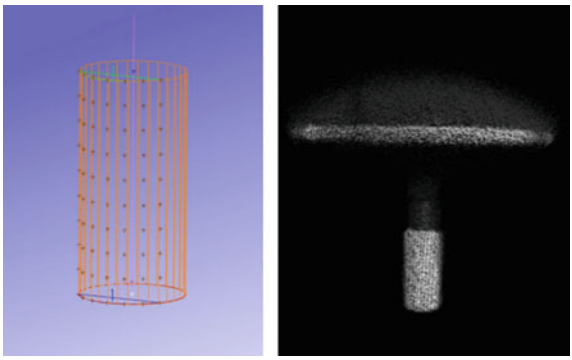


Fig. 3 Cooperation targets for laboratory calibration experiment

(a) 3D points measured by TMS (b) 3D Point cloud captured by TLS

information while LS attains points cloud spatial information as well as retro-reflective information. So the measurement data must be preprocessed before the application into the calibration procedure.

2.2 *Details of Calibration Procedure*

The total extrinsic calibration of GIL is complex, which consists of subordinate sensor calibration and inter-sensor relationship calibration. Subordinate sensor calibration mainly focuses on the improvement of survey-grade sensors (e.g. IMU, laser scanner and camera) with the corresponding corrections (e.g. IMU drift error correction [19], laser scanner ranging constant correction [20] and camera orientation element correction [21]). Inter-sensor calibration mainly discusses about the extrinsic relationship between GNSS antenna [22], IMU [23] and measuring sensors (including laser scanner [24] and camera [25]).

GNSS lever arm is the difference between the different sensing point of GNSS and IMU. The lever arm calibration is necessary in GIL, because its error will influence the accuracy of navigation solution directly [26]. It is assumed in this

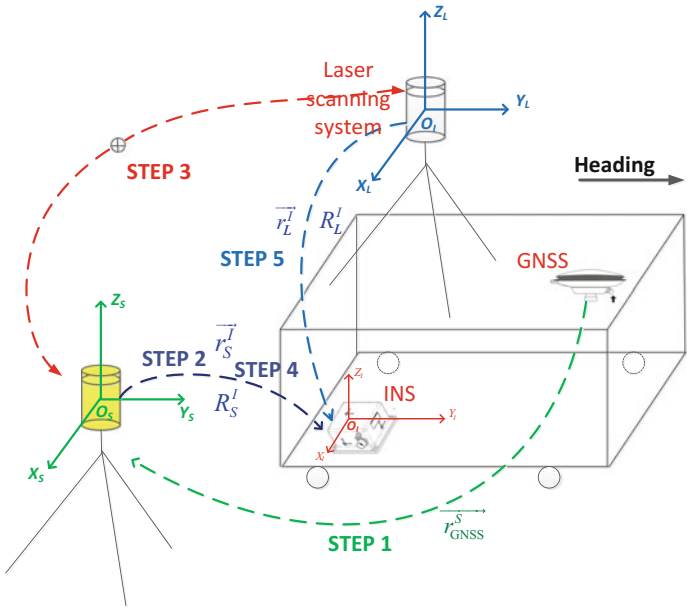


Fig. 4 Procedure of laboratory calibration

paper that the center of POS is in coincident with that of IMU. The definition of the relevant coordinate system can be found in [27], the whole calibration procedure is shown in Fig. 4.

The extrinsic calibration in this paper consists of GNSS lever arm calibration and laser sensor mounting parameters calibration, with the following specific scheme:

(1) GNSS lever arm calibration.

Step 1: Attain the geometric center of GNSS antenna through AMS.
Step 2: Transfer the geometric center into POS coordinate, with the connection between POS coordinate and AMS coordinate established from common points between POS and AMS.

(2) LS mounting parameters calibration.

Step 1: Establish the coordinate connection between LS and AMS.
Step 2: Similar to Step 2 of GNSS lever arm calibration, establish the coordinate connection between POS and AMS.
Step 3: Calculate the mounting parameters of LS through the coordinate connection between LS and POS, deduced from Steps 1 and 3.

3 Laboratory Extrinsic Calibration Mathematical Model

3.1 GNSS Lever Arm Calibration Model

As shown in Fig. 5, lever arm calibration model can be divided into two parts: (1) Circle fitting of GNSS geometric center; (2) coordinate transformation from AMS to POS.

1. Circle fitting of GNSS geometric center
   Based on the fact that the edge of GNSS antenna is a regular shape, the geometric center of GNSS can be deduced with the axial information from circle fitting.

(1) Least square based circle fitting

Assume the radio of antenna is R, and its center is $\overrightarrow{r^S_{P_{ANT}}} = (x_0, y_0, z_0)$, the corresponding equation can be expressed as:

$$\begin{cases} (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 - R^2 = 0 \\ a(x - x_0) + b(y - y_0) + z - z_0 = 0 \end{cases} \tag{1}$$

The nonlinear model of GNSS antenna edge is established by removing the higher order terms through Taylor series expansion. After that, least square method is applied to calculate the coordinate of GNSS antenna edge's center.

(2) Extraction of axis normal vector
There exists a relative distance between the center of antenna edge and that of antenna. An axis normal vector is needed for the deducing of GNSS geometric center.
The axis normal vector can be extracted from the circle equation of antenna's edge as formula (2):

$$\overrightarrow{r} = \left(\frac{-a}{\sqrt{a^2 + b^2 + 1}} \quad \frac{-b}{\sqrt{a^2 + b^2 + 1}} \quad \frac{1}{\sqrt{a^2 + b^2 + 1}} \right) \tag{2}$$

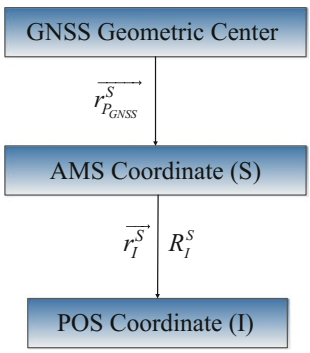


Fig. 5 GNSS lever arm parameters calibration model

Besides, the axis normal vector can be extracted from the fitting of GNSS connecting rod, due to its coaxial character with GNSS antenna edge, which can also be used to validate the accuracy of axis normal vector extraction [28].

2. Attain the coordinate of GNSS geometric center

   With the relative distance ΔL, GNSS antenna edge's center $\overrightarrow{r^S_{P_{ANT}}}$ and axis normal vector $\vec{r}$, GNSS geometric center can be calculated through formula (3):

$$\overrightarrow{r^S_{P_{GNSS}}} = \overrightarrow{r^S_{P_{ANT}}} + \Delta L \cdot \vec{r} \tag{3}$$

3. Establish the coordinate connection between POS and AMS.
   With the common observation points between POS and AMS, the transformation can be expressed as formula (4):

$$\overrightarrow{r_{I_j}} = k_2 \cdot R^I_S \cdot \overrightarrow{r_{S_j}} + \overrightarrow{r^I_S} \tag{4}$$

4. Transfer the geometric center into POS coordinate
   Substitute (3) into (4), we get formula (5), namely, GNSS lever arm in POS coordinate:

$$\overrightarrow{r^I_{P_{GNSS}}} = k_2 \cdot R^I_S \cdot \overrightarrow{r^S_{P_{GNSS}}} + \overrightarrow{r^I_S} \tag{5}$$

3.2 LS Relative Mounting Parameters Calibration Model

LS mounting parameters consist of the LS position parameters, as well as orientation parameters, in relevant with POS coordinate. Through AMS, this paper establish the transformation between POS coordinate and AMS coordinate, from which we can deduce position parameters from translation vector and orientation parameters from rotation matrix. The calibration can be seen in Fig. 6.

1. Establish the coordinate connection between LS and AMS.

$$\overrightarrow{r_{S_i}} = k_1 \cdot R^S_L \cdot \left(R(\theta, \varphi) \cdot \overrightarrow{r_{L_i}}\right) + \overrightarrow{r^S_L} \tag{6}$$

2. Establish the coordinate connection between POS and AMS.

$$\overrightarrow{r_{I_j}} = k_2 \cdot R^I_S \cdot \overrightarrow{r_{S_j}} + \overrightarrow{r^I_S} \tag{7}$$

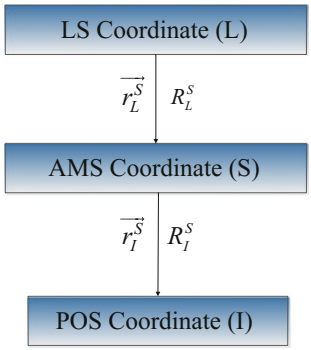


Fig. 6 LS mounting parameters calibration model

3. Establish the coordinate connection between POS and LS.
   Substitute (6) into (7), we get formula (8), namely, the transformation between POS coordinate and LS coordinate.

$$\overrightarrow{r_{I_j}} = k_3 \cdot R_M \cdot \overrightarrow{r_{L_j}} + \overrightarrow{r_L^I} \tag{8}$$

where $k_3 = k_1 \cdot k_2$ represents the scale change between LS coordinate and POS coordinate; $\overrightarrow{r_L^I} = k_2 \cdot R_S^I \cdot \overrightarrow{r_L^S} + \overrightarrow{r_S^I}$ represents the translation vector, namely position parameters of LS; $R_M = R_S^I \cdot R_L^S$ represents the rotation matrix, from which we can calculate the orientation parameters through the following formula:

$$\begin{aligned} \alpha_M &= \arctan\left(\frac{R_M(3,2)}{R_M(3,3)}\right) \\ \beta_M &= -\arcsin(R_M(1,3)) \\ \gamma_M &= \arctan\left(\frac{R_M(2,2)}{R_M(2,1)}\right) \end{aligned} \tag{9}$$

3.3 *Extrinsic Calibration Procedure Error Model*

While the vehicle is in motion state, there exists no obvious line vibration or angular vibration, making the trajectory relatively smooth. Therefore, we can filter the measurement data based on the condition of smooth trajectory to remove part of surveying error in the measurement system.

From formulas (5) and (8), we can find that the essence of lever arm and mounting parameters is the transformation between different coordinate. In this paper, we deduce the error model from the angle of coordinate transformation.

The transformation between coordinate A and coordinate B can be expressed as formula (10):

$$\overrightarrow{r_B} = k \cdot R_A^B \cdot \overrightarrow{r_A} + \overrightarrow{r_B^A} \tag{10}$$

where α, β, γ represents the rotation angle from A to B; k represents the scale factor between A and B; X_0, Y_0, Z_0 represents the translation from A to B; All of them form the seven-parameter transformation vector $\overrightarrow{n}$.

The expansion of matrix R_A^B can be expressed as:

$$R_A^B = \begin{pmatrix} \cos\gamma\cos\beta & -\cos\beta\sin\gamma & \sin\beta \\ \cos\alpha\sin\gamma + \cos\gamma\sin\beta\sin\alpha & \cos\gamma\cos\alpha - \sin\gamma\sin\beta\sin\alpha & -\cos\beta\sin\alpha \\ \sin\gamma\sin\alpha - \cos\gamma\cos\alpha\sin\beta & \cos\gamma\sin\alpha + \cos\alpha\sin\gamma\sin\beta & \cos\beta\cos\alpha \end{pmatrix}$$

The expansion of formula (10) can be expressed as:

$$\begin{cases} X_B = f(n) = k \cdot (R_{11} \cdot X_A + R_{12} \cdot Y_A + R_{13} \cdot Z_A) + X_0 \\ Y_B = g(n) = k \cdot (R_{21} \cdot X_A + R_{22} \cdot Y_A + R_{23} \cdot Z_A) + Y_0 \\ Z_B = h(n) = k \cdot (R_{31} \cdot X_A + R_{32} \cdot Y_A + R_{33} \cdot Z_A) + Z_0 \end{cases} \tag{11}$$

The linearization of formula (11) can be expressed as:

$$\begin{cases} V_{X_B} = \frac{\partial f(n)}{\partial X_0}\delta X_0 + \frac{\partial f(n)}{\partial Y_0}\delta Y_0 + \frac{\partial f(n)}{\partial Z_0}\delta Z_0 + \frac{\partial f(n)}{\partial \alpha}\delta\alpha + \frac{\partial f(n)}{\partial \beta}\delta\beta + \frac{\partial f(n)}{\partial \gamma}\delta\gamma + \frac{\partial f(n)}{\partial k}\delta k - l_{X_A} \\ V_{Y_B} = \frac{\partial g(n)}{\partial X_0}\delta X_0 + \frac{\partial g(n)}{\partial Y_0}\delta Y_0 + \frac{\partial g(n)}{\partial Z_0}\delta Z_0 + \frac{\partial g(n)}{\partial \alpha}\delta\alpha + \frac{\partial g(n)}{\partial \beta}\delta\beta + \frac{\partial g(n)}{\partial \gamma}\delta\gamma + \frac{\partial g(n)}{\partial k}\delta k - l_{Y_A} \\ V_{Z_B} = \frac{\partial h(n)}{\partial X_0}\delta X_0 + \frac{\partial h(n)}{\partial Y_0}\delta Y_0 + \frac{\partial h(n)}{\partial Z_0}\delta Z_0 + \frac{\partial h(n)}{\partial \alpha}\delta\alpha + \frac{\partial h(n)}{\partial \beta}\delta\beta + \frac{\partial h(n)}{\partial \gamma}\delta\gamma + \frac{\partial h(n)}{\partial k}\delta k - l_{Z_A} \end{cases} \tag{12}$$

Its residuals matrix can be expressed as $Vr_B = A \cdot \delta n - lr_A$; Its relevant weight matrix P; The number of common points i ($i \geq 3$). Then, we can deduce the unit weight variance as formula (13):

$$\delta^2 = \frac{Vr_B^T P Vr_B}{3i - 7} \tag{13}$$

Usually, P is a unit weight matrix. So the accuracy of translation, rotation and scale factor is relevant. We assume that the Maximum range of common points is L and the value of rotation is in radians. The relevance of the accuracy can be expressed as formula (14):

$$m_\alpha^2 = m_\beta^2 = m_\gamma^2 = m_k^2 = \frac{m_{X_0}^2}{L^2} = \frac{m_{Y_0}^2}{L^2} = \frac{m_{Z_0}^2}{L^2} = \frac{\delta^2}{L^2} \tag{14}$$

The corresponding translation error and rotation error can be expressed as formula (15):

$$\begin{cases} m_T^2 = m_{X_0}^2 + m_{Y_0}^2 + m_{Z_0}^2 = 3\delta^2 \\ m_R^2 = m_\alpha^2 + m_\beta^2 + m_\gamma^2 = \dfrac{3\delta^2}{L^2} \end{cases} \tag{15}$$

1. **Error model of GNSS lever arm**
   From formula (5), we can find that the accuracy of lever arm is mainly influenced by the transformation error m_{trf1} (between S and I) and GNSS center fitting error m_{fit}. We assume the transformation vector (between S and I) consists of k_1, α_1, β_1, γ_1, X_{01}, Y_{01}, Z_{01}, with its corresponding error m_{k_1}, m_{α_1}, m_{β_1}, m_{γ_1}, $m_{X_{01}}$, $m_{Y_{01}}$, $m_{Z_{01}}$ and the maximum range is L_1.
   Differential method has been applied in formula (5) to develop GNSS lever arm error model:

$$\begin{aligned} m_{r_{P_{GNSS}}^I}^2 = & \left| R_S^I \cdot \overrightarrow{r_{P_{GNSS}}^S} \right|^2 \cdot m_k^2 + \left| \frac{\partial R_S^I}{\partial \alpha_1} \cdot \overrightarrow{r_{P_{GNSS}}^S} \right|^2 \cdot m_{\alpha_1}^2 + \left| \frac{\partial R_S^I}{\partial \beta_1} \cdot \overrightarrow{r_{P_{GNSS}}^S} \right|^2 \cdot m_{\beta_1}^2 \\ & + \left| \frac{\partial R_S^I}{\partial \gamma_1} \cdot \overrightarrow{r_{P_{GNSS}}^S} \right|^2 \cdot m_{\gamma_1}^2 + m_{X_{01}}^2 + m_{Y_{01}}^2 + m_{Z_{01}}^2 + \left| k_2 \cdot R_S^I \right|^2 \cdot m_{fit}^2 \end{aligned} \tag{16}$$

2. **Error model of LS mounting parameters**
   From formula (8), we can find that the accuracy of LS mounting parameters is mainly influenced by the transformation error m_{trf1} (between S and I) and the translation error m_{T_2} (between L and S). We assume the transformation vector (between S and L) consists of k_2, α_2, β_2, γ_2, X_{02}, Y_{02}, Z_{02}, with its corresponding error m_{k_2}, m_{α_2}, m_{β_2}, m_{γ_2}, $m_{X_{02}}$, $m_{Y_{02}}$, $m_{Z_{02}}$ and the maximum range is L_2.
   Differential method has been applied in formula (8) to develop LS position parameters error model:

$$\begin{aligned} m_{r_L^I}^2 = & \left| R_S^I \cdot \overrightarrow{r_L^S} \right|^2 \cdot m_k^2 + \left| \frac{\partial R_S^I}{\partial \alpha_1} \cdot \overrightarrow{r_L^S} \right|^2 \cdot m_{\alpha_1}^2 + \left| \frac{\partial R_S^I}{\partial \beta_1} \cdot \overrightarrow{r_L^S} \right|^2 \cdot m_{\beta_1}^2 \\ & + \left| \frac{\partial R_S^I}{\partial \gamma_1} \cdot \overrightarrow{r_L^S} \right|^2 m_{\gamma_1}^2 + m_{X_{01}}^2 + m_{Y_{01}}^2 + m_{Z_{01}}^2 \\ & + \left| R_S^I \cdot \frac{\partial \overrightarrow{r_L^S}}{\partial X_{02}} \right| \cdot m_{X_{02}}^2 + \left| R_S^I \cdot \frac{\partial \overrightarrow{r_L^S}}{\partial Y_{02}} \right| \cdot m_{Y_{02}}^2 + \left| R_S^I \cdot \frac{\partial \overrightarrow{r_L^S}}{\partial Z_{02}} \right| \cdot m_{Y_{02}}^2 \end{aligned} \tag{17}$$

Orientation parameters can be deduced from the rotation matrix $R_M = R_S^I \cdot R_L^S$, which can be specifically expressed as $R_M = R(a_1, \beta_1, \gamma_1) \cdot R(a_2, \beta_2, \gamma_2)$. But, differential method is not suitable for nonlinear and complex equation, so, error of orientation parameters is discussed using Monte-Carlo method in this paper. The extraction of orientation parameters α_M, β_M, γ_M can be expressed as $(\alpha_M, \beta_M, \gamma_M) = F(\mu)$. Vector μ consists of the orientation parameters α_1, β_1, γ_1 and α_2, β_2, γ_2, which represent the rotation between S and I, and rotation between S and L respectively. The rotation error can be expressed as $\Delta\mu$, which can realize 0 of bias mean value and m_μ of mean square error. The error model of LS orientation parameters can be expressed as:

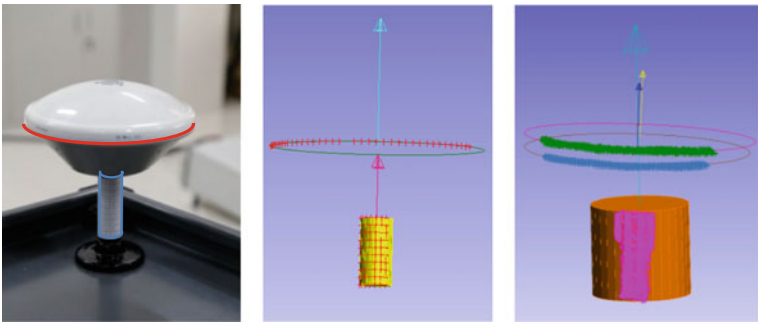
$$(\Delta\alpha_M, \Delta\beta_M, \Delta\gamma_M) = F(\mu + \Delta\mu) - F(\mu) \tag{18}$$

4 Experiments and Results

Experiment has been carried out with Trimble Zephyr II GNSS antenna, as shown in Fig. 7a. LS and TMS have been chosen as AMS in the calibration procedure. TMS has a measuring precision better than 0.1 mm, fulfill the need of GIL calibration [29]. While the accuracy of LS is influenced by the retro-reflective properties of non-cooperative targets.

4.1 GNSS Lever Arm Calibration Results

1. **Fitting results of GNSS antenna edge center**
   The visual effects of fitting based on TMS and LS (RigeL VZ400) can be seen in Fig. 7b, c, with the corresponding fitting results of GNSS geometric center be seen in Table 2.
   From Table 2, we can find that the fitting accuracy of TMS is 0.06 mm, while the fitting accuracy of LS is only 3.6 mm, due to the influence of the non-cooperative targets [30].
2. **Transformation parameters between S and I**
   Retro-reflective targets have been chosen as the common observation points between POS and AMS. We can find the relevant translation and rotation parameters in Tables 3 and 4 and deducing the corresponding error through the proposed error model.
   From Tables 4 and 5, we can find that the transformation accuracy, supported by TMS, is better than 1 mm, while the accuracy of LS-supported transformation is better than 2 mm. we can calculate the translation error and rotation error



(a) GNSS antenna (b) Circle fitting to points from TMS (c) Circle fitting to point cloud from LS

Fig. 7 Circle fitting of GNSS geometric center

Table 2 Fitting results of GNSS antenna edge center (mm)

	LS	TMS
Circle center	(1120.3, −1708.7, −304.1)	(1569.2, −2540.2, −477.5)
Circle fitting accuracy	3.6	0.06
Axial vector of circle fitting	(0.01, −0.02, 1)	(0.0004, −0.0033, 1)
Axial vector of cylinder fitting	(−0.03, 0.02, 1)	(−0.001, −0.0002, 1)

Table 3 Transformation parameters from coordinate S to coordinate I (supported by LS)

	X axis	Y axis	Z axis
Translation (mm)	956.26	−1259.17	742.40
Rotation (°)	0.649	0.337	176.409
RMSE (mm)	0.75	0.12	0.47
Scale factor	0.996038		

through formulas (18) and (19): $m_{T_{POS1}} = 1.38$ mm, $m_{R_{POS1}} = 0.046$; $m_{T_{POS2}} = 0.82$ mm, $m_{R_{POS2}} = 0.017°$.

3. **GNSS lever arm parameters**

By applying GNSS geometric center and transformation parameters (between S and I) into formula (12), we can attain GNSS lever arm parameters, as shown in Table 5.

With GNSS lever arm error model, the accuracy of the LS-supported GNSS lever arm calibration and that of TMS-supported calibration can be calculated: $mx_1 = 2.21$ mm, $my_1 = 2.35$ mm, $mz_1 = 2.35$ mm; $mx_2 = 0.63$ mm, $my_2 = 0.57$ mm, $mz_2 = 0.69$ mm. We can find that TMS-based scheme is more accurate than LS-based scheme, and it can be used as a reference. By comparing the lever arm parameters in Table 5, we can tell that the accuracy of GNSS lever arm (LS-supported calibration) is $mx_{GNSS} = 3.5$ mm, $my_{GNSS} = 2.2$ mm,

Table 4 Transformation parameters from coordinate S to coordinate I (supported by TMS)

	X axis	Y axis	Z axis
Translation (mm)	1767.63	−1867.75	924.63
Rotation (°)	0.284	0.238	−174.272
RMSE (mm)	0.46	0.17	0.21
Scale factor	1.001921		

Table 5 GNSS lever arm parameters (mm)

Lever arm	Laser scanner	Theodolite
X axis	−56.2	−52.7
Y axis	512.7	510.5
Z axis	466.7	465.5

mz_{GNSS} = 1.2 mm, which is in consistent with the accuracy deduced from the proposed error model.

4.2 LS Mounting Parameters Calibration Results

Retro-reflective targets and spherical targets can be used as the common observation points between LS (FARO FOCUS) and TMS (Rigel VZ400).

The recognition of different types of common point can be seen in Fig. 8 and the transformation accuracy between LS and TMS can be seen in Tables 6 and 7.

With LS mounting parameters error model, the accuracy of the LS-supported LS mounting parameters calibration and that of TMS-supported calibration can be calculated:

(1) Position parameters accuracy: mx_1 = 1.82 mm, my_1 = 1.83 mm, mz_1 = 1.83 mm; mx_2 = 0.66 mm, my_2 = 0.67 mm, mz_2 = 0.67 mm.
(2) Orientation parameters accuracy: $m\alpha_1 = 0.0288°$, $m\beta_1 = 0.0287°$, $m\gamma_1 = 0.0292°$; $m\alpha_2 = 0.0102°$, $m\beta_2 = 0.0101°$, $m\gamma_2 = 0.0103°$.

We can find that TMS-based scheme is more accurate than LS-based scheme, and it can be used as a reference. By comparing the transformation parameters in

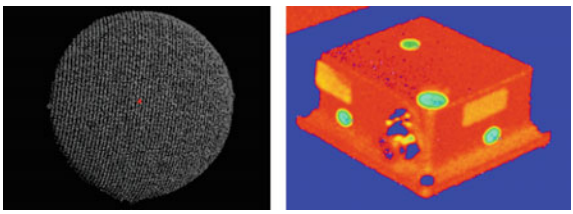


Fig. 8 Recognition of different cooperation targets

(a) Recognition of spherical target

(b) Recognition of retro-reflect target

Table 6 Transformation parameters from coordinate L to coordinate S (supported by LS)

	X axis	Y axis	Z axis
Translation (mm)	673.18	−970.49	219.15
Rotation (°)	0.413	0.176	−79.685
RMSE (mm)	1.62	1.11	0.66
Scale factor	1.002285		

Table 7 Transformation parameters from coordinate L to coordinate S (supported by TMS)

	X axis	Y axis	Z axis
Translation (mm)	1246.66	−1744.03	36.41
Rotation (°)	0.016	0.012	−89.022
RMSE (mm)	0.35	0.47	0.40
Scale factor	1.00026		

Tables 6 and 7, we can tell that the accuracy of LS position parameters (LS-supported calibration) is mx_{LS} = 1.7 mm, my_{LS} = 1.0 mm, mz_{LS} = 0.7 mm, the accuracy of LS orientation parameters (LS-supported calibration) is $m\alpha_{LS}$ = 0.042°, $m\beta_{LS}$ = 0.037°, $m\gamma_{LS}$ = 0.02°, which is in consistence with the accuracy deduced from the proposed error model.

4.3 Summary

From the proposed error model, we can deduce the accuracy of LS-based GNSS lever arm calibration is 4.0 mm, and the accuracy of LS-based laser scanner position parameters and orientation parameters is 3.2 mm and 0.36°, respectively. While the comparison between TMS and LS, the accuracy of LS-based GNSS lever arm calibration is 4.3 mm and the accuracy of LS-based laser scanner position parameters and orientation parameters is 2.1 mm and 0.43°. Comparison results are in coincident with the deductions form the proposed error model, further verifying the feasibility of LS-based GIL rapid extrinsic calibration scenario.

5 Conclusion

(1) Researches about extrinsic calibration of seamless multi-sensor navigation system have been carried in this paper, and a LS-based calibration scheme has been proposed, which is independent of GNSS signal, more manipulable and more adaptive to complex environment. Error analyses and TMS-supported experiments have proved the feasibility and efficiency of the extrinsic calibration scheme.

(2) TMS-based comparison experiments have proved advantages of LS-based scheme: (1) High speed calibration, the whole procedure can be completed within a few minutes; (2) High efficiency calibration, automatic recognizing targets have helped to reduce workload and save time; (3) Environmental adaptive calibration, LS-based calibration has few requests on precise leveling and centering, more operative in practice; (4) Abundant measuring information, LS can attain common observation points, as well as additional points cloud, which can be used to testify the calibration results.
(3) LS-based calibration scheme has its deficiencies at the same time: (1) Limited measuring precision, due to the influence of the non-cooperative targets, the accuracy of LS decreases dramatically, as can be found in the GNSS geometric center's fitting results; (2) Weak versatility, different laser sensor has different measurement principle, making it hard to find one appropriate common observation target, which can be automatically recognized by all those laser scanners. S0, automatic target recognition technology and image fusion technology should be enhanced to compensate the weakness of LS-based calibration scheme.
(4) This paper mainly focuses on the rapid extrinsic calibration of multi-sensor navigation system. Future implementation will be extended to filtering algorithm based online calibration as positioning and orientating information from POS will be added for completing the whole calibration scheme of seamless indoor/outdoor multi-sensor navigation and positioning system.

Acknowledgements This work was supported by the Key Fund of National Key Laboratory of Geo-Information Engineering, No. SKLGIE2014-Z-2-1 and the National Science Foundation of China, No. 41274014.

References

1. Nam T, Pardo TA (2011) Conceptualizing smart city with dimensions of technology, people, and institutions. In: International conference on digital government research
2. Deng Z, Li X, Ma W (2014) Design and implementation of indoor and outdoor positioning service platform based on Beidou and base station system. In: 2014 proceedings of China satellite navigation conference (CSNC)
3. Chen Y, Chen R, Pei L, Kröger T (2010) Knowledge-based error detection and correction method of a multi-sensor multi-network positioning platform for pedestrian indoor navigation 37(16):873–879
4. Pesyna KM Jr, Kassas ZM, Bhatti JA, Humphreys TE (2011) Tightly-coupled opportunistic navigation for deep urban and indoor positioning. In: Proceedings of international technical meeting of the satellite division of the institute of navigation 10(1):3605–3616
5. BAE (2012) BAE navigation system identifies suspicious signals. Engineer
6. Zhongliang D, Zhang S, Jiao J, Lianming XU (2016) Research and application of high-precision indoor location-aware big data. J Comput Appl
7. Kuusniemi H, Liu J, Pei L, Chen Y (2012) Reliability considerations of multi-sensor multi-network pedestrian navigation. IET Radar Sonar Navig 6(3):157–164

8. Chu T, Guo N, Backén S, Akos D (2012) Monocular camera/IMU/GNSS integration for ground vehicle navigation in challenging GNSS environments. Sensors 12(3):3162–3185
9. Gong X, Wei D, Li X, Xu Y, Yuan H (2015) Visual-aided pedestrian/GNSS coupled navigation for smart terminals. In: 2015 proceedings of China satellite navigation conference (CSNC)
10. Chow JCK (2014) Multi-sensor integration for indoor 3D reconstruction. University of Calgary
11. He SJ, Zhao ST, Bai F, Wei J (2013) A method for spatial data registration based on PCA-ICP algorithm. Adv Mater Res 718–720:1033–1036
12. Montemerlo MS (2015) Fastslam: a factored solution to the simultaneous localization and mapping problem with unknown data association. Arch Environ Contam Toxicol 50(2): 240–248
13. Mahler SP, Reece JL (2007) Global positioning systems, inertial navigation, and integration. Wiley
14. SPAN-CPT Product Sheet [Online]. Available from: http://www.novatel.com/assets/Documents/Papers/SPAN-CPT.pdf. Accessed 10 June 2016
15. FARO Laser Scanner Focus3D X130 HDR [Online]. Available from: http://www.faroasia.com/resource-centre/registeredUserFrm.php. Accessed 10 June 2016
16. Rogel VZ400 DataSheet [Online]. Available from: http://www.riegl.com/uploads/tx_pxpriegldownloads/10_DataSheet_VZ-400_2014-09-19.pdf. Accessed 10 June 2016
17. Wang L (2010) Automatic registration of terrestrial lidar data using planar targets. Information Engineering University
18. Li Z, Li G, Wu X (2003) Review of accuracy measuring technology of antenna's surface. Bull Survey Mapp 6:16–19
19. Yu MJ (2012) INS/GPS integration system using adaptive filter for estimating measurement noise variance. IEEE Trans Aerosp Electron Syst 48(2):1786–1792
20. Lichti DD (2007) Error modelling, calibration and analysis of an AM–CW terrestrial laser scanner system. ISPRS J Photogrammetry Rem Sens 61(5):307–324
21. Kavzoglu T, Karsli F (2008) Calibration of a digital single lens reflex (SLR) camera using artificial neural networks. In: ISPRS
22. Han S, Wang J (2011) Quantization and colored noises error modeling for inertial sensors for GPS/INS integration. IEEE Sens J 11(6):1493–1503
23. Zetian YE, Yang Y, Zhao W (2011) Dynamic calibration of exterior orientations for vehicle GPS/IMU/LS laser imaging system. Acta Geod Cartogr Sin 40(3):344–345
24. Fu L, Zhang J, Li R (2015) A novel vision-aided GNSS/INS integration method using line features for aviation approach and landing
25. Groves PD (2015) Principles of GNSS, inertial, and multisensor integrated navigation systems, 2nd edn. IEEE Aerosp Electron Syst Mag 30(2):26–27
26. Liu S (2012) Research and implementation of GPS/INS integrated navigation algorithms. Information Engineering University
27. Zhou Y, Wang L, Li G, Zong W, Wang S (2015) Global error model and accuracy analysis of MLS. Bull Survey Mapp 9:19–23
28. Li G, Li Z (2011) Principles and applications of the industrial measuring system. Surveying and Mapping Press
29. SPAN on OEM6 Firmware Reference Manual [Online]. Available from: http://www.novatel.com/assets/Documents/Manuals/OM-20000144UM.pdf. Accessed 10 June 2016
30. Zhang Q (2012) Research of methods to the performance behavior of terrestrial laser scanner. Information Engineering University

The GPS/INS Integrated Navigation Method Based on Adaptive SSR-SCKF Cubature Kalman Filter

Zhe Yue, Baowang Lian and Chengkai Tang

Abstract There are many methods aiming at the nonlinear problem of GPS/INS integrated navigation, such as EKF and UKF, however, these methods have low positioning accuracy and instability. On the study of SCKF and nonlinear model of GPS/INS integrated navigation, aiming at the issues that the state equation of GPS/INS is nonlinear while the measured equation is linear, and the measured noise changes owing to the changing number of visible satellites or multipath. Therefore, this paper promotes the integrated method based on adaptive SSR-SCKF, which uses the spherical simple-radial cubature rule (SSRCR) to set the cubature sampling points. We also provide a linear measured update process on the basis of singular value decomposition (SVD), and it avoids choosing the cubature sampling points. Combining the moving window method, it can adjust the covariance matrix of measurement noise in real-time. The experiment results show that the proposed method has lower computational complexity, while higher estimated accuracy, numerical stability and better adaptive ability to the changing noise than SCKF in the same conditions.

Keywords GPS/INS integrated navigation · SSR-SCKF · Singular value decomposition · Adaptive

1 Introduction

GPS/INS integrated navigation system usually uses the Kalman filter to estimate the system [1], and the estimation method assumes that the equation is linear, however, the GPS/INS integrated navigation system is nonlinear, therefore, the nonlinear

Z. Yue (✉) · B. Lian · C. Tang
School of Electronics and Information,
Northwestern Polytechnical University, Xi'an, China
e-mail: yuezhe5471@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_32

model can more completely spread the feature of system error [2]. Aiming to the nonlinear model, we usually use the extend kalman filter (EKF) to make the nonlinear equation [3]. However, the linear process results in relatively higher truncation error, and the EKF also require calculating the complex Jacobi matrix. The unscented kalman filter (UKF) uses the unscented transformation (UT) to make Gaussian approximation for the probability density function [3, 4], but if the dimension is more than three, i.e., $n \geq 4$, the suggested tuning parameter will less than zero, i.e., $k = 3 - n < 0$, this phenomenon may make the covariance matrix negative and break the filter. Arasaratnam [5] firstly proposes the cubature kalman filter (CKF), and employs the three-order spherical-radial cubature rule to transform the integral calculation of the Gauss weighted multidimensional nonlinear function in cartesian coordinates to the cubature approximation, and it has higher estimated accuracy than UKF, in addition, the weights of the sampling points of CKF are the same and positive, and the numerical stability is better than UKF. It is said that the CKF has the highest estimated accuracy in the Gauss filters [6]. The superior nonlinear performance of CKF attracts more and more scholars in the field of GPS/INS integrated navigation [7–9]. The existing integrated navigation methods based on CKF usually utilize three order spherical-radial cubature rule to choose the cubature points in the state prediction and measured update process, and the calculation is complex and the numerical stability is not high. Additionally, there is little research on the adaptive CKF about the changing measured noise, owing to the changing number of the visible satellites or the multipath.

Thus we promote the GPS/INS integrated navigation method of the adaptive spherical simplex-radial square-root cubature filter (SSR-SCKF) based on the spherical simplex-radial cubature rule (SSRCR). This method utilizes the SSRCR instead of the three-order spherical-radial cubature rule to achieve higher estimated accuracy. The measured equation of GPS/INS loosely integrated navigation is linear, thus we promote a simple filter based on singular value decomposition (SVD), which reduces the computation complexity, because it does not need to calculate the sampling points in the measured update. Moreover, the SVD can also avoid interruption when the new information covariance matrix cannot be inversed, and it also improves the numerical stability of the method. Meanwhile, in order to coping with the changing measured noise, we use the moving window method to estimate the measured noise covariance matrix in real-time, thus can greatly improve the adaptive ability and estimated accuracy of the changing noise. The experiment shows that, under the same conditions, the method we proposed has greater estimated accuracy, higher numerical stability and lower computational complexity than SCKF. When the measured noise changed, this method has better adaptive ability, and we also prove the superiorities of the method in this paper.

2 The Nonlinear Model of GPS/INS Integrated Navigation

In this paper, we set the east-north-up coordinate as the navigation coordinate of INS and the right-forward-up coordinate as the body coordinate (b coordinate). The idea navigation coordinate is called n coordinate, and the calculated navigation coordinate is named n' coordinate. The front right coordinate system of the moving object is the carrier coordinate system, i.e., the b system, and the geocentric inertial coordinate system is i system, and the earth coordinate named e coordinate, we also choose the navigation error as the state variables of the filter. The GPS/INS integrated navigation system is loosely, it utilizes the errors of attitude, velocity, positioning, gyroscope drift and constant bias of accelerometer as the 15 error state variables of the filter. The nonlinear error model of the attitude error, velocity error and positioning error are expressed as the follows:

$$\begin{cases} \dot{\phi} = C_{\omega}^{-1}\left(\left(I - C_{n}^{n'}\right)\hat{\omega}_{in}^{n} + C_{n}^{n'}\delta\omega_{in}^{n} - C_{b}^{n'}\delta\omega_{ib}^{b}\right) \\ \delta\dot{v} = \left(I - C_{n'}^{n}\right)C_{b}^{n'}\hat{f}_{b} + C_{b}^{n}\delta f_{b} \\ \qquad -\left(2\hat{\omega}_{ie}^{n} + \hat{\omega}_{en}^{n}\right) \times \delta v - \left(2\delta\omega_{ie}^{n} + \delta\omega_{en}^{n}\right) \times v \\ \delta\dot{L} = \frac{1}{R_M + h}\delta v_N - \frac{v_N}{(R_M + h)^2}\delta h \\ \delta\dot{\lambda} = \frac{\sec L}{R_N + h}\delta v_E + \frac{v_E \sec L \tan L}{R_N + h}\delta L - \frac{v_E \sec L}{(R_N + h)^2}\delta h \\ \delta\dot{h} = \delta v_U \\ \dot{\varepsilon}^{b} = 0 \\ \dot{\nabla}^{b} = 0 \end{cases} \tag{1}$$

$$C_{\omega}^{-1} = \frac{1}{\cos\phi_E}\begin{bmatrix} \cos\phi_N\cos\phi_E & 0 & \sin\phi_N\cos\phi_E \\ \sin\phi_N\sin\phi_E & \cos\phi_E & -\cos\phi_N\sin\phi_E \\ -\sin\phi_N & 0 & \cos\phi_N \end{bmatrix} \tag{2}$$

$$C_{n}^{n'} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{3}$$

$$\begin{cases} a_{11} = \cos\phi_N\cos\phi_U - \sin\phi_E\sin\phi_N\sin\phi_U \\ a_{12} = \cos\phi_N\sin\phi_U - \sin\phi_E\sin\phi_N\cos\phi_U \\ a_{13} = -\cos\phi_E\sin\phi_N \\ a_{21} = -\cos\phi_E\sin\phi_U \\ a_{22} = \cos\phi_E\cos\phi_U \\ a_{23} = \sin\phi_E \\ a_{31} = \sin\phi_N\cos\phi_U + \sin\phi_E\cos\phi_N\sin\phi_U \\ a_{32} = -\sin\phi_E\cos\phi_N\cos\phi_U + \sin\phi_N\sin\phi_U \\ a_{33} = \cos\phi_E\cos\phi_N \end{cases} \tag{4}$$

$$\begin{cases} \hat{\omega}_{in}^{n} = \omega_{in}^{n} + \delta\omega_{in}^{n} \\ \omega_{in}^{n} = \omega_{ie}^{n} + \omega_{en}^{n} \\ \delta\omega_{in}^{n} = \delta\omega_{ie}^{n} + \delta\omega_{en}^{n} \\ \omega_{ie}^{n} = [0 \;\; \omega_{ie}\cos L \;\; \omega_{ie}\sin L]^{T} \\ \omega_{en}^{n} = \left[\frac{-v_N}{R_M+h} \;\; \frac{v_E}{R_N+h} \;\; \frac{v_E \tan L}{R_N+h}\right]^{T} \\ \delta\omega_{ie}^{n} = [0 \;\; -\omega_{ie}\sin L\cdot\delta L \;\; \omega_{ie}\cos L\cdot\delta L]^{T} \\ \delta\omega_{en}^{n} = \begin{bmatrix} -\frac{\delta v_N}{R_M+h} + \frac{v_N\delta h}{(R_M+h)^2} \\ \frac{\delta v_E}{R_N+h} - \frac{v_E\delta h}{(R_N+h)^2} \\ \frac{\tan L\cdot\delta v_E}{R_N+h} + \frac{v_E \sec^2 L\cdot\delta L}{R_N+h} - \frac{v_E\tan L\cdot\delta h}{(R_N+h)^2} \end{bmatrix} \end{cases} \tag{5}$$

$\phi = [\phi_E \;\; \phi_N \;\; \phi_U]^T$ denotes the misalignment angles of east, north and up-vertical direction, $\delta v = [\delta v_E \;\; \delta v_N \;\; \delta v_U]$ indicates the velocity errors of east, north and up-vertical direction, $\delta L, \delta\lambda, \delta h$ represent the latitude, longitude and altitude errors, respectively. $C_n^{n'}$ is the coordinate transformation matrix from the idea navigation coordinate to the calculated navigation coordinate. $\hat{\omega}_{in}^{n}$ means the value which the sum of the earth rotation rate ω_{ie}^{n} and the navigation coordinate rotation rate ω_{en}^{n} in the calculated navigation coordinate. $\delta\omega_{in}^{n}, \delta\omega_{ie}^{n}, \delta\omega_{en}^{n}$ are the calculated errors of $\hat{\omega}_{in}^{n}, \hat{\omega}_{ie}^{n}, \hat{\omega}_{en}^{n}$, respectively. L and h represent the latitude and altitude. R_M and R_N suggest the radius of the meridian and prime vertical of the earth. The gyro error $\delta\omega_{ib}^{b}$ is mainly include the gyro constant bias ε^b and the zero mean Gauss white noise w_g^b. The accelerometer error δf_b is mainly include the accelerometer constant bias ∇^b and the zero mean Gauss white noise w_a^b.

We choose the state vector as Eq. (6):

$$X = [\phi_E \;\; \phi_N \;\; \phi_U \;\; \delta v_E \;\; \delta v_N \;\; \delta v_U \;\; \delta L \;\; \delta\lambda \;\; \delta h \;\; \varepsilon_E^b \;\; \varepsilon_N^b \;\; \varepsilon_U^b \;\; \nabla_E^b \;\; \nabla_N^b \;\; \nabla_U^b] \tag{6}$$

Then, discretizing Eq. (1) we can obtain the GPS/INS state equation. When take the positioning and velocity errors difference between GPS and INS into consideration, we can get the measurement equation as the follows.

$$Z = HX + V \tag{7}$$

$$Z = [\delta v_E \;\; \delta v_N \;\; \delta v_U \;\; \delta L \;\; \delta\lambda \;\; \delta h]^T \tag{8}$$

$$H = [0_{6\times3} \;\; I_{6\times6} \;\; 0_{6\times6}] \tag{9}$$

where V denotes the measured noise.

3 Filtering Method

3.1 The Filtering Method of SCKF

In the filtering method of SCKF, we take the following nonlinear equation into consideration.

$$x_k = f(x_{k-1}) + w_{k-1} \tag{10}$$

$$z_k = h(x_k) + v_k \tag{11}$$

where x_k is the unobservable state vector, z_k is the observable one. $f(\cdot)$ and $h(\cdot)$ represent the known nonlinear state and measured equation. The process noise w_{k-1} and measured noise v_k are white, and they are irrelevant, the mean of them is zero, the covariance matrixes are Q_{k-1} and R_k, respectively. The sampling points and weights calculated by the three-order spherical-radial cubature rule can describe as:

$$\begin{cases} \xi_i = \bar{x} + \sqrt{nP_x}e_i \\ \omega_i = \dfrac{1}{2n} \\ \xi_{i+n} = \bar{x} + \sqrt{nP_x}e_i \\ \omega_{i+n} = \dfrac{1}{2n} \end{cases} i = 1, 2, \ldots, n \tag{12}$$

where e_i is the unit vector of the ith element which is 1. The update steps of the SCKF can be found in [2].

3.2 The Filtering Method of Adaptive SSR-SCKF

In order to get the higher estimated accuracy of the nonlinear equation, this paper uses the SSRCR to choose the sampling points and weights as:

$$S_k^T = qr\left(\left[X'_{k/k-1} - G_k Z_{k/k-1} \quad G_k S_{R,k}\right]^T\right) \tag{13}$$

where $\left[a_{i,1}, a_{i,2}, \ldots, a_{i,n}\right]^T$, $i = 1, 2, \ldots, n+1$, is the column vector, which is composed with n + 1 vertexes of the n-spherical simplex and can be described as:

$$a_{i,j} = \begin{cases} -\sqrt{\frac{n+1}{n(n-j+2)(n-j+1)}}, & j < i \\ \sqrt{\frac{(n+1)(n-j+1)}{n(n-j+2)}}, & j = i \\ 0, & j > i \end{cases} \tag{14}$$

The initialization module and measured update module are the same as the SCKF except that the sampling rules, it uses the (13) to replace Eq. (12). Aiming at the measured equation of the GPS/INS loosely integrated navigation, we have no necessary to calculate the complex 2n cubature points in the measured update module. In order to improve the numerical stability and avoid the interruption when the new information covariance matrix does not have the inversion, this paper proposes a simple measured update process based on SVD.

A is the square-root matrix of the covariance matrix P. We can uses the SVD to $A_{m\times n}$, when $m<n$, $A = U[D\ \ 0]V^T$, U and V are unitary matrix, D is diagonal matrix. P^{-1} can be described as:

$$\begin{aligned} P^{-1} &= \left(AA^T\right)^{-1} \\ &= U\left(DD^T\right)^{-1}U^T \end{aligned} \tag{15}$$

The inversion of P can be described as (15), and we only need to calculate the inversion of diagonal element to get the inverse matrix of DD^T. Using the SVD can reduce the computation complexity and improve the numerical stability.

The measured update process based on SVD can be described as:

$$[U, D, V] = svd\left(\left[H_k S_{k/k-1}\ \ S_{R,k}\right]\right) \tag{16}$$

$$S_{zz,k/k-1} = UD \tag{17}$$

$$P_{xz,k/k-1} = S_{k/k-1}S_{k/k-1}^T H_k^T \tag{18}$$

$$G_k = P_{xz,k/k-1}P_{zz,k/k-1}^{-1} = P_{xz,k/k-1}U\left(DD^T\right)^{-1}U^T \tag{19}$$

$$\hat{x}_k = \hat{x}_{k/k-1} + G_k\left(z_k - \hat{z}_{k/k-1}\right) \tag{20}$$

$$S_k^T = qr\left(\left[S_{k/k-1} - G_k H_k S_{k/k-1}\ \ G_k S_{R,k}\right]^T\right) \tag{21}$$

The measured noise will change owing to the changing number of visible satellites and multipath, if we use the measured noise covariance matrix which is set by experience, it will result in relatively bigger estimated errors. Thus, in this paper we add the moving window method to the SSR-SCKF to change the measured noise covariance matrix adaptively.

We employ the new measured vector z_k and the $\hat{x}_{k/k-1}$ of the predicted module to define the new measured vector.

$$y_k = z_k - h\left(\hat{x}_{k/k-1}\right) = z_k - H_k\hat{x}_{k/k-1} \tag{22}$$

Therefore, the covariance matrix of the new measured vector is

$$R_{y_k} = R_k + H_k P_{k/k-1} H_k^T = S_k S_k^T + H_k S_{k/k-1} S_{k/k-1}^T H_k^T \tag{23}$$

The R_{y_k} can be received by the mean value in the length of the window with the moving window method.

$$R_{y_k} = \frac{1}{N} \sum_{j=0}^{N} y_{k-j} y_{k-j}^T, \quad k > N \tag{24}$$

The square-root of the measured noise covariance matrix can be received by the moving window method through (23) and (24).

$$S_{R,k} = chol\left(R_{y_k} - H_k S_{k/k-1} S_{k/k-1}^T H_k^T\right) \tag{25}$$

When taking $S_{R,k}$ into the (16) and (21) of the SSR-SCKF, we can obtain the GPS/INS integrated navigation method based on adaptive SSR-SCKF cubature kalman filter.

4 Simulation and Verification

The simulation parameters are set as the follows: the initial position of the body is $(L, \lambda, h) = (34.246°, 108.910°, 380\,\text{m})$, the gyro drift is $0.1°/h$, the random walk of gyro is $0.05°/\sqrt{h}$, the constant bias of the accelerometer is 100 μg, the random walk of accelerometer is $50\,\mu\text{g}/\sqrt{\text{Hz}}$, the initial positioning error is $(10, 10, 10)$ m/s, the initial velocity error is $(0.5, 0.5, 0.5)$ m/s, the initial attitude error is $(2°, -2°, 3°)$, the position noise of GPS is $(10, 10, 10)$ m, and the velocity noise is $(0.1, 0.1, 0.1)$ m/s. The flight path of the body is include acceleration, uniform speed, left turn, right turn, climb, descent, deceleration and static. The total time of the simulation is 600 s, the flight path is in Fig. 1. The sample frequency of INS is 100 Hz, and the sample frequency of GPS is 1 Hz.

The positioning and velocity error of SCKF and the method of this paper can be seen in Figs. 2 and 3.

From Figs. 2 and 3, we can conclude that the two methods both have relatively good integrated navigation performance, and the square-root errors are shown in Table 1. From the statistic results of Table 1, it can be seen that the method proposed in this paper is better than SCKF both in the positioning and velocity accuracy. The reason is that the sampling rule of this paper is SSRCR, not the three-order spherical-radial cubature rule as the SCKF. And in the measured update process, SCKF require to calculate the 30 cubature sampling points to make the cubature approximation each time, but the method of this paper do not need to, therefore it reduces the computation complexity.

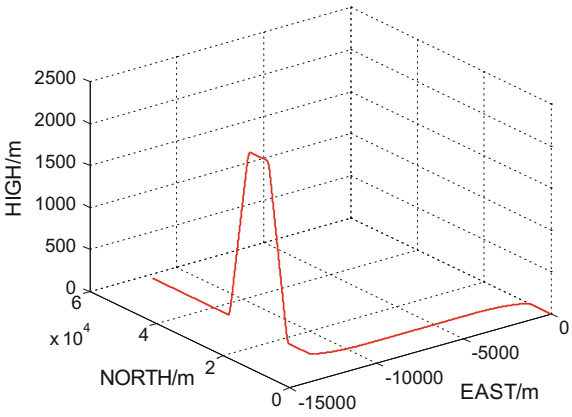


Fig. 1 Flight path

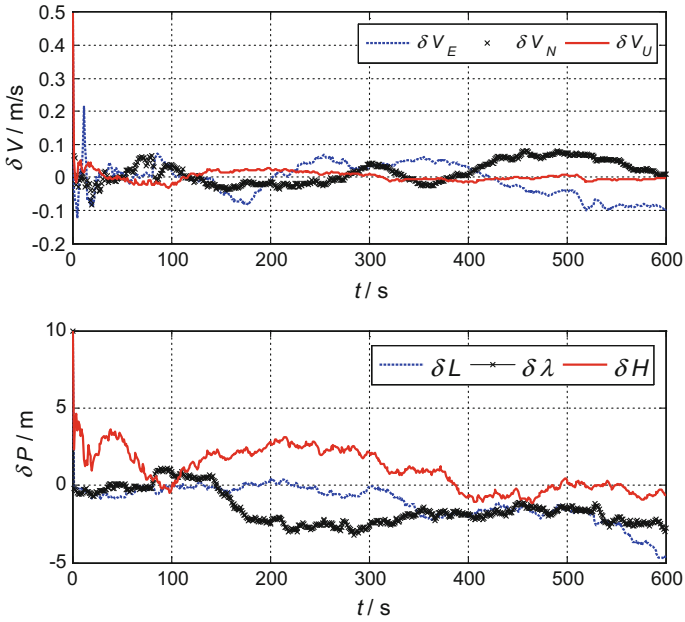


Fig. 2 Position and velocity error of SCKF

The measured noise will change owing to the changing numbers of the visible satellite and multipath, so in order to verify the adaptive ability to the changing measured noise of the method proposed in this paper, we set three different noise in the three stages. The first 200 s the measured noise of position and velocity are $(10, 10, 10)$ m and $(0.1, 0.1, 0.1)$ m/s, respectively. Then in the range of 200–400 s,

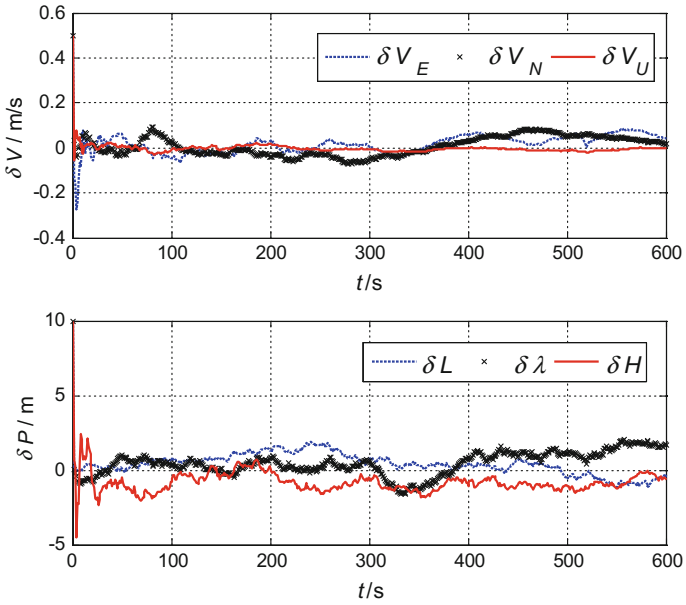


Fig. 3 Position and velocity error of the method of this paper

Table 1 The comparison of root mean square error of SCKF of two methods

	Positioning error (m)	Velocity error (m/s)	The number of sampling points in each update time
SCKF	3.0238	0.0732	60
The method of this paper	1.7307	0.0712	32

the noise becomes the three times to the initial noise, and the last 200 s the noise becomes 2 times to the initial noise. The simulation diagram of the position and velocity errors of SCKF and this paper are in Figs. 4 and 5.

From Fig. 4, we can suggest that, the filtering performance of SCKF is severely affected when the measured noise changes. Through Fig. 5 we can conclude that, under the same conditions, the method of this paper has better adaptive ability. The statistic results of RMS errors are in Table 2. From Table 2, we can also see that when the measured noise changes, the position and velocity errors of SCKF become bigger, because the measured noise does not update correspondingly, The method of this paper can adjust the covariance matrix of the measured noise in real-time, so it can receive the better accuracy and the better anti interference ability.

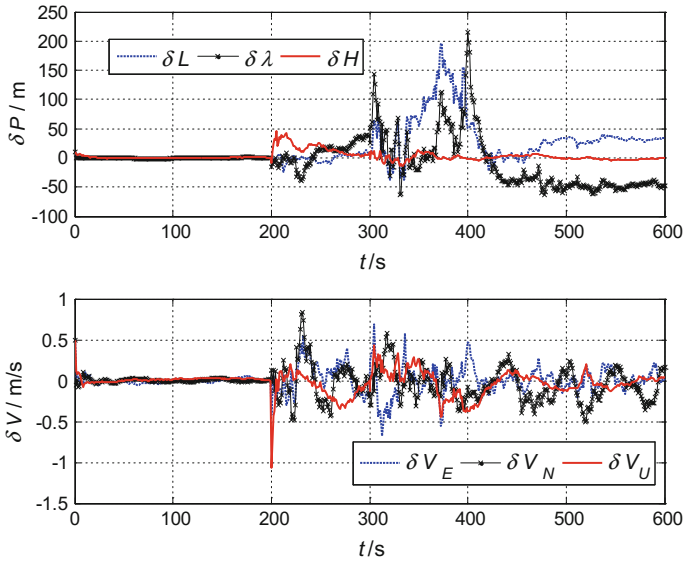


Fig. 4 Position and velocity error of SCKF in different noise

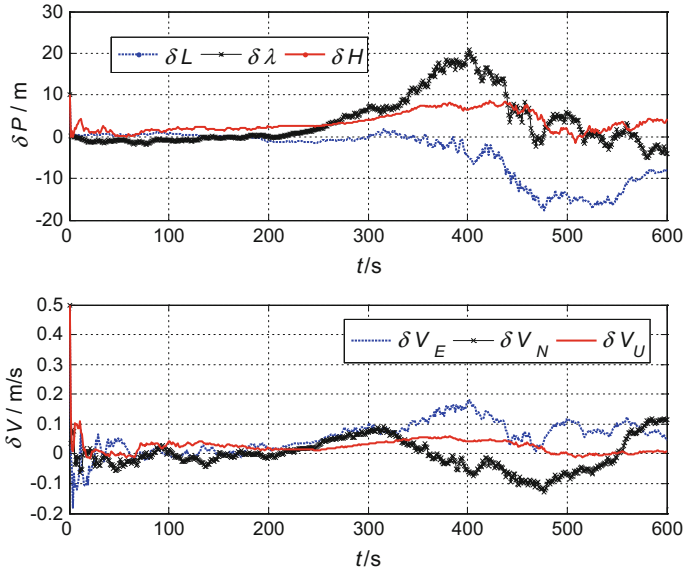


Fig. 5 Position and velocity error of this paper in different noise

Table 2 The comparison of root mean square error of SCKF of two methods in different noise

	Positioning error (m)	Velocity error (m/s)	The number of sampling points in each update time
SCKF	54.4903	0.7578	60
The method of this paper	10.7369	0.1043	32

5 Summary

According to the characteristic that the state equation of GPS/INS integrated navigation is nonlinear and the measured equation is linear, we promote an integrated navigation method based on adaptive SSR-SCKF cubature kalman filter in this paper. Firstly, we use the SSSRCR to choose the sampling points instead of the three-order spherical-radial cubature rule of SCKF to get the higher estimated accuracy. Secondly, aiming at the problem that the measured equation is linear, we propose a new measured update process based on SVD, it does not need to choose the cubature sampling points so that the computation complexity is simple. And the SVD can also improve the numerical stability of the filter. Finally, combing the moving window method, the filter can adjust the covariance matrix of measured noise in real-time, so it can improve the adaptive ability of the method proposed in this paper.

References

1. Johansen TA, Fossen TI (2016) Nonlinear observer for tightly coupled integration of pseudo-range and inertial measurements. In: IEEE transactions on control systems technology, pp 1–8
2. Sun F, Tang L (2012) INS/GPS integrated navigation filter algorithm based on cubature Kalman filter. In: Control and decision, pp 1032–1036
3. Alexander K, Dirk A (2016) Adaptive EKF-based vehicle state estimation with online assessment of local observability. In: IEEE transactions on control system technology, pp 1368–1381
4. Ivan V, Nabil A (2014) Fault diagnosis for MEMS INS using unscented Kalman filter enhanced by Gaussian process adaptation. In: 2014NASA/ESA conference on adaptive hardware and systems, pp 120–126
5. Arasaratnam I, Haykin S (2009) Cubature Kalman filters. IEEE Trans Autom Cont 54(6): 510–518
6. Chandra K, Gu D (2013) Postlethwaite I. Square root cubature information filter. IEEE Sens J 13(2):750–758
7. Wang Z, Ling H, Zhou S, Sun F (2014) The study of square root cubature Kalman smoother and its application on INS/GPS integrated navigation. In: 2014 IEEE international conference on mechatronics and automation, pp 1827–1832
8. Jiang W, Li Y, Chris R (2015) Optimal data fusion algorithm for navigation using triple integration of PPP-GNSS, INS, and terrestrial ranging system. IEEE Sens J: 5634–5644
9. Zhao Y (2015) Cubature + extended hybrid kalman filtering method and its application in PPP/IMU tightly coupled navigation systems. IEEE Sens J 15(12):6973–6985

Cramér-Rao Lower Bound for Cooperative Positioning in Non-line-of-Sight Environments

Shiwei Tian, Zhi Xiong, Guangxia Li, Jing Lv and Shi Yu

Abstract The availability of positional information is of great importance in many applications. Global Navigation Satellite System (GNSS) is the most popular positioning techniques. However, determination of devices' positions is still a challenging task in harsh environments due to radio blockage. In this situation, cooperative positioning can be utilized to improve the ability to compute position. This paper derives the expression for Cramér-Rao lower bound (CRLB) in cooperative positioning, where parts of the propagation paths between devices are non-line-of-sight (NLOS). Numerical examples illustrate our results, and the results lead to a deep understanding of the benefits of cooperation.

Keywords Cooperative positioning · GNSS · CRLB

1 Introduction

Currently, global positioning system (GPS) is the most important technology to provide location-awareness around the globe through a constellation of at least 24 satellites [1]. Unfortunately, since the received GPS signal is extremely low power, the satellite signals can be easily obstructed. For instance, GPS-based techniques fail to provide satisfactory performance due to signal blockage in many scenarios such as urban terrain, forests, or jungles. At the same time, GPS receivers are also

S. Tian · G. Li (✉) · J. Lv
College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China
e-mail: satlab_13905177686@163.com

S. Tian · Z. Xiong
College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

S. Yu
Graduate School of Chongqing Communication Institute, Chongqing 400035, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_33

susceptible to jamming and unintentional interference for this same reason [2]. Driven by the success of cooperative techniques in many research areas, cooperative techniques have been introduced in the research of positioning and navigation.

Cooperative positioning methods have been recognized as part of an effort to improve the performance of GNSS devices in hostile environments, such as accuracy, availability, integrity etc., relying on information exchange and/or direct measurements among devices. They can be used not only when GNSS is unavailable, but also in combination with GNSS [3].

It is well known that the CRLB sets a fundamental lower limit to the covariance of any unbiased estimator for a vector parameter [4]. Several previous works deal with CRLB in case of cooperative positioning (e.g., [5] for wide-band systems, [6] for wireless sensor networks, and [3] for hybrid GNSS-terrestrial systems), however, the derivation of the CRLB mainly follows that of the line-of-sight (LOS) case (the reference [5] considered the NLOS case, however, the analysis therein exploited the received waveforms rather than specific signal metrics), where an NLOS induced error due to the propagation properties of the environment is not considered. In practice, however, NLOS measurements can not be circumvented for cooperative positioning in hostile environments and must be accounted for. NLOS measurement can come from complete direct path (DP) blockage or DP excess delay, and both of them add a positive bias to the true range, resulting in a larger measured range than the true value.

Specially, related discussions given in [3, 6] can be interpreted as special cases of our results. Therefore the present analysis, while applying the same theoretical framework of [3, 6], extends it to a more general and comprehensive scenario.

The paper is organized as follows: Sect. 2 illustrates model and problem formulation; in Sect. 3 expressions for the Fisher information matrix are derived for the LOS case and the NLOS case; a numerical example is presented in Sect. 4; some conclusions are driven in Sect. 5.

2 Problem Formulation

Consider a GNSS positioning system utilizing cooperative positioning techniques, including satellite nodes with known clock bias and known position, and devices with unknown clock bias and unknown position. Generally, errors in range measurements from satellite to terrestrial devices are considered, resulting from imperfect synchronization of device clock with respect to satellites. In addition, the positive errors due to NLOS measurements between cooperative devices should be considered also.

Let M be the set of devices, S the set of satellites; denote by S_m the set of satellites device m can observe, by M_m the set of peers device m can communicate with, and M_m by the cardinality of M_m. Without loss of generality, we assume the first L measurements in the M_m device m related terrestrial measurements are

corrupted by NLOS. Positional state of satellite $s \in S$, and of device $m \in \mathrm{M}$, are indicated respectively by $\mathbf{x}_s = [x_s \ \ y_s \ \ z_s]$, $\mathbf{x}_m = [x_m \ \ y_m \ \ z_m]$. The variable c_m represents the clock bias, and b_m denotes the peer to peer NLOS measurement errors of agent m, expressed in distance units. In this paper, c_m is treated as additional parameters to be estimated as [3]. On the other hand, b_m is treated as additional noise terms, and the rationales for this assumption can be found in [7]. Similar as assumptions in [4], we assume b_m has the a prior known probability density function (pdf) $p(b_m)$, with its mean and variance denoted by $\bar{b}_m$ and κ_m^2. The information device m obtained from satellites (denoted by $\mathbf{P}_{sm}$) and from cooperative peer devices (denoted by $\mathbf{P}_m$) is denoted by $\mathbf{M}_m = \mathbf{P}_{sm} \cup \mathbf{P}_m$, a set of unknowns $\mathbf{U}_m = [\mathbf{x}_m \ \ c_m]$ for agent m are to be determined.

In this subsection, we describe the measurement models in our system. Pseudorange measurements, i.e., distance between agents and satellites can be expressed as

$$\rho_{sm} = \|\mathbf{x}_s - \mathbf{x}_m\| + c_m + v_{sm} \tag{1}$$

Under Gaussian measurement noise:

$$\ln p(\rho_{sm}|\mathbf{x}_m, c_m) = C - \frac{|\rho_{sm} - \|\mathbf{x}_s - \mathbf{x}_m\| - c_m|^2}{2\sigma_m^2} \tag{2}$$

Note that, if terrestrial anchors exist in our system, the measurement model and the following CRLB can be derived in the same way. Under the assumption that the first L range measurements are biased, range measurements between device m and its cooperative devices $n \in \mathrm{M}_m$ can be expressed as

$$r_{nm} = \begin{cases} \|\mathbf{x}_n - \mathbf{x}_m\| + b_n + v_{nm} & n = 1, \ldots L \\ \|\mathbf{x}_n - \mathbf{x}_m\| + v_{nm} & n = L+1, \ldots M_m \end{cases} \tag{3}$$

Under Gaussian measurement noise:
When $n = L+1, \ldots M_m$, the following expression can be derived:

$$\ln p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n) = C' - \frac{|r_{nm} - \|\mathbf{x}_n - \mathbf{x}_m\||^2}{2\sigma_{nm}^2} \tag{4}$$

When $n = 1, \ldots L$, we can show that

$$\begin{aligned} p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n, b_n) &= \frac{1}{\sqrt{2\pi\sigma_{nm}^2}} \exp\left[\frac{|r_{nm} - \|\mathbf{x}_n - \mathbf{x}_m\| - b_n|^2}{2\sigma_{nm}^2}\right] \\ p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n) &= \int_{-\infty}^{\infty} p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n, b_n) p(b_n) db_n \end{aligned} \tag{5}$$

where the symbol $\|\cdot\|$ denotes Euclidean distance, $d_{nm} = \|\mathbf{x}_n - \mathbf{x}_m\|$, $n, m \in \mathrm{M}$, $s \in S$; v_{sm}, v_{nm} are random Gaussian noise, with zero mean and variance σ_{sm}^2, σ_{nm}^2; C, C' are constant terms.

3 Cramér-Rao Lower Bound on the Positioning Accuracy

The CRLB is defined as the inverse of the Fisher information matrix (FIM) $\mathbf{F}$. Let $\boldsymbol{\theta} = [\theta_1 \ \ \theta_2 \ \ \ldots \ \ \theta_p]^T$ (superscript T indicates transpose) are the parameters to be estimated from the measurements, then $\mathbf{F}$ is a $p \times p$ matrix, and it can be expressed as

$$[\mathbf{F}(\boldsymbol{\theta})]_{ij} = -E\left[\frac{\partial^2 \ln p(\mathbf{X}|\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right], \quad i = 1, 2, \ldots, p; \ \ j = 1, 2, \ldots, p \tag{6}$$

where $\mathbf{X}$ are the random measurements, $p(\mathbf{X}|\boldsymbol{\theta})$ is statistical model of the random measurements, where $E[\cdot]$ indicates expected value with respect to $p(\mathbf{X}|\boldsymbol{\theta})$.

Consider the problem in this paper, the parameters to be estimated are $\mathbf{U} = [\mathbf{U}_1 \ \ \mathbf{U}_2 \ \ \ldots \ \ \mathbf{U}_M]^T$.

We focus on a single device, say m. The log-likelihood function of its measurements with respect to satellites and cooperative devices is

$$\begin{aligned}&\log p(\{\{\rho_{sm}\}_{s\in S_m}, \{r_{nm}\}_{n\in M_m}\}_{m\in M}|\mathbf{X}, \mathbf{c}) \\ &\quad = \sum_{m\in M}\sum_{s\in s_m} \log p(\rho_{sm}|\mathbf{X}_m, c_m) + \sum_{m\in M}\sum_{n\in M_m} \log p(r_{nm}|\mathbf{X}_m, \mathbf{X}_n)\end{aligned}$$

Assume that $\sum_{m\in M}\sum_{n\in M_m} \log p(r_{n\to m}|\mathbf{X}_m, \mathbf{X}_n) = \Lambda_{coop}(\mathbf{X})$.

The Fisher information matrix (FIM) can be expressed as: $\mathbf{F} = \mathbf{F}_{non-coop} + \mathbf{F}_{coop}$. Following the analysis found in [3], $\mathbf{F}_{non-coop}$ is the FIM in non-cooperative case, and can be expressed as:

$$\mathbf{F}_{non-coop} = \begin{pmatrix} \mathbf{F}_1 & & \\ & \ddots & \\ & & \mathbf{F}_m \end{pmatrix} \tag{7}$$

where $\mathbf{F}_m = \begin{pmatrix} F_{x_m} & F_{x_m,b_m} \\ F_{x_m b_m}^T & F_{b_m} \end{pmatrix}$, $F_{xm} = \sum_{s\in S_m} \frac{1}{\sigma_{sm}^2} q_{sm} q_{sm}^T$, $F_{b_m} = \sum_{s\in S_m} \frac{1}{\sigma_{sm}^2}$, $F_{x_m,b_m} = \sum_{s\in S_m} -\frac{1}{\sigma_{s\to m}^2} q_{sm}$, $q_{im} = \frac{x_i - x_m}{\|x_i - x_m\|}$.

and the FIM due to cooperative can be expressed as

$$\mathbf{F}_{coop} = \begin{pmatrix} \mathbf{F}'_1 & \mathbf{0} & \mathbf{K}_{12} & \mathbf{0} & \cdots & \mathbf{K}_{1m} & \mathbf{0} \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & & \mathbf{0}^T & 0 \\ \mathbf{K}_{21} & \mathbf{0} & \mathbf{F}'_2 & \mathbf{0} & & & \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & & & \\ \vdots & & & & \ddots & & \\ \mathbf{K}_{m1} & \mathbf{0} & & & & \mathbf{F}'_m & \mathbf{0} \\ \mathbf{0}^T & 0 & & & & \mathbf{0}^T & 0 \end{pmatrix} \tag{8}$$

Case 1: when $n = L+1, \ldots M_m$, we can show that

$$\mathbf{F}'_m = \sum_{n \in M_m} \frac{1}{\sigma_{nm}^2} \mathbf{q}_{nm} \mathbf{q}_{nm}^T \tag{9}$$

$$\mathbf{K}_{mn} = \begin{cases} -\frac{1}{\sigma_{nm}^2} \mathbf{q}_{nm} \mathbf{q}_{nm}^T, & n = L+1, \ldots M_m \\ 0, & otherwise \end{cases} \tag{10}$$

Case 2: when $n = 1, \ldots L$, we can show that

$$\begin{aligned} &\frac{\partial \ln p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n)}{\partial(\mathbf{x}_m, \mathbf{x}_n)} \\ &= \frac{1}{p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n)} \int_{-\infty}^{\infty} \left[\frac{\partial p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n, b_n)}{\partial(\mathbf{x}_m, \mathbf{x}_n)} \right] p(b_n) db_n \\ &= \int_{-\infty}^{\infty} \frac{p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n, b_n) p(b_n)}{p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n)} \left[\frac{\partial \ln p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n, b_n)}{\partial(\mathbf{x}_m, \mathbf{x}_n)} \right] db_n \\ &= \mathrm{E}_{b_n|r_{nm}, \mathbf{x}_m, \mathbf{x}_n} \left[\frac{\partial \ln p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n, b_n)}{\partial(\mathbf{x}_m, \mathbf{x}_n)} \right] \\ &= \sigma_{nm}^{-2} a_n \mathbf{q}_{nm} \end{aligned} \tag{11}$$

where

$$\begin{aligned} a_n &= \mathrm{E}_{b_n|r_{nm}, \mathbf{x}_m, \mathbf{x}_n} [r_{nm} - d_{nm} - b_n] \\ &= \int_{-\infty}^{\infty} (r_{nm} - d_{nm} - b_n) \frac{p(r_{nm}|b_n, \mathbf{x}_m, \mathbf{x}_n) p(b_n)}{p(r_{nm}|\mathbf{x}_m, \mathbf{x}_n)} db_n \end{aligned}$$

$\mathrm{E}_{b_n|r_{nm},\mathbf{x}_m,\mathbf{x}_n}[\cdot]$ denotes expectation operator with respect to the posterior pdf $p(b_n|r_{nm},\mathbf{x}_m,\mathbf{x}_n)$. Then, the FIM can be denoted as

$$\mathbf{F}'_m = \sum_{n=1}^{L} \lambda_{nm}\mathbf{q}_{nm}\mathbf{q}_{nm}^T \tag{12}$$

$$\mathbf{K}_{mn} = \begin{cases} -\lambda_{nm}\mathbf{q}_{nm}\mathbf{q}_{nm}^T, & n = 1,\ldots L \\ 0, & otherwise \end{cases} \tag{13}$$

where

$$\mathbf{K}_{mn} = \begin{cases} -\lambda_{nm}\mathbf{q}_{nm}\mathbf{q}_{nm}^T, & n = 1,\ldots L \\ -\frac{1}{\sigma_{nm}^2}\mathbf{q}_{nm}\mathbf{q}_{nm}^T, & n = L+1,\ldots M_m \\ 0 & otherwise \end{cases}$$

$$\lambda_{nm} = \sigma_{nm}^{-4} \cdot \mathrm{E}_{r_{nm}|\mathbf{x}_m,\mathbf{x}_n}\left[a_n^2\right]$$

$$\lambda_{nm} = \sigma_{nm}^{-4} \cdot \int_{-\infty}^{\infty} a_n^2 p(r_{nm}|\mathbf{x}_m,\mathbf{x}_n)dr_{nm}$$

$$\mathbf{F}'_m = \sum_{n=1}^{L} \lambda_{nm}\mathbf{q}_{nm}\mathbf{q}_{nm}^T + \sum_{n=L+1}^{M} \frac{1}{\sigma_{nm}^2}\mathbf{q}_{nm}\mathbf{q}_{nm}^T$$

Let $\mathbf{J}$ be the CRLB matrix by inversion of $\mathbf{F}$, and $\mathbf{J}_m$ be the block of $\mathbf{J}$ corresponding to device m. Then, the positioning accuracy for each device m can be decomposed into: a horizontal component,

$$\sigma_{CRLB-hor}(m) = \sqrt{\mathbf{J}_m(1,1) + \mathbf{J}_m(2,2)} \tag{14}$$

a vertical component

$$\sigma_{CRLB-vert}(m) = \sqrt{\mathbf{J}_m(3,3)} \tag{15}$$

4 Numerical Examples

For the sake of comparison, we utilize the scenario depicted in reference [3] to illustrate the previous analytical results. We assume measurement r_{12} between device 1 and device 2 is biased, and the other peer to peer measurements are unbiased. In addition, the bias pdf $p(b)$ is assumed to have the same shape as the measured bias distribution reported in [4, 7], and it can be expressed as

$$p(b) = \begin{cases} 0, & if\ b \in (-\infty, \Omega_0] \cup (\Omega_9, \infty) \\ \frac{P_i}{\Delta} & if\ b \in (\Omega_i, \Omega_{i+1}]\ \ with\ \ i = 0,\ldots,8 \end{cases} \tag{16}$$

where $\Omega_i = 0.1 + i\Delta$, and P_i is shown in Table 1, $i = 0,\ldots,8$. In addition, it can be calculated that $\bar{b} = 0.1 + 3.49\Delta$, $\kappa = 1.83\Delta$.

Table 1 Values of P_i

P_0	P_1	P_2	P_3	P_4
0.12	0.03	0.31	0.12	0.24
P_5	P_6	P_7	P_8	
0.12	0.03	0	0.03	

To illustrate the influence of bias on the CRLB, we have evaluated the CRLB from the analysis in Sect. 3 numerically, compared to the case when the measurement between device 1 and device 2 is unbiased, and discarded. The horizontal component and vertical component are shown with respect to the ratio of standard deviation of bias to that of Gaussian noise (κ/σ_{nm}) in Figs. 1 and 2 respectively.

It can be seen that when κ/σ_{nm} goes to 0, all the two components of CRLB converges to the case when r_{12} is unbiased. On the other hand, the two converges to the upper line (the case when r_{12} is discarded), when κ/σ_{nm} goes to infinity. Also, this is expected since the information from NLOS measurement reduces as κ/σ_{nm} grows larger. However, it is interesting to note that the NLOS measurement can help significantly in reducing the CRLB, especially when κ/σ_{nm} is small (NLOS measurement has small bias) or when r_{12} is large (the standard deviation of the range measurement is large compared to the bias). Therefore the range information from NLOS measurement should not be dismissed, as it can greatly improve the localization accuracy.

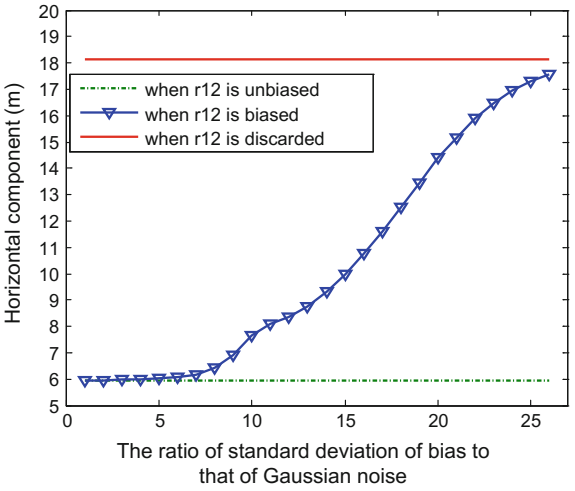


Fig. 1 Horizontal component of CRLB for device 2

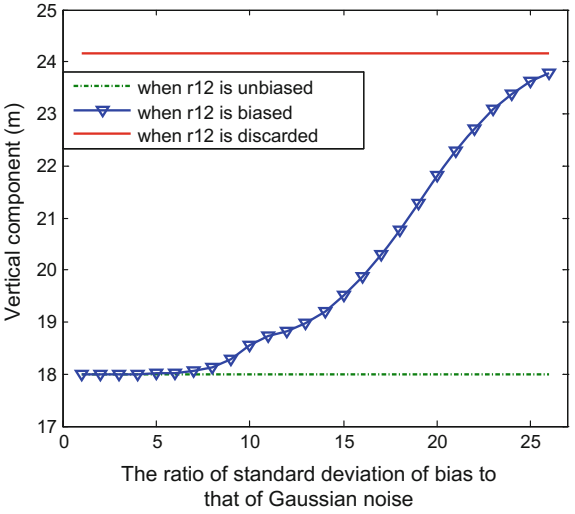


Fig. 2 Vertical component of CRLB for device 2

5 Conclusion

In this paper, we developed an expression for the Cramér-Rao bound of cooperative positioning system considering the existence of NLOS measurements (indicating that the range measurements can be positively biased), which extends the traditional relative methodology and provides new insights into the discussion of the limit on the accuracy of cooperative positioning. In particular, we studied the property of the bound in the case that which range measurement is biased and the general statistical characterization of the positive bias can be obtained a prior. Results show that the utilizing of information from NLOS measurements can be very valuable, i.e., taking the NLOS measurement into account often yields a significantly lower localization bound. These results can be used to yield important insights into the benefits of cooperation, as well as guidelines for a cooperative positioning system design.

Acknowledgements This work is supported by National Natural Science Foundation of China (61601511). Specially, we would like to thank the anonymous reviewers for their constructive comments.

References

1. David LK (2006) Global node selection for localization in a distributed sensor network. IEEE Trans Aerosp Electron Syst 42(1):113–135
2. Quebe S, Campbell J, DeVilbiss S et al (2010) Cooperative GPS navigation. In: IEEE/ION plans, pp 834–837
3. Penna F, Caceres MA, Wymeersch H (2010) Cramér-Rao bound for hybrid GNSS-terrestrial cooperative positioning. IEEE Commun Lett 14(11):1005–1007

4. Wang T (2012) Cramér-Rao bound for localization with a priori knowledge on biased range measurements. IEEE Trans Aerosp Electron Syst 48(1):468–476
5. Shen Y, Wymeersch H, Win MZ (2010) Fundamental limits of wideband localization—part II: cooperative networks. IEEE Trans Inf Theory 56(10):4981–5000
6. Larsson EG (2004) Cramér-Rao bound analysis of distributed positioning in sensor networks. IEEE Sig Process Lett 11(3):334–337
7. Jourdan DB, Dardari D, Win MZ (2008) Position error bound for UWB localization in dense cluttered environments. IEEE Trans Aerosp Electron Syst 44(2):613–628

WLAN-Aided BDS Location Algorithm

Dengao Li, Zheng Wei, Jumin Zhao, Zhiyin Ma and Ya Liu

Abstract In order to improve the positioning accuracy when the BDS signals are blocked and disturbed in complex urban environments, a WLAN-aided BDS location algorithm to fuse the location information of BDS and WLAN by federated Kalman filter is proposed in this paper. Firstly, the system state was modeled by analyzing the movement features, and the BDS observation equation is established based on pseudo-range and Doppler shift equation. Then, the WLAN observation equation is established by analyzing the fingerprinting location algorithm. Finally, the global optimum estimation of position and velocity is obtained by the principal Kalman filter. The experimental results show that this method could improve the location accuracy of BDS with better robustness in complex urban environment.

Keywords BDS · WLAN · Federated Kalman filter · Information fusion

1 Introduction

BeiDou Navigation Satellite System (BDS) is a independent-developed satellite navigation system in China, which can provide high-precision and reliability positioning, navigation and time service for a great variety of users around the world in all weather. However, in a complex urban environment, when the BDS signal is blocked and interference, there will often be a serious decline in positioning accuracy [7].

In order to solve the problem, some scholars have put forward many solutions. For example, Zhang et al. [11] merged the GPS and Zigbee locating trajectories by extended Kalman filter to improve the positioning accuracy of GPS in complex environments. Zou et al. [12] fused the position information of GPS and UWB by particle filter, which improves the positioning accuracy by 19%. However, the

D. Li (✉) · Z. Wei · J. Zhao · Z. Ma · Y. Liu
Taiyuan University of Technology, Yuci District, Jinzhong, Shanxi, China
e-mail: lidengao@tyut.edu.dn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_34

assistance of Zigbee and UWB need to re-laying the network facilities, add additional equipment, and mobile terminals which are expensive and poor universality. Deng et al. [3] fused the GPS and the base station location information by the federated Kalman filter, which effectively replenishes the GPS signal blind spot. However, the base station needs to make a big change for precise positioning, and the generalization is poor. Bi et al. [1] selected GPS or WiFi for positioning by judging the regional environment of the location point. However, due to the instability of the WiFi signal, a separate location of the WiFi positioning accuracy is low.

In order to solve the above problems, WLAN (Wireless Local Area Networks) aided BDS location algorithm is presented in this paper. Most public places in the city have WLAN deployment which is wireless and low cost. Therefore, the integration of BDS and WLAN by federated Kalman filter can increase the coverage of positioning and the positioning accuracy, which is an effective means to realize the precision positioning in complex urban environment. The position information of BDS and WLAN is fused by the Federal Kalman subsystem to obtain the local optimal estimate, and the local optimal estimate are fused by the main filter according to certain weight to obtain the global optimal estimate. The validity of the algorithm is verified by the simulation results.

2 Federated Kalman Filtering Algorithm in BDS/WLAN

2.1 The Principle of Federal Kalman Fusion

The federated Kalman filter is composed by some sub-filters and a main filter which has two data processing stages. The basic structure is shown in Fig. 1.

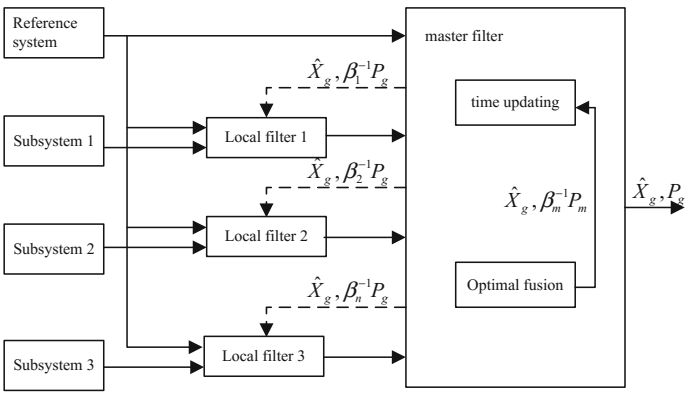


Fig. 1 General structure of federated Kalman filter

On the one hand, the output of the common reference system is transmitted directly to the main filter, and on the other hand it can be output to the sub-filters as the measured values. First each sub-filter obtains respective local estimation values $\hat{X}_i$ and a covariance matrix P_i by local filter. Then they are fused together with the estimated values of the main filter for obtaining global optimal estimates $\hat{X}_{\mathrm{g}}$ and covariance matrices P_{g}. At last they are fed back into the sub-filters by the principle of information distribution in order to weight the sub-filters [2, 4, 9].

$$\hat{X}_i = \hat{X}_{\mathrm{g}}, \quad P_i = \beta_i^{-1} P_g \tag{1}$$

In addition, the covariance matrix of the main filter error can also reset the global covariance matrix with a weighting factor of β_m^{-1}. When β_m^{-1} takes different values, different structures and different characteristics of the federated filter can be obtained. The β_i and β_m can be expressed as

$$\sum_{i=2}^{N} \beta_i + \beta_m = 1 \quad 0 \le \beta_i \le 1, \quad 0 \le \beta_m \le 1 \tag{2}$$

The Kalman filter is used to update the state estimation and covariance matrix. But the main filter does not have measure information, so the Kalman equation is just used to update the time. The measurement update is the time update. Measurement updates can be written as

$$\begin{cases} \hat{X}_{\mathrm{m},k} = \hat{X}_{\mathrm{m},k/k-1} \\ P_{\mathrm{m},k} = P_{\mathrm{m},k/k-1} \end{cases} \tag{3}$$

2.2 System Federated Kalman Filtering Model

Federated Kalman filter model of BDS/WLAN consists of two sub-Kalman filter and a master filter. The structure is shown in Fig. 2.

Firstly, the respective positioning information of BDS and WLAN is filtered for getting the local Kalman optimal estimation $\hat{X}_B$, $\hat{X}_W$ and the estimation error covariance matrix P_B, P_W by the local filters. Secondly, the main filter fuses the state information of the main filter and the local optimal estimate of the sub-filter for obtaining global optimal estimation and global estimation error. Finally, the local filters are weighted according to the principle of information distribution.

The system uses a federated Kalman filter of federated—feedback. The local filter estimates of the BDS and WLAN are fused with the estimated values of the main filter for getting the global optimal estimates $\hat{X}_g$ and covariance matrices P_g. They can be expressed as

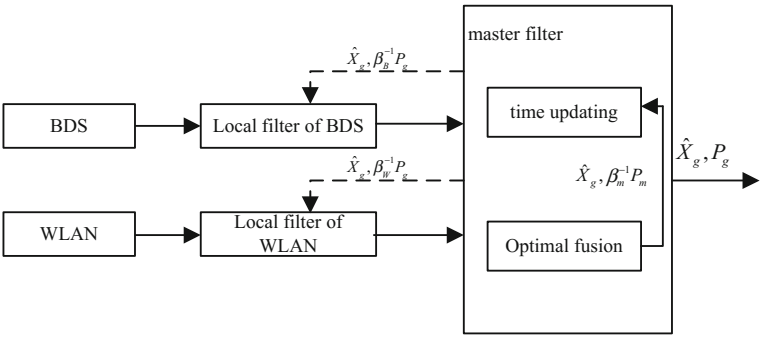


Fig. 2 BDS/WLAN federal Kalman filter model

$$\begin{cases} \hat{X}_g = P(P_B \hat{X}_B + P_W \hat{X}_W) \\ P_g = (P_B^{-1} + P_W^{-1})^{-1} \end{cases} \tag{4}$$

The global filter estimate and the local filter estimate are connected by β_i For the system, the relationship is as follows

$$X = \beta_B X_B + \beta_W X_W \tag{5}$$

where the relationship between β_B and β_W is $\beta_B + \beta_W = I$.

From (4) and (5), we know that the information distribution coefficients β_B and β_W are

$$\begin{cases} \beta_B = PP_B^{-1} \\ \beta_W = PP_W^{-1} \end{cases} \tag{6}$$

From Eq. (6), when the local estimation error P_i is larger, the information distribution coefficient of the sub-filter is smaller. So when the location information of BDS or WLAN is not accurate, its influence to the filtering result is smaller. It can effectively improve the positioning accuracy.

2.3 Establishment of System State Equation

In WLAN-Aided BDS positioning, the state vector of the system is expressed as

$$X = [\delta x, \delta y, \delta z, \delta v_x, \delta v_y, \delta v_z] \tag{7}$$

where the $\delta x, \delta y, \delta z$ are the x, y, z position error on the three directional components and $\delta v_x, \delta v_y, \delta v_z$ are the speed error in each direction. State equation is expressed as

$$X_{k+1} = AX_k + W_k \tag{8}$$

where A is state transition matrix and W represents the process noise matrix.

The mean of W is zero and the standard deviation is the Gaussian distribution. The state transition matrix A is defined as

$$A = \text{diag}[1, 1, 1, 1, 1, 1] \tag{9}$$

2.4 The Establishment of the Equation of BDS Subsystem

2.4.1 The Establishment of Measurement Equation of BDS

In the BDS subsystem, the pseudorange equation can be used to obtain the position information of the target. The Doppler shift equation is used to get the velocity information of the target. The known factors in pseudorange equations include satellite clocks $\delta t^{(s)}$, ionospheric delays $I(t)$, and tropospheric delay $T(t)$. These known factors can be used to correct the measured pseudoranges [8]. The calibration equation is expressed as

$$\rho^i = \sqrt{(x - x^i)^2 + (y - y^i)^2 + (z - z^i)^2} + \delta t_u + \varepsilon_\rho^i \tag{10}$$

where ρ^i is the pseudorange correction, (x, y, z) is the position coordinates of the receiver and (x^i, y^i, z^i) is the position coordinates of the satellite. δt_u is the receiver clock error and ε_ρ^i is pseudorange measurement noise. Equation (10) can be expressed as a linearization:

$$\rho^i = r + l^i \delta x + m^i \delta y + n^i \delta x + \delta t_u + \varepsilon_\rho^i \tag{11}$$

where r is the geometric distance between the satellite and the receiver. l^i, m^i, n^i are the direction cosine of the receiver to the satellite.

Doppler shift equations can be obtained by performing a differential operation on Formula (11). It can be expressed as

$$\lambda\phi = l^i \delta v_x + m^i \delta v_y + n^i \delta v_z + \delta t_f + \varepsilon_\phi^i \tag{12}$$

where λ is carrier phase wavelength, ϕ is satellite Doppler observations and δt_f is receiver frequency offset.

δt_u and δt_f are the unknown factors in Eqs. (11) and (12). The pseudorange and Doppler shift equations of different satellites are subtracted by single difference principle. The system of equations can be expressed as

$$\begin{cases} \rho^{ij} = (l^i - l^j)\delta x + (\mathrm{m}^i - m^j)\delta y + (\mathrm{n}^i - n^j)\delta z + \varepsilon_\rho^{ij} \\ \lambda\phi^{ij} = (l^i - l^j)\delta v_x + (m^i - m^j)\delta v_y + (n^i - n^j)\delta v_z + \varepsilon_\phi^{ij} \end{cases} \tag{13}$$

So measurement equation of BDS subsystem can be written as

$$Z_k^B = H^B X_k^B + V_k^B \tag{14}$$

where V^B is measurement noise matrix and H^B is measurement relationship matrix. The H^B can be written as

$$H^B = \begin{bmatrix} l^i - l^j & m^i - m^j & n^i - n^j & 0 & 0 & 0 \\ 0 & 0 & 0 & l^i - l^j & m^i - m^j & n^i - n^j \\ & & & \vdots & & \\ l^m - l^n & m^m - m^n & n^m - n^n & 0 & 0 & 0 \\ 0 & 0 & 0 & l^m - l^n & m^m - m^n & n^m - n^n \end{bmatrix} \tag{15}$$

2.4.2 The Establishment of Measurement Equation of WLAN

The core of WLAN positioning method which is based on the received signal strength is to convert the signal strength of the reference point to distance. In this paper, fingerprint positioning algorithm with high positioning accuracy is adopted, and the algorithm is improved.

The core of positional fingerprinting is the idea of machine learning. The technology is divided into two phases which are Offline Phase and Online Phase. The Offline Phase process establishes a fingerprint library by position data of reference point and the RSSI signal characteristic value of the access point. RSSI signal characteristic values may include RSSI, MAC address, the most value, mean, variance, direction, probability and so on. Online Phase is to match the received RSSI signal characteristic value of the target point and the fingerprint library obtained by offline sampling [5, 6, 10].

The location fingerprint library data format can be expressed as

$$\mathrm{MAP}_j(\mathrm{L}_j, \mathrm{F}_j) \tag{16}$$

where L_j is position of the reference point and the F_j is the signal characteristic value fingerprint matrix. The F_j can be written as

$$F_j = \begin{bmatrix} mac_1 & \mu_1 & d_1 \\ mac_2 & \mu_2 & d_2 \\ & \vdots & \\ mac_n & \mu_n & d_n \end{bmatrix} \tag{17}$$

where *mac* is the Media Access Control of the transmit signal point, μ is the received signal strength and d is the distance which corresponds to the received signal strength value.

Assume that the fingerprint matrix of the received RSSI signal is F_i Euclidean distance can be expressed as

$$l = \min\left\|\mathrm{F}_i - \mathrm{F}_j\right\| \tag{18}$$

The position of F_j corresponding to L_j is the target point when d takes the minimum value.

In this paper, we only use the matrix F_j at the minimum d to find the distance between the tested point and the access point. The minimum distance value is subjected to Kalman filtering. Equation (19) is derived from the distance between two points in the space.

$$d_j = \sqrt{(x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2} + \varepsilon_\rho^i \tag{19}$$

The measurement equation of WLAN subsystem is similar to that of BDS subsystem. However, there is no receiver clock δt_u in the fingerprint-based WLAN location. Equation (20) is obtained by linearizing the distance between two points in space

$$d_i = r + l_i \delta x + m_i \delta y + n_i \delta x + \varepsilon_\rho^i \tag{20}$$

Geometric distance between two points in the distance equation can be eliminated by the principle of single difference. The observation equation can be express as

$$Z_k^W = H^W X_k^W + V_k^W \tag{21}$$

where the H^W is

$$H^W = \begin{bmatrix} l^i - l^j & m^i - m^j & n^i - n^j & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ & & \vdots & & & \\ l^m - l^n & m^m - m^n & n^m - n^n & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{22}$$

3 Experiment and Result Analysis

The experimental scenario is two areas within the university campus. Area A is between Buildings 12 and 13, and area B is the Information Institute downstairs. Buildings 12 and 13 are taller and closer. What's worse, Signals in the south and north between them are severely obscured. The number of BDS visible satellites is small. However, visible satellite number is more in area B, and BDS signal is strong. Data collection was performed for two regions on October 1 and 2 at 13:00 pm, respectively.

The collected data are simulated and analyzed by MATLAB. Figures 3 and 4 show the results of positioning errors in the BDS system and the system designed in this paper.

In order to reflect the anti-gross error performance of the fusion algorithm in this paper, we add the gross error of +200 m to the 100th epoch in B region. Using the fusion algorithm designed in this paper, the simulation results are shown in Figs. 5 and 6.

The simulation results show that the location error of the BDS is larger in A region. Because the number of visible satellites is small, the signal is weak. The average error of (x, y, z) is 6.7671, 7.9234 and 11.2600 m in three directions respectively. But after the application of the federated Kalman filter, the position

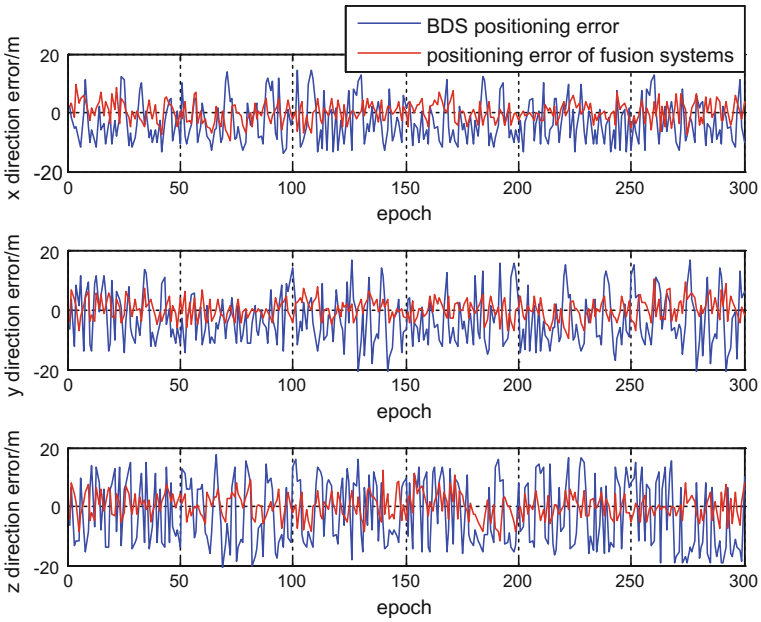


Fig. 3 A regional comparison of error curve

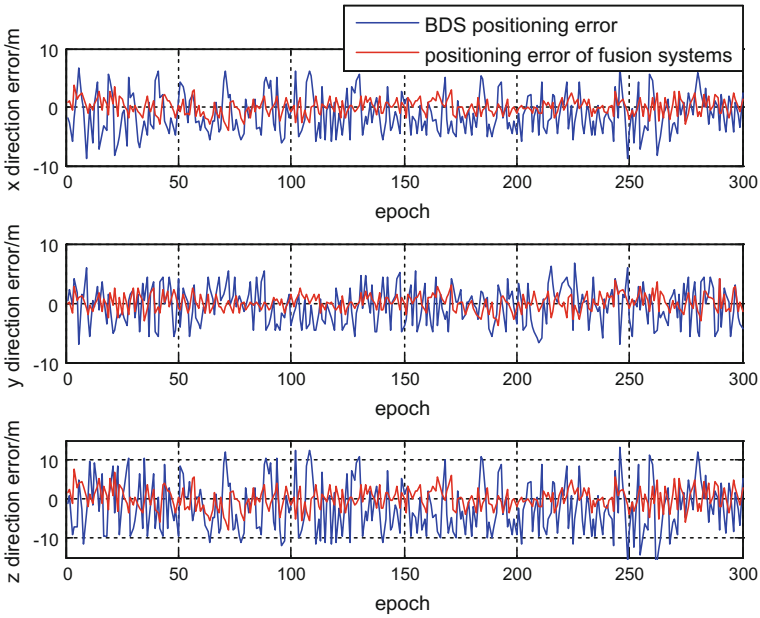


Fig. 4 B regional comparison of error curve

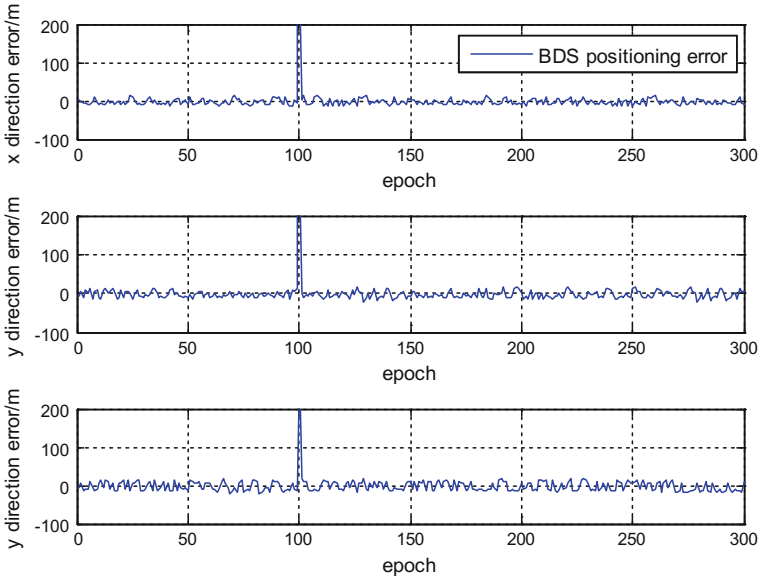


Fig. 5 The robustness of the BDS system in the B region

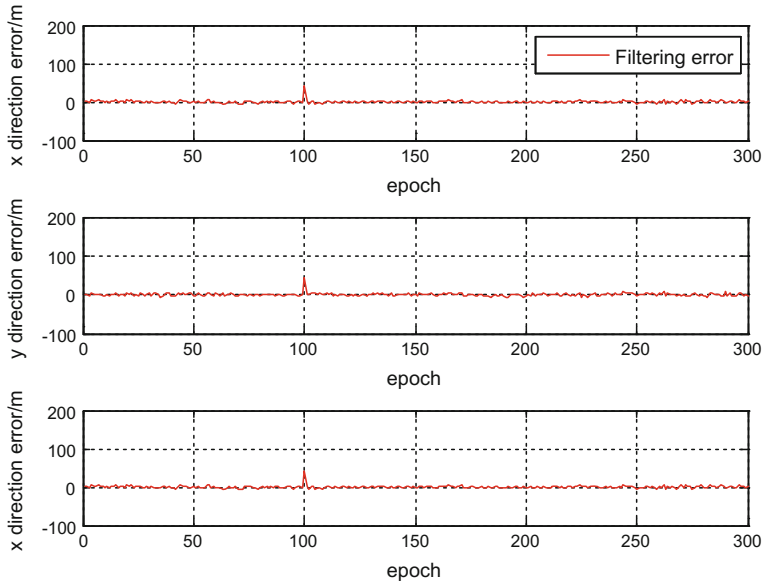


Fig. 6 The robustness of the fusion systems in the B region

accuracy is obviously improved. The mean errors of the three directions are 2.7982, 2.8150 and 3.3780 m respectively. It can be seen from Fig. 4 that BDS signal is ideal and BDS system positioning error is small in B region. The mean errors of the three directions (x, y, z) are 3.1021, 2.7028 and 3.943 m respectively. The accuracy will be improved when using the filter designed in this paper. The mean error of localization in three directions becomes 1.1193, 1.1326, 1.4052 m respectively. It can be seen from Fig. 5 that the BDS single system will produce a lot of errors when the gross error is added. But application fusion algorithm is less affected by the gross error, the simulation results shown in Fig. 6. Therefore, the proposed algorithm has good anti-gross error performance and improves the system stability.

4 Concluding Remarks

According to the characteristics of WLAN location and BDS location, a localization fusion algorithm based on federated Kalman filter is proposed. In this paper, fingerprint matching can find the distance between the point to be measured and the transmitting point to establish the distance equation. The distance equation is taken as the measurement equation of the WLAN sub-system, and the optimal estimation of the WLAN location is obtained by sub-Kalman filtering. The global optimal estimate can be obtained from the main filter by fusing the BDS and WLAN filter

estimates. The experimental results show that the proposed algorithm achieves the function of BDS location based on WLAN and improves the accuracy and robustness of the system.

References

1. Bi J (2015) Indoor and outdoor seamless positioning technology based on GPS and WiFi integrated on android smart phone. Doctoral dissertation, Shandong University of Science and Technology
2. Carlson NA, Berarducci MP (1994) Federated Kalman filter simulation results. Navigation 41 (3):297–322
3. Deng ZL, Lu Y, Lei Y (2013) Location information fusion algorithm for GPS/base-station positioning system based on federated Kalman filter. J Beijing Univ Posts Telecommun
4. Gao S, Zhong Y, Zhang X (2009) Multi-sensor optimal data fusion for INS/GPS/SAR integrated navigation system. Aerosp Sci Technol 13(4–5):232–237
5. Huang J, Millman D, Quigley M (2011) Efficient, generalized indoor WiFi graph SLAM. In: Robotics and automation IEEE, 2011, pp 1038–1043
6. Ouyang RW, Wong KS (2011) Indoor localization via discriminatively regularized least square classification. Int J Wirel Inf Netw 18(2):57–72
7. Teng Y, Shi Y, Zhi Z (2011) Research on GPS receiver positioning algorithm under bad conditions. Yi Qi Yi Biao Xue Bao/Chin J Sci Instrum 32(8):1879–1884
8. Xie G (2009) Principles of GPS and receiver design. Electronic Industry Press
9. Zhang L, Qian S, Zhang S (2016) Federated nonlinear predictive filtering for the gyroless attitude determination system. Adv Space Res 58(9):1671–1681
10. Zhu N, Zhao H, Feng W (2015) A novel particle filter approach for indoor positioning by fusing WiFi and inertial sensors. Chin J Aeronaut 28(6):1725–1734
11. Zhang W, Chen X, Li Y (2013) Seamless localization method research based on GPS and Zigbee fusion. J Ningxia Univ
12. Zou C, Zhang G, Liu Jie J et al (2012) Research on DGPS positioning accuracy seamlessly mixed with UWB technology. Transducer Microsyst Technol 31(3):74–77

Gaussian Mixture Filter Based on Variational Bayesian Learning in PPP/SINS

Qing Dai, Lifen Sui, Yuan Tian and Tian Zeng

Abstract Aiming at the problem of inaccurate estimation of parameters and noise covariance for the noise mixture model encountered when using Gaussian filter during operation of PPP/SINS integrated navigation system in the non-Gaussian noise environment, this paper proposes a Gaussian mixture adaptive filtering algorithm that is based on variational Bayesian learning. The algorithm allows accurate and efficient adaptive estimation of the parameters of Gaussian mixture filtering model based on the variational learning theory. It also further refines the Gaussian mixture filtering stochastic model, which greatly improves the filtering estimation precision; reduces the computational complexity; and effectively improves the solving efficiency. Data simulation was carried out on the PPP/SINS tightly coupled integrated navigation system. The results showed that compared to the conventional Gaussian mixture filtering algorithms, the new algorithm further improved the estimation accuracy, which was also computationally fast and less burdensome. Our results provided some theoretical support for the future application and extension of the PPP/SINS integrated filtering algorithm in the non-Gaussian noise environment.

Keywords Variational Bayesian learning · Gaussian mixture filter · Adaptive filter · Parameter estimation · PPP/SINS

1 Introduction

Modern navigation technology, which centers around Precise Point Positioning and Strapdown Inertial Navigation System (PPP/SINS), is attracting increasing attention of numerous geodesists with its remarkable flexibility and high-precision dynamic navigation and positioning performance. However, in actual practice, there exist typical non-Gaussian measurement noise environment in the navigation system, and

Q. Dai (✉) · L. Sui · Y. Tian · T. Zeng
Institute of Geospatial Information, Information Engineering University, Zhengzhou, China
e-mail: Geo_DaiQing@hotmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_35

approximation of filter model with Gaussian white noise will degrade the stability and positioning precision [1]. Some scholars used the Gaussian mixture model (GM) for non-Gaussian noise modeling, but the estimation of model parameters was rather complex [2]. At present, studies on parameter estimation of Gaussian mixture model are concentrated mainly on maximum likelihood estimation (ML), Markov chain Monte Carlo (MCMC), expectation maximization (EM), etc. Among them, ML and EM algorithms cannot adaptively estimate the model order and are prone to data over-fitting. MCMC algorithm has large computational complexity and slow convergence rate, which is unsuitable for engineering applications with high real-time requirements. On the other hand, the inaccuracy of Gaussian mixture model parameters will directly affect the estimation accuracy of Gaussian mixture filtering. To solve the problem of unknown or inaccurate measurement noise covariance in filters, some scholars have proposed the optimal Bayesian filter that contains measurement noise. However, in practical application, it faces problems such as complex integrand and high variable dimension. Moreover, the introduction of noise variance matrix makes its solution more difficult, thus affecting its popularization [4].

Variable Bayesian learning (VB) originated from machine learning theory has been proposed by Beal, which solves the complexity of integrand calculation for the Bayesian estimation in the practical application. It is suitable for high dimensional observation environment. By obtaining the suboptimal solution through approximation, it enables high precision parameter estimation with less computational complexity; accelerates the convergence; and effectively improves the estimation speed [5]. VB learning has made some progress in the fields like stochastic system parameter estimation, blind source separation, speech recognition and independent component analysis [6].

Inspired by the above research, this paper further refines the Gaussian mixture filter model by considering the non-Gaussianity of random errors in the PPP/SINS integrated navigation system. A VB learning-based Gaussian mixture adaptive filtering scheme is proposed, which simultaneously solves the problem of mixture model parameter estimation in non-Gaussian measurement noise and the self-adaptation problem under inaccurate sub-filter noise covariance by utilizing the VB learning theory. Simulation results verify that the new algorithm can improve the precision, real-timeliness and robustness of the PPP/SINS tightly coupled navigation system, which provide some theoretical support for its future application and extension in the precise integrated navigation field.

2 PPP/SINS Tightly Coupled System

In the PPP/SINS tightly coupled navigation system, the GNSS receiver provides pseudorange and carrier observation information. After ionosphere-free correction, precision ephemeris and clock offset pre-processing, the corrected GNSS observations are obtained. Meanwhile, equivalent GNSS observations are derived based

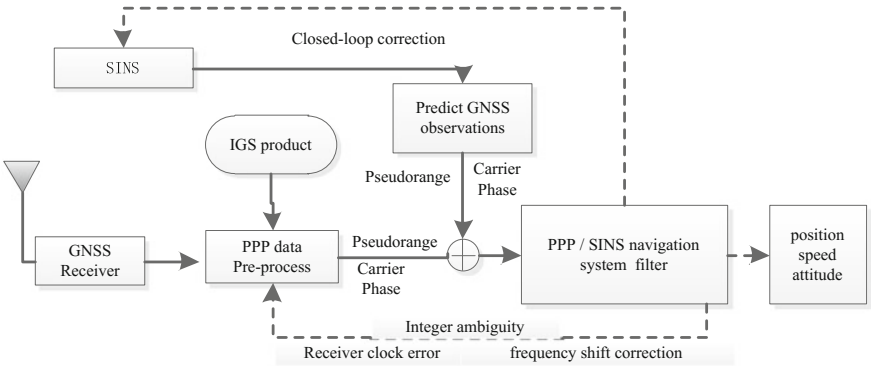


Fig. 1 PPP/SINS tightly coupled navigation system

on the SINS output information of position and velocity combined with precise ephemeris calculation. The two sets of observations together provide the measurement input for the filtering algorithm, while the output error estimators are closed-loop corrected for SINS and GNSS respectively [7, 8] (Fig. 1).

2.1 System State Model

State model of PPP/SINS tightly coupled navigation system

$$X_k = \mathrm{F}_{k|k-1} X_{k-1} + W_k \tag{1}$$

$$F = \begin{bmatrix} F_{INS} & \\ & F_{GNSS} \end{bmatrix} \tag{2}$$

$$F_{GNSS} = \begin{bmatrix} M_w & 1 & 0 & 0_{1\times n} & 0_{1\times n} \\ M_w & 1 & 0 & 0_{1\times n} & 0_{1\times n} \\ M_w & 1 & 0 & 0_{1\times n} & 0_{1\times n} \\ M_w & 1 & 0 & 0_{1\times n} & 0_{1\times n} \end{bmatrix} \tag{3}$$

where $X = [X_{INS}\, X_{GNSS}]^T$. X_{INS} includes the platform error angle φ; velocity error v; position error p; gyro random constant error ε_b; gyro first-order Markov random noise vector ε_r; and acceleration first-order Markovian error b_a. X_{GNSS} includes the tropospheric wet component d_w; equidistant distances d_t and $\dot{d}_t$ resulting from receiver clock offset and frequency; and the integer ambiguity N_i. F_{INS} denotes the state transition matrix for INS, which is set up specifically referring to [9]. F_{GNSS} denotes the state transition matrix corresponding to the GNSS state quantities; M_w denotes the wet component of tropospheric delay; and W_k is the system noise.

2.2 System Measurement Model

After precise ephemeris and clock offset processing of the GNSS receiver's pseudorange and carrier phase observation information, they are estimated together with tropospheric wet component delay to derive the ionospheric-free observation equations for pseudorange and carrier phase

$$P_{G_i} = \rho_i - Cd_t + M_w \cdot d_w + v_{p_i} \tag{4}$$

$$\varphi_i = \rho_i - Cd_t + M_w \cdot d_w - \lambda \cdot N_i + v_{\varphi_i} \tag{5}$$

where ρ_i represents the geometrical real distance from the i-th satellite to the GNSS receiver; $\lambda \cdot N_i$ is the ambiguity; and v_{p_i}, v_{φ_i} are the observation errors of pseudorange and carrier phase, respectively.

The carrier phase, pseudorange output by GNSS and SINS predictions together constitute the measurement input of the PPP/SINS tightly coupled observation model

$$Z = \begin{bmatrix} P_j^{GNSS} - P_j^{SINS} \\ \varphi_j^{GNSS} - \varphi_j^{SINS} \\ \vdots \end{bmatrix} \tag{6}$$

where P_j^{GNSS}, φ_j^{GNSS} correspond to the j-th satellite's ionosphere-free pseudorange and carrier phase observations in GNSS; and P_j^{SINS}, φ_j^{SINS} correspond to the j-th satellite's ionosphere-free pseudorange and carrier phase observations estimated by SINS. Hence, we get the tightly coupled observation model of the PPP/SINS system:

$$Z_k = HX_k + e_k \tag{7}$$

where H is the observation matrix; and e_k is the measurement noise part.

2.3 Gaussian Mixture Filter

When a measurement noise e_k does not conform to the Gaussian white noise distribution, it can be approximated using the Gaussian mixture model:

$$p(\overline{e}_k) = \sum_{i=1}^{m} \omega_k^i \times N\left[\mu_{e_k}^i, \sum_k^i\right] \tag{8}$$

$$\sum_{i=1}^{m} \omega_k^i = 1, \quad \omega_k^i \geq 0 \tag{9}$$

by parallel estimation with multiple extended Kalman filters (EKF) that satisfy the white noise requirements, the optimal posterior probability distribution of state parameters are obtained:

$$P\left(X_k | L^k\right) = \sum_{i=1}^{m} \omega_k^i N\left[X_k - \widehat{X}_k^i, \sum_{\widehat{X}_k^i}^{i}\right] \tag{10}$$

where $\widehat{X}_k^i$ and $\sum_{\widehat{X}_k^i}^{i}$ represent the state estimator and covariance of sub-filter; and ω_k^i denotes the corresponding weight. The weights of Gaussian components in the Gaussian mixture filters are subject to innovation-based adaptive adjustment, which makes the system robust. However, estimation of unknown parameter $\left\{\omega_i, \mu_{e_k}^i, \sum_k^i\right\}_{i=1}^{m}$ in the Gaussian mixture model is a crucial issue. Meanwhile, high-dimensional observations in the tightly coupled navigation further complicate the computation. Additionally, when the estimation accuracy of mixture model parameters is affected, the estimation effect of Gaussian mixture filtering will also be restricted directly.

3 VB Learning

The basic idea of VB learning is to transform the complicated multidimensional numerical integration problem into the optimization problem of approximate distribution, which is a good approximation optimization algorithm. In the Bayesian estimation criteria, the observation data set is Y; Θ is a set of d number of unknown model parameters $\Theta = [\theta_1\, \theta_2 \ldots \theta_d]$; and Z is a hidden variable. When the posterior distribution $p(\Theta, Z|Y)$ is complex in form, and analytical expression is difficult to obtain, the new distribution $q(\Theta, Z)$ can be approximated by minimizing the KL divergence into a real posterior distribution $p(\Theta, Z|Y)$. According to the theory of variational function, the expression of general solution is obtained

$$q(\theta_i) = \frac{\exp\left(E_q(\theta_{j\neq i})(\ln p(Y, \Theta, Z))\right)}{\int \exp\left(E_q(\theta_{j\neq i})(\ln p(Y, \Theta, Z))\right) d\theta_i} \tag{11}$$

The properties of conjugate exponential domain function allow the distribution form of each calculated $q(\theta_i)$ to be the same as the real posteriori. We can find from the above formula that the general solution expression is an implicit solution, and each calculation of θ_i is related to $\theta_{j\neq i}$. So iterative calculation is needed

$$q_m(Z) \propto \exp\left[E_{q(\theta)}[\ln p(Y, Z, \theta)]\right] \tag{12}$$

$$q_m(\theta) \propto \exp\left[E_{q(z)}[\ln p(Y, Z, \theta)]\right] \tag{13}$$

After M times of iteration, then the convergence $|F_M(q(\theta)) - F_{M-1}(q(\theta))| \leq t$ is calculated [4–6]. Figure 2 illustrates the flow of VB learning.

3.1 VB-Based Parameter Estimation for Gaussian Mixture Model

Based on the VB learning method, this paper designs a fast and effective algorithm for Gaussian mixture (GM) parameter estimation. Suppose that $\pi = [\omega_k^1, \omega_k^2, \ldots, \omega_k^n], \mu = [\mu_{e_k}^1, \mu_{e_k}^2, \ldots, \mu_{e_k}^n], \Lambda = [\sum_k^1, \sum_k^2, \ldots, \sum_k^n]$ According to the VB procedure described in Fig. 2.

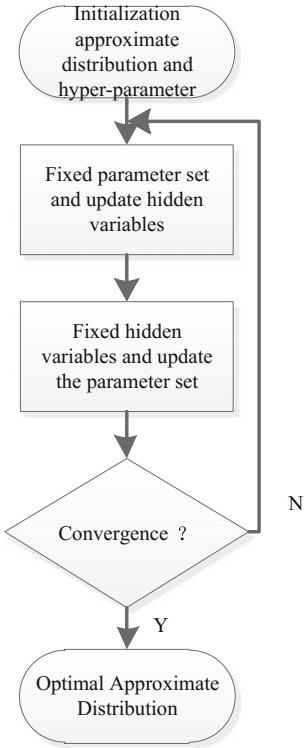


Fig. 2 Variational Bayesian learning algorithm

3.1.1 Initialization of Hyper-parameters $\alpha_0, \beta_0, v_0, m_0, w_0$

After initialization, the order of the mixture model is assumed to be T.

3.1.2 Fixed Parameter Set $\alpha_t, \beta_t, v_t, m_t, v_t$ and Updated Hidden Variable Z

The approximate posterior distribution of hidden variable Z is obtained from Eq. (12)

$$\begin{aligned} q(Z) \propto \exp\Bigg\{ & \sum_{n=1}^{N}\sum_{t=1}^{T} Z_{nt}\left[\frac{1}{2}E_{\Lambda_t}\ln|\Lambda_t| - \frac{D}{2}\ln 2\pi\right. \\ & \left. -\frac{1}{2}E_{\Lambda_t,\mu_t}[(y_n-\mu_n)^T\Lambda_t(y_n-\mu_n)] + E_{\pi_t}[\ln\pi_t]\right]\Bigg\} \\ = & \prod_{n=1}^{N}\prod_{t=1}^{T}\gamma_{nt}^{Z_{nt}} \end{aligned} \tag{14}$$

where

$$E_{\Lambda_t}\ln|\Lambda_t| = \sum_{i=1}^{D}\psi\left(\frac{v_t+1-i}{2}\right) + D\ln 2 + \ln|w_t| \tag{15}$$

$$E_{\Lambda_t,\mu_t}[(y_n-\mu_n)^T\Lambda_t(y_n-\mu_n)] = v_t(y_n-m_t)^T w_t(y_n-m_t) + D\beta_t^{-1} \tag{16}$$

$$E_{\pi_t}[\ln\pi_t] = \psi(\alpha_t) - \psi\left(\sum_{t=1}^{T}\alpha_t\right) \tag{17}$$

where $\psi(\alpha_t)$ is a double gamma function.

3.1.3 Fixed Hidden Variable Z and Updated Parameter Set $\alpha_t, \beta_t, v_t, m_t, v_t$

The following is obtained from Eq. (13)

$$\begin{aligned} q(\pi,\Lambda,\mu) \propto \exp\Bigg[& \sum_{n=1}^{N}\sum_{t=1}^{T} E_{Z_{nt}}[Z_{\mathrm{nt}}](\ln N(y_n|\mu_t,\Lambda_t^{-1}) + \ln\pi_t \\ & + \ln Dir(\pi|\alpha_t)] + \ln\pi_t) + \ln Dir(\pi|\alpha_t)\Bigg] \end{aligned} \tag{18}$$

where $E_{Z_{\mathrm{nt}}}[Z_{\mathrm{nt}}] = \gamma_{\mathrm{nt}}$. The approximate distribution of Eq. (18) is

$$q(\pi) = C\prod_{\mathrm{t}=1}^{T}\pi_k^{\left(\alpha_0+\sum_{n=1}^{N}\gamma_{nt}-1\right)} = Dir(\pi|\alpha) \tag{19}$$

In the formula set $N_t = \sum_{n=1}^{N}\gamma_{nt}, \overline{y}_t = \frac{1}{N_t}\sum_{n=1}^{N}\gamma_{nt}y_{\mathrm{n}}, S_t = \frac{1}{N_t}\sum_{n=1}^{N}\gamma_{nt}\ (y_n-\overline{y}_t)(y_n-\overline{y}_{\mathrm{t}})^T, \alpha_t = \alpha_0 + N_t$

Then

$$\begin{aligned} q(\mu_t|\Lambda_t) &\propto \exp\left[-\frac{1}{2}\left(\mu_t^T\beta_t\Lambda_t\mu_t - m_t^T\beta_t\Lambda_t\mu_t - \mu_t^T\beta_t\Lambda_t m_t + m_t^T\beta_t\Lambda_t m_t\right)\right] \\ &= N\left(\mu_t|m_t,(\beta_t\Lambda_t)^{-1}\right) \end{aligned} \tag{20}$$

In the formula set $\beta_t = \beta_0 + N_t$, $m_t = \frac{\beta_0 m_0 + N_t\overline{y}_t}{\beta_0+N_t}$

where

$$\begin{aligned} q(\Lambda_t)\propto\exp\Bigg\{\frac{1}{2}Tr\Bigg[\Bigg[w_0^{-1}+\beta_0(\mu_t-m_0)(\mu_t-m_0)^T+\sum_{n=1}^{N}E_{Z_{nt}}[Z_{nt}](y_n-\mu_t)(y_n-\mu_t)^T \\ -\beta_t(\mu_t-m_t)(\mu_t-m_t)^T\Bigg]\Lambda_t+\sum_{n=1}^{N}E_{Z_{nt}}[Z_{nt}](y_n-\mu_t)(y_n-\mu_t)^T-\beta_t(\mu_t-m_t)(\mu_t-m_t)^T\Bigg]\Lambda_t \\ +\frac{1}{2}\left[\sum_{n=1}^{N}E_{Z_{\mathrm{nt}}}[Z_{\mathrm{nt}}]+v_0-D-1\right]\ln|\Lambda_t|\Bigg\} = W(\Lambda_t|v_t,w_t) \end{aligned} \tag{21}$$

In the formula set $v_t = v_0 + N_t$, where

$$w_t^{-1} = w_0^{-1} + N_tS_t + \frac{\beta_0 N_t}{\beta_0+N_t}(\overline{y}_t - m_0)(\overline{y}_t - m_0)^T$$

3.1.4 Determination of Lower Convergence Bound

From the joint density function of approximate solution

$$q(Z,\pi,\Lambda,\mu) = q(Z)q(\pi)\prod_{t=1}^{T}q(\mu_t|\Lambda_t)q(\Lambda_t) \tag{22}$$

$F(q(\theta,Z))$ is substituted to calculate ΔF, and threshold is set.

3.2 VB-Based Self-adaptive EKF Algorithm

As a parameter estimation algorithm for multivariate statistical analysis, VB learning can simultaneously estimate the mean and variance of measurement noise when they are unknown or inaccurate. And VB-based EKF (VBEKF) algorithm is designed.

3.2.1 Time Prediction

$$\begin{cases} \widehat{x}_{k|k-1} = F_{k,k-1}\widehat{x}_{k-1} \\ P_{k|k-1} = F_{k,k-1}P_{k-1}F_{k,k-1}^T + Q_{k-1} \end{cases} \tag{23}$$

3.2.2 Measurement Updating

$$\begin{cases} K_k = P_{k|k-1}H_k^T\left(H_kP_{k|k-1}H_k^T + \widehat{R}_{k-1}\right)^{-1} \\ \widehat{x}_k = \widehat{x}_{k|k-1} + K_k\left(y_k - H_k\widehat{x}_{k|k-1} - \widehat{\mu}_{k-1}\right) \\ P_k = P_{k|k-1} - K_kH_kP_{k|k-1} \end{cases} \tag{24}$$

3.2.3 Estimation of Measurement Noise Mean and Variance

$$\overline{v}_k = \frac{1}{k}\sum_{j=1}^{k}\left(y_j - H_j\widehat{x}_j\right) \tag{25}$$

$$s_k = \sum_{j=1}^{k}\left(\overline{v}_k - \left(y_j - H_j\widehat{x}_j\right)\right)_{\cdot}^2 + \sum_{j=1}^{k} diag\left(H_jP_jH_j^T\right) \tag{26}$$

$$\widehat{\mu}_k = \frac{k_0\mu_0 + k\overline{v}_k}{k_0 + k} \tag{27}$$

$$\widehat{\alpha}_k = \alpha_0 + \frac{k}{2} \tag{28}$$

$$\widehat{\beta}_k = \beta_0 + \frac{1}{2}s_k + \frac{k_0k}{k_0 + k}\frac{\left(\mu_0 - \overline{v}_k\right)_{\cdot}^2}{2} \tag{29}$$

$$\widehat{R}_k = diag\left(\widehat{\beta}_{k\cdot}/\widehat{\alpha}_k\right) \tag{30}$$

where μ_0 is the initial value; $\widehat{\mu}_k$ is the mean estimator of measurement noise; $\widehat{R}_k$ is the covariance estimator; α_0, β_0 are the initial noise variance parameters; and k_0 can be taken as 1. Compared to the residual-based self-adaptive filtering method which

achieves summation operation through the memory of innovation series in the sliding window, the estimation of the present algorithm is related only to the previous estimator and is easier to iterate.

4 Simulation and Analysis

PPP/SINS tightly coupled simulation system was set up. There were 5 visible satellites in the simulation. The initial position of motion carrier was: longitude 125°, latitude 50°, height 300 m, initial misalignment angle (5° 5° 20°), and simulation time last 1000 s. Settings of simulation parameters for IMU are listed in Table 1

4.1 Performance Analysis of VB-Based Parameter in Gaussian Mixture Model

To verify the performance of the GM model parameter estimation algorithm proposed in this paper, non-Gaussian noise stochastic model of GNSS observations can be approximated by the Gaussian mixture model of two components referenced the idea in [10]. The pseudorange observation error data of five satellites following Gaussian mixture distribution were simulated. The parameters of the mixture model were set as: $\pi = [0.2\,0.8]$, $\mu = [\mu_{e_k}^1, \mu_{e_k}^2]$, $\Lambda = [\sum_1, \sum_2]$. Where $\mu_{e_k}^1 = \mu_{e_k}^2 = [0\,0\,0\,0\,0]$, $\sum_1 = \mathrm{diag}(0.04\,0.04\,0.04\,0.04\,0.04)$, $\sum_2 = \mathrm{diag}(0.08\,0.08\,0.08\,0.08\,0.08)$. K was set as 3.

Two effective components were obtained after the computation of convergence. Table 2 lists the VB-based estimation results of GM parameters, as well as the MCMC estimation results at a cycle number of 50 with sample size of 1000 and 5000 respectively. Comparison found that the estimation accuracy of GM parameters based on VB learning proposed herein was overall better than that of MCMC method with a size of 1000; and slightly lower than the MCMC method with a size of 5000. Nonetheless, it took the shortest time. It was thus clear that the VB algorithm has the advantages of low computational complexity, fast speed and good accuracy in the estimation of mixture model parameters. It can effectively identify the statistical properties of non-Gaussian noise in the high-dimensional data and make use of them, thereby improving the noise processing ability. So it has better engineering application prospects than the sampling method.

Table 1 Simulation parameters of IMU

Parameter	Gyroscope	Accelerometer
Zero bias	20°/h	50 mg
Random walk	0.07°/h/sqrt(Hz)	6 mg/h/sqrt(Hz)

Table 2 Parameter estimation of Gaussian mixture model

Estimation method	Component	Weight	Mean	Covariance	Time
Variation estimation	1	0.19	[0.01 0.02 0.01 0.02 0.01]	diag(0.04 0.04 0.04 0.03 0.03)	Approximately 19 s
	2	0.76	[0.03 0.02 0.04 0.01 0.02]	diag(0.07 0.08 0.08 0.09 0.08)	
MCMC (1000 size)	1	0.14	[0.11 0.15 0.01 0.08 0.04]	diag(0.05 0.04 0.06 0.02 0.01)	Approximately 20 s
	2	0.85	[0.09 0.07 0.02 0.05 0.05]	diag(0.07 0.06 0.08 0.09 0.08)	
MCMC (5000 size)	1	0.18	[0.03 0.01 0.01 0.02 0]	diag(0.04 0.04 0.04 0.03 0.02)	Approximately 180 s
	2	0.81	[0.01 0.02 0.02 0.02 0]	diag(0.08 0.08 0.08 0.08 0.08)	

4.2 VB-Based Adaptive Filtering Performance Analysis

When the covariance of the satellite measurement data is unknown or inaccurate, adaptive filtering is needed to make compensation. To verify the performance of adaptive filtering based on VB, the idea in [8] was referenced. Despite varying degrees of jamming received by satellites in the actual observation, the variation trends of measurement noise variances were the same. Therefore, drastic covariance variation of PRN-8 and PRN-29 satellite measurement noises was set between 600 and 800 s. The simulation results are shown in the figures below (Table 3).

The attitude, speed and position error estimation curves given in Figs. 3, 4 and 5, the dashed blue line represents the residual based AKF algorithm, while the red solid line represents the VBEKF algorithm. As can be seen, the filtering precisions of VBEKF and AKF are close. Figure 6 shows when the variance changed in the 600 s, VBEK can still track variance. Therefore, the two algorithms can both suppress the decrease of estimation accuracy when the satellite measurement noise variance is inaccurate or jumps. However, VBEKF occupies less computing resources than the AKF since it does not need to store the residual or innovation sequence, which is more suitable for iterative computation. So it is less time-consuming and more suitable for real-time computation of output.

Table 3 Computing time comparison of three filters

Scheme	Computation time (s)
AKF	25.3
VBEKF	19.8

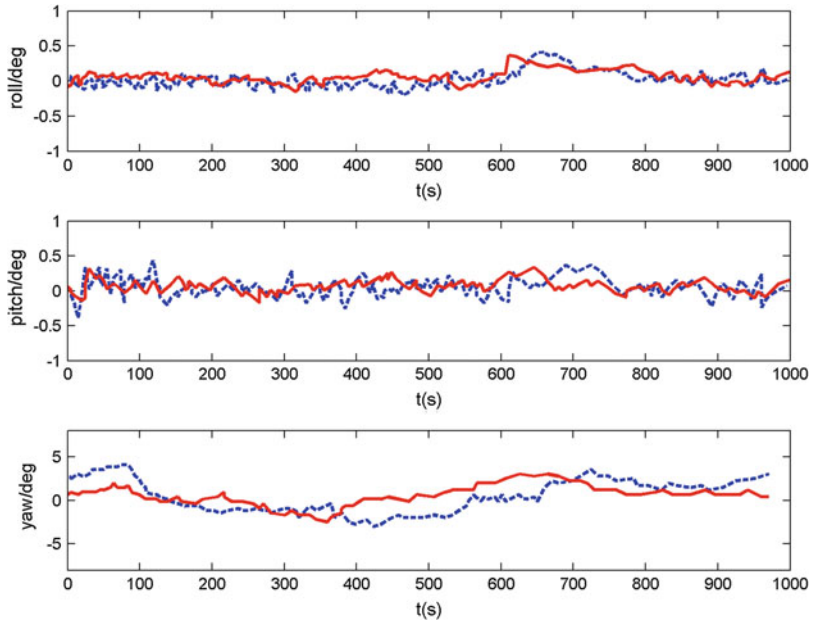


Fig. 3 Curve of estimated attitude error

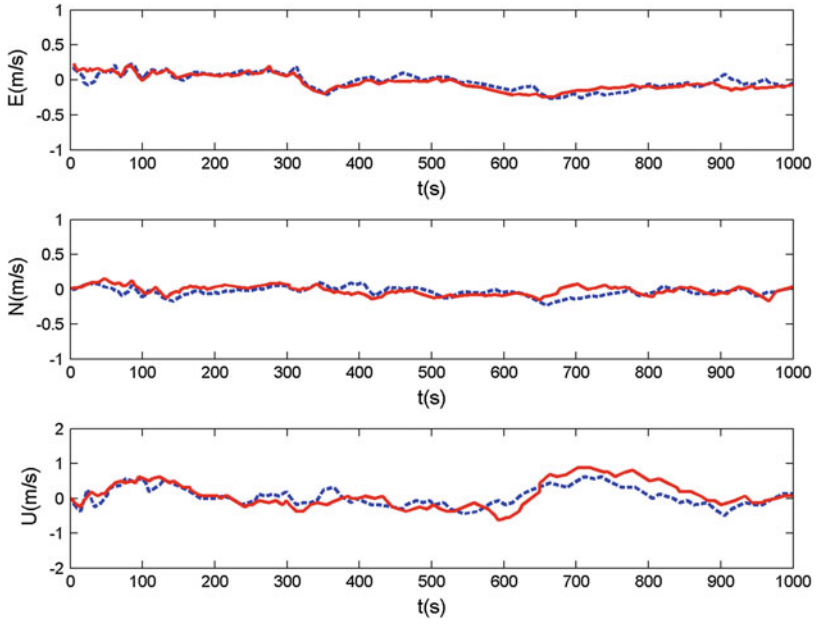


Fig. 4 Curve of estimated speed error

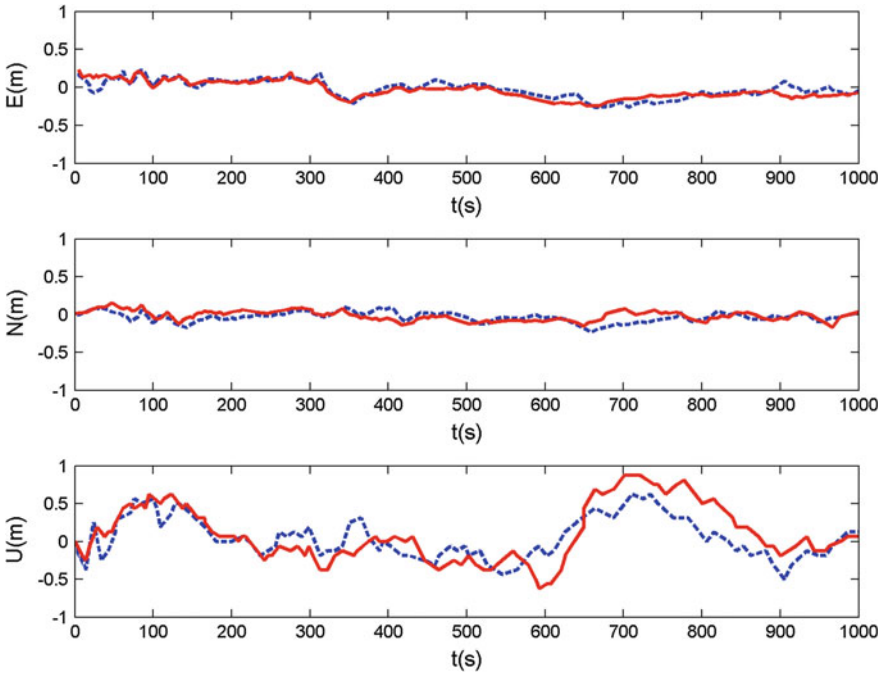


Fig. 5 Curve of estimated position error

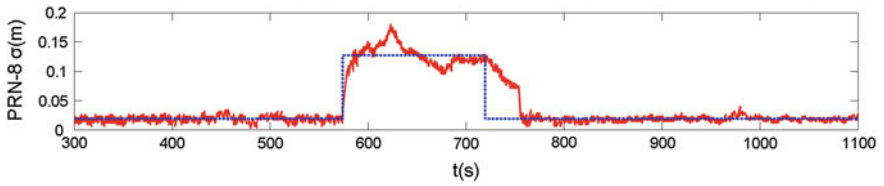


Fig. 6 Estimated variances for PRN8 pseudorange noise

4.3 VB-Based Performance Analysis of Gaussian Mixture Filtering Algorithm

This experiment considers the performance of PPP/SINS system under the non-Gaussian noise influenced by Gaussian mixture adaptive filter on VB learning. From the model

$$0.9\mathrm{N}\left(0, 0.5^2\right) + (1 - 0.1)\mathrm{N}\left(0, 0.8^2\right)$$

Table 4 Estimation errors (RMS) and simulation time

	Scheme 1	Scheme 2	Scheme 3
Attitude error (°)	0.22	0.11	0.08
	0.21	0.12	0.13
	3.01	2.75	2.67
Speed error (m/s)	0.45	0.33	0.21
	0.31	0.28	0.22
	0.34	0.31	0.32
Position error (m)	0.21	0.19	0.13
	0.34	0.28	0.23
	0.88	0.69	0.41
Simulation time (s)	48	129	85

non-Gaussian noise was generated, and the noise variance changes were set at 600 s. Three filtering schemes were designed in the experiment (Table 4).

Scheme 1: Extended Kalman filter (EKF)
Scheme 2: VB-based Gaussian mixture filtering (VBGM-EKF)
Scheme 3: VB-based Gaussian mixture adaptive filtering (VBEM-VBEKF).

Experimental results were analyzed. As shown in Fig. 7, the estimation accuracy of conventional EKF algorithm decreased sharply due to imprecise modeling of measurement noise resulting from the addition of non-Gaussian noise, which sometimes even exceeded the 3σ error limit. As shown in Fig. 8, VBGM-EKF and VBGM-VBEKF could effectively identify the non-Gaussian noise parameters and make use of them. These algorithms could also improve the filter stability by way of Gaussian mixture filtering and enhance the filtering precision obviously. VBGM-EKF exhibited slightly lower precision than the VBGM-VBEKF algorithm. This was because VBGM-VBEKF could well track the variances and improves the algorithm robustness when the measurement noise covariance estimation was inaccurate or the noise variance changed abruptly. To sum up, in the PPP/SINS tightly coupled navigation system under the non-Gaussian noise environment, the novel algorithm proposed in this paper can ensure the stability and precision of filters, which is more suitable for engineering computing applications.

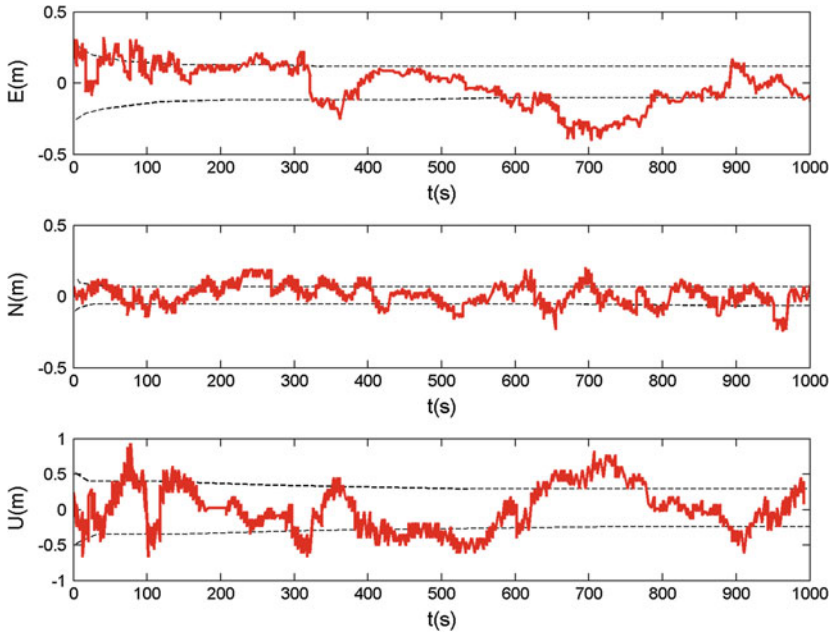


Fig. 7 Position error curve and 3σ limit (scheme 1)

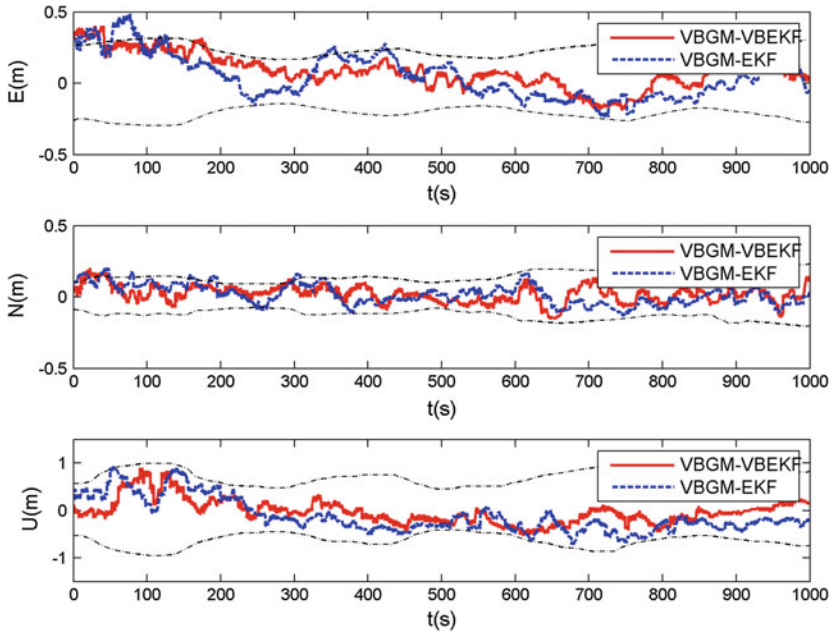


Fig. 8 Position error curve and 3σ limit (schemes 2 and 3)

5 Conclusion

To further improve the filter precision of PPP/SINS tightly integrated navigation system, the stochastic model under non-Gaussian measurement noise influence is refined. In this paper, VB learning theory is utilized to estimate the mixture model parameters; and to track the filter variances. A VB-based Gaussian mixture adaptive filtering algorithm is proposed. The simulation results show that the proposed method can effectively implement precise adaptive estimation of mixture model parameters and the adaptive sub-filter estimation under inaccurate or abruptly changing measurement variance compared to the conventional Gaussian mixture filters. It has better stability and robustness under the non-Gaussian noise condition, as well as higher filtering precision. It is also computationally less burdensome as it applies the VB learning principle, which has better practical value in the engineering field. In addition, future research works will focus on further enhancing the real-timeliness and accuracy of the algorithm.

References

1. Yi C (2011) Research theory and application of real time precise point positioning. Central South University
2. Zhu H, Han Y (2006) Bayesian multivariate statistical inference theory. Science Press, Beijing, pp 2–5
3. Shikawa Y, Takeuchi L, Nakanob R (2010) Multi-directional search from the primitive initial point for Gaussian mixture estimation using variational Bayesian method. Neural Netw 23 (3):356–364
4. Vrettas MD, Cornford D, Opper M (2011) Estimating parameters in stochastic systems: a variational Bayesian approach. Physica D 240(23):1877–1900
5. Beal MJ (2003) Variational algorithms for approximate Bayesian inference. University College London
6. Shen Y, Cornford D (2012) Variational Markov chain Monte Carlo for Bayesian smoothing of nonlinear diffusions. Comput Stat 27(1):149–176
7. Rabbou MA, El-Rabbany A (2015) Integration of GPS precise point positioning and MEMS-based ins using unscented particle filter. Sensors 15(4):7228–7245
8. Du SH, Gao Y (2012) Inertial aided cycle slip detection and identification for integrated PPP GPS and INS. Sensors 12(11):14344–14362
9. Fu M, Deng Z, Yan L (2010) Kalman filtering theory and its application in navigation system, 2nd edn. Science Press, Beijing, pp 58–61
10. Cao Y (2012) Research of NonGaussian/nonlinear filtering algorithms and its applications in GPS kinematic positioning. PLA Information Engineering University

Vision-Aided Inertial Navigation System with Point and Vertical Line Observations for Land Vehicle Applications

Zhenbo Liu, Qifan Zhou, Yongyuan Qin and Naser El-Sheimy

Abstract Various aiding sensors can be integrated with the inertial navigation system (INS) to reduce its error growth when the vehicle is operating in GNSS denied environments. This paper developed a method to use the vanishing point from vertical line observations of building blocks in order to further improve point-based visual-inertial navigation system (VINS) for land vehicle applications. First, we presented the formulations of tightly coupled point-based VINS based on the Multi-State Constraint Kalman Filter (MSCKF) in the local-level frame. Second, we developed the relationship between the INS roll angle and vanishing point coordinates from vertical line observations. The roll angle measurement model is formulated. Finally, loosely coupled vertical line aiding module is added to the existing VINS, and the integration scheme is presented. Real world experiments demonstrated the validity of the mixed VINS method and the improved accuracy of the attitude and position estimation when compared with the solution without vertical line vanishing point aiding.

Keywords Visual-inertial navigation system · Vanishing point · Roll angle · MSCKF · Localization

1 Introduction

In the last decade there has been an increasing use of Autonomous Ground Vehicles (AGVs) in many navigation applications. One of the representative civil application is in intelligent transport systems (ITSs); more specifically in autonomous cars or driverless cars [1]. In the vehicle's Guidance, Navigation, and Control (GNC) system, the navigation state estimation is an essential component, not only

Z. Liu (✉) · Q. Zhou · N. El-Sheimy
Department of Geomatics Engineering, University of Calgary, Calgary, Canada
e-mail: zhenbo.liu2@ucalgary.ca

Z. Liu · Y. Qin
School of Automation, Northwestern Polytechnical University, Xi'an, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_36

for the vehicle's autonomous navigation, but also for cooperative safety communications among vehicles [2–4]. Among different navigation systems, the inertial navigation system (INS) is the only one that has the capability to produce a complete and continuous set of navigation state data with high precision during a short time span. However, low-cost INS has to be integrated with other aiding sensors to reduce its error growth. INS and Global Navigation Satellite System (GNSS) integration is commonly used for vehicle outdoor navigation. However, GNSS signal can suffer from obstruction and multipath errors in city centers and mountainous regions, and is prone to the possibility of being jammed or spoofed [5].

To mitigate the error growth of INS in GNSS denied environment, aiding sensors such as wheel odometers and motion constraints (e.g., Non-holonomic Constraints (NHC) as a "virtual" sensor) were utilized for land vehicles [6, 7]. Recently, researchers proposed to use vision sensor aiding in the navigation system, in loosely coupled form, with the visual odometry (VO) module [8, 9], or in tightly coupled form [10–13]. Most research used point features in the tightly coupled VINS, including simultaneous localization and mapping (SLAM) based methods [10, 14] and structure-less methods, which focus on the ego-motion only [12, 15]. The representative algorithm of filter based methods is called multi-state constraint Kalman filter (MSCKF) [12], which augments camera poses (a sliding window of poses) into the state vector to make most use of measurements of feature points in the geometry constraint. Apart from the point-based VINS, there are approaches using the vanishing point (VP) module [16, 17], another attitude aiding source for INS working in the environment with structured buildings or in corridors. These methods rely on the observation of parallel lines in structured buildings.

In this paper, we present MSCKF in the local-level frame (LLF), which is common to the navigation community. To improve the performance of point-based MSCKF, we propose to add a VP module, which observes vertical lines of building blocks, for land vehicle navigation. Based on that, a roll angle measurement model is derived. Therefore, a mixed VINS using points and line observations is developed, and its performance is verified by real data experiments.

2 MSCKF Algorithm in Local-Level Frame

MSCKF is considered as a structure-less visual-inertial navigation system, which means the vehicle localization is of the main interest, rather than recovering the structure of the environment. They take into account all camera poses at which observations of the same 3D point occurred. MSCKF augments camera poses (a sliding window of poses) into the state vector to make most use of measurements of feature points in the geometry constraint. Every feature has a corresponding stack of misclosure vectors (residuals). Feature position error states are cancelled out by projecting the residual stack on the left null space of the design matrix corresponding to the feature error state [12]. Here we present the MSCKF algorithm in LLF with the notations common to navigation engineers and researchers, in order to

bridge between the navigation community and robotics community. It should be mentioned that some components in the filtering formulations should be modified with respect to original MSCKF.

2.1 MSCKF States and Covariance

For the near-Earth navigation, the navigation frame is usually the local-level frame (n-frame; east-north-up is used in this paper). The navigation state is presented by the position (latitude L, longitude λ and height h), velocity (east velocity, north velocity, and up velocity), and attitude (pitch, roll, and azimuth). Differential equations of position $\boldsymbol{P}$, velocity $\boldsymbol{V}^n$ and attitude matrix $\boldsymbol{C}_b^n$ are given by

$$\dot{\boldsymbol{P}} = \boldsymbol{D}^{-1}\boldsymbol{V}^n = \begin{bmatrix} 0 & \frac{1}{R_M+h} & 0 \\ \frac{1}{(R_N+h)\cos L} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{V}^n \tag{2.1}$$

$$\dot{\boldsymbol{V}}^n = \boldsymbol{f}^n - \left(2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n\right) \times \boldsymbol{V}^n + \boldsymbol{g}^n \tag{2.2}$$

$$\dot{\boldsymbol{C}}_b^n = \boldsymbol{C}_b^n \left[\boldsymbol{\omega}_{ib}^b \times\right] - \left[\left(\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n\right) \times\right] \boldsymbol{C}_b^n \tag{2.3}$$

where R_M and R_N are meridian radius of curvature and prime vertical radius respectively, $\boldsymbol{f}^n$ is the specific force in n-frame, $\boldsymbol{\omega}_{ie}^n$ is the earth rotation angular rate vector, $\boldsymbol{\omega}_{en}^n$ is the angular rate of the navigation frame with respect to (w.r.t.) ECEF frame, and $\boldsymbol{\omega}_{ib}^b$ is the angular rate of the body frame (b) w.r.t. the inertial frame (i). The notation $(\boldsymbol{a})\times$ is the asymmetric matrix of a vector $\boldsymbol{a}$. The navigation states are propagated using Eqs. (2.1–2.3), which is called INS mechanization where the IMU measurement is the input signal.

The navigation states can be considered as the full states, and the associated errors are called error states. For convenience, let $\boldsymbol{X}$ denote the error states, and the full states will be presented by their quantities notions. The state vector of the MSCKF is:

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_I^T & \boldsymbol{X}_{C_1}^T & \boldsymbol{X}_{C_2}^T & \cdots & \boldsymbol{X}_{C_m}^T \end{bmatrix}^T \tag{2.4}$$

where $\boldsymbol{X}_I$ is the current INS error states, and $\boldsymbol{X}_{C_i}, i = 1 \ldots m$ are camera poses errors at the times the last m images were recorded. INS error states are defined as

$$\boldsymbol{X}_I = \begin{bmatrix} \delta\boldsymbol{V}^{nT} & \boldsymbol{\phi}^{nT} & \delta\boldsymbol{P}^T & \boldsymbol{b}_g^T & \boldsymbol{b}_a^T \end{bmatrix}^T \tag{2.5}$$

where $\delta\boldsymbol{V}^n$, $\boldsymbol{\phi}^n$, and $\delta\boldsymbol{P}$ are the velocity error, misalignment error, and position error (latitude, longitude and height error). $\boldsymbol{b}_g$ and $\boldsymbol{b}_a$ are biases of gyroscopes and

accelerometers. In this case, the dimension of the state vector is 15 + 6 m. Misalignment error $\boldsymbol{\phi}^n$ (considered as the small angle) is defined by

$$\hat{\boldsymbol{C}}_b^n = [\boldsymbol{I} - \boldsymbol{\phi}^n\times]\boldsymbol{C}_b^n \tag{2.6}$$

For EKF, the INS error state propagation is based on the INS error dynamic model. The IMU sensor biases are modelled as random constants, random walks or Gauss-Markov processes. The *i*th camera pose errors are defined as:

$$\boldsymbol{X}_{C_i} = \begin{bmatrix} \boldsymbol{\phi}_{C_i}^{nT} & \delta\boldsymbol{P}_{C_i}^T \end{bmatrix}^T \tag{2.7}$$

where $\boldsymbol{\phi}_{C_i}^n$ and $\delta\boldsymbol{P}_{C_i}$ are the camera's attitude error and position error at the *i*th recorded instance. The augmented *i*th camera pose errors are modeled as random constant, since the previous camera poses will not change over time.

The camera pose is augmented into the state when recording a new image. It is calculated from current IMU pose and the known relative pose between IMU and the camera using Eqs. 2.8–2.9.

$$\hat{\boldsymbol{C}}_C^n = \hat{\boldsymbol{C}}_b^n\boldsymbol{C}_C^b \tag{2.8}$$

$$\hat{\boldsymbol{P}}_C = \hat{\boldsymbol{P}} + \boldsymbol{D}^{-1}\hat{\boldsymbol{C}}_b^n\boldsymbol{P}_{bC}^b \tag{2.9}$$

where $\boldsymbol{C}_C^b$ and $\boldsymbol{P}_{bC}^b$ are the rotation matrix and lever-arm between the camera and the IMU; $\hat{\boldsymbol{P}}_C$ and $\hat{\boldsymbol{C}}_c^n$ are the position and attitude matrix of the camera. Then the camera pose estimate is appended to the state vector, and the covariance matrix is augmented accordingly:

$$\boldsymbol{X}_{(15+6(m+1))} \leftarrow \left[\boldsymbol{X}_{(15+6m)\times 1}; \boldsymbol{X}_{C(6\times 1)}\right]^T \tag{2.10}$$

$$\boldsymbol{P}_{k|k} \leftarrow \begin{bmatrix} \boldsymbol{I}_{15+6m} \\ \boldsymbol{J}_{6\times(15+6m)} \end{bmatrix} \boldsymbol{P}_{k|k} \begin{bmatrix} \boldsymbol{I}_{15+6m} \\ \boldsymbol{J}_{6\times(15+6m)} \end{bmatrix}^T \tag{2.11}$$

where

$$\boldsymbol{J}_{6\times(15+6m)} = \begin{bmatrix} \boldsymbol{O}_{3\times 3} & \boldsymbol{I}_{3\times 3} & \boldsymbol{O}_{3\times 3} & \boldsymbol{O}_{3\times 6} & \boldsymbol{O}_{3\times 6m} \\ \boldsymbol{O}_{3\times 3} & \boldsymbol{D}^{-1}\left(\widehat{\boldsymbol{C}}_b^n\boldsymbol{P}_{bC}^b\right)\times & \boldsymbol{I}_{3\times 3} & \boldsymbol{O}_{3\times 6} & \boldsymbol{O}_{3\times 6m} \end{bmatrix}$$

2.2 MSCKF Procedures

Assuming that the *i*th feature, which has been observed in a set of M_i images, is no longer detected. An EKF measurement update will be triggered after forming a

residual corresponding to this ith feature. Suppose we have the measurement $\tilde{\boldsymbol{z}}_i^j$ for $j = 1 \ldots M_i$. The predicted measurement of the ith feature in jth image is

$$\hat{\boldsymbol{z}}_i^j = h\left(\widehat{\boldsymbol{P}}_{C_j}, \widehat{\boldsymbol{C}}_{C_j}^n, \widehat{\boldsymbol{P}}_{f_i}\right) \tag{2.12}$$

where $h(\cdot)$ is the nonlinear camera projection equation, and $\widehat{\boldsymbol{P}}_{f_i}$ is the computed feature position. So the residuals of ith feature are

$$\delta \boldsymbol{z}_i^j = \hat{\boldsymbol{z}}_i^j - \tilde{\boldsymbol{z}}_i^j \approx \boldsymbol{H}_{X_i}^j \boldsymbol{X} + \boldsymbol{H}_{f_i}^j \delta \boldsymbol{P}_{f_i} + n_i^j \tag{2.13}$$

where the Jacobian matrix is

$$\boldsymbol{H}_{X_i}^j = \begin{bmatrix} \boldsymbol{O}_{2\times 15} & \boldsymbol{O}_{2\times 6} & \cdots & \boldsymbol{J}_i^j \left(\widehat{\boldsymbol{P}}_{f_i}^{C_j} \times\right) \widehat{\boldsymbol{C}}_n^{c_j} & -\boldsymbol{J}_i^j \widehat{\boldsymbol{C}}_n^{c_j} \boldsymbol{D} & \cdots & \boldsymbol{O}_{2\times 6} \end{bmatrix}_{2\times(15+6\mathrm{m})}$$

$$\boldsymbol{H}_{f_i}^j = \boldsymbol{J}_i^j \widehat{\boldsymbol{C}}_n^{c_j} \boldsymbol{D}$$

where

$$\boldsymbol{J}_i^j = \frac{1}{\widehat{Z}_i^{C_j}} \begin{bmatrix} 1 & 0 & -\frac{\widehat{X}_i^{C_j}}{\widehat{Z}_i^{C_j}} \\ 0 & 1 & -\frac{\widehat{Y}_i^{C_j}}{\widehat{Z}_i^{C_j}} \end{bmatrix}, \quad \begin{bmatrix} \widehat{X}_i^{C_j} \\ \widehat{Y}_i^{C_j} \\ \widehat{Z}_i^{C_j} \end{bmatrix} = \widehat{\boldsymbol{P}}_{f_i}^{C_j} = \widehat{\boldsymbol{C}}_n^{C_j} \boldsymbol{D} \left(\widehat{\boldsymbol{P}}_{f_i} - \widehat{\boldsymbol{P}}_{C_j}\right)$$

Stack the $\delta \boldsymbol{z}_i^j$ of all M_i measurements of the ith feature to form the block residual vector:

$$\delta \boldsymbol{z}_i = \boldsymbol{H}_{X_i} \boldsymbol{X} + \boldsymbol{H}_{f_i} \delta \boldsymbol{P}_{f_i} + \boldsymbol{n}_i \tag{2.14}$$

To remove the effect of feature position error, define a residual vector $\delta \boldsymbol{z}_i^O = \boldsymbol{A}^{\mathrm{T}} \delta \boldsymbol{z}_i$, where $\boldsymbol{A}$ is a matrix whose columns form the basis of the left null space of $\boldsymbol{H}_{f_i}$. The new residual will be

$$\delta \boldsymbol{z}_i^O = \boldsymbol{A}^{\mathrm{T}} \delta \boldsymbol{z}_i = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{H}_{X_i} \boldsymbol{X} + \boldsymbol{A}^{\mathrm{T}} \boldsymbol{n}_i = \boldsymbol{H}_i^O \boldsymbol{X} + \boldsymbol{n}_i^O \tag{2.15}$$

After the measurement update in the Kalman filter, do the feedback corrections to the navigation states and camera poses.

The MSCKF can be implemented and operated in different ways [18]. For example, minimum feature track lengths can be set to trigger an update. Large minimum feature track length means that the landmarks will be observed with a larger base line, while some near landmarks will be neglected. Potentially, large minimum feature track may have better accuracy than the small minimum feature

track, but the former will also have the problem of losing useful information, which is often better than none. Usually, a large minimum feature track length will require more features to be detected to have superior accuracy than the small minimum feature track length.

3 Roll Angle Aiding with Vertical Line Observations

Most approaches of VINS have focused on the point features observation. Line features are omnipresent and can be reliably extracted and tracked in structured urban environments. Here we propose to use vanishing point coordinates from buildings' vertical line observations as an attitude aiding to improve point-based VINS on-board the land vehicles.

3.1 Roll Measurement from Vertical Line Observations

The advantage of using buildings' vertical line observations is that the measurement is absolute roll, since the directions of these vertical lines are known. A vanishing point is the intersection of image projections of a set of parallel 3D lines in the scene. Each set of parallel lines is associated to a VP in an image [19].

The vertical lines in the world frame (local-level frame here) intersect at the infinity point which can be represented as $\boldsymbol{U} = [0 \quad 0 \quad 1 \quad 0]^{\boldsymbol{T}}$. The corresponding vanishing point $\boldsymbol{u}$ in the image plane is the projection of the infinity point $\boldsymbol{U}$. The projection equation is written as:

$$\boldsymbol{u} = \boldsymbol{K}\left[\boldsymbol{C}_n^C \quad \boldsymbol{T}\right]\boldsymbol{U} \tag{3.1}$$

where $\boldsymbol{K}$ is the camera calibration matrix, $\boldsymbol{C}_n^C$ is the camera rotation matrix, and $\boldsymbol{T}$ is the translation vector of the camera. Equation (3.1) can be rewritten as

$$\boldsymbol{u} = \boldsymbol{K}\boldsymbol{C}_b^C\boldsymbol{C}_n^b[0 \quad 0 \quad 1]^T \tag{3.2}$$

Then INS roll (γ) and pitch (θ) can be represented as

$$\begin{bmatrix} -\sin\gamma\cos\theta \\ \sin\theta \\ \cos\gamma\cos\theta \end{bmatrix} = \boldsymbol{C}_C^b\boldsymbol{K}^{-1}\boldsymbol{u} \tag{3.3}$$

The roll angle can be calculated as

$$\gamma = \tan^{-1}\left(-\frac{\boldsymbol{e}_1^T \boldsymbol{C}_C^b \boldsymbol{K}^{-1}\boldsymbol{u}}{\boldsymbol{e}_3^T \boldsymbol{C}_C^b \boldsymbol{K}^{-1}\boldsymbol{u}}\right) \tag{3.4}$$

where $\boldsymbol{e}_1 = [1, \quad 0, \quad 0,]^T$, and $\boldsymbol{e}_3 = [0, \quad 0, \quad 1]^T$.

Suppose the camera has been calibrated, so $\boldsymbol{C}_C^b \boldsymbol{K}^{-1}$ is known. Vanishing point $\boldsymbol{u}$ is determined by the coordinates of the intersection of multiple detected vertical lines in the image. In this paper, we use the Line Segment Detection (LSD) [20] to get the lines segments, which are then clustered into multiple VP classes based on the J-Linkage model [21]. The vertical VP $\boldsymbol{u}$ is one of the dominant VP classes, and can be easily selected and estimated [22]. Note that some erroneous line segments exist in the vertical VP class. They can be further cancelled by making additional assumptions, for example on the orientation or length of the line segments.

3.2 Roll Angle Measurement Model

In the INS mechanization, the roll angle γ can be calculated by

$$\hat{\gamma} = \tan^{-1}\frac{-\hat{c}_{31}}{\hat{c}_{33}} \tag{3.5}$$

where $\hat{c}_{31}$ and $\hat{c}_{33}$ are elements of attitude matrix $\widehat{\boldsymbol{C}}_b^n$, which can be written as

$$\widehat{\boldsymbol{C}}_b^n = [\boldsymbol{I} - \boldsymbol{\phi}^n\times]\boldsymbol{C}_b^n = \begin{bmatrix} 1 & \phi_U & -\phi_N \\ -\phi_U & 1 & \phi_E \\ \phi_N & -\phi_E & 1 \end{bmatrix}\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \tag{3.6}$$

So the roll can be expressed as

$$\hat{\gamma} = \tan^{-1}\frac{-\hat{c}_{31}}{\hat{c}_{33}} = \tan^{-1}\frac{-(c_{11}\phi_N - c_{21}\phi_E + c_{31})}{c_{13}\phi_N - c_{23}\phi_E + c_{33}} \tag{3.7}$$

Then the INS roll error $\delta\gamma$ can be expressed as

$$\delta\gamma = \frac{\partial\hat{\gamma}}{\partial\phi_E}\phi_E + \frac{\partial\hat{\gamma}}{\partial\phi_N}\phi_N + \frac{\partial\hat{\gamma}}{\partial\phi_U}\phi_U \tag{3.8}$$

where

$$\frac{\partial\hat{\gamma}}{\partial\phi_E} \approx \frac{\hat{c}_{21}\hat{c}_{33} - \hat{c}_{31}\hat{c}_{23}}{\hat{c}_{31}^2 + \hat{c}_{33}^2}, \quad \frac{\partial\hat{\gamma}}{\partial\phi_N} \approx \frac{\hat{c}_{31}\hat{c}_{13} - \hat{c}_{11}\hat{c}_{33}}{\hat{c}_{31}^2 + \hat{c}_{33}^2}, \quad \frac{\partial\hat{\gamma}}{\partial\phi_U} = 0$$

Subtracting the VP derived roll angle from the INS roll angle, we have the residual

$$Z_\gamma = \hat{\gamma} - \tilde{\gamma} = \delta\gamma - v_\gamma \tag{3.9}$$

Therefore, the design matrix of the roll measurement is

$$\boldsymbol{H}_\gamma = \begin{bmatrix} \boldsymbol{O}_{1\times3} & \frac{\partial\hat{\gamma}}{\partial\phi_E} & \frac{\partial\hat{\gamma}}{\partial\phi_N} & \frac{\partial\hat{\gamma}}{\partial\phi_U} & \boldsymbol{O}_{1\times9} & \boldsymbol{O}_{1\times6m} \end{bmatrix} \tag{3.10}$$

As we can see from the Jacobian matrices, the roll measurement will contribute to the estimation of both horizontal attitude errors but not the azimuth error.

3.3 Navigation Procedure

The navigation flow chat is illustrated in Fig. 1. When a new image is observed, some key functions in MSCKF and VP aiding will work: state augmentation and feature detection and matching, and vertical line detection, respectively. If the measurement update is triggered in either vision modules, the states will be corrected.

4 Results and Analysis

The point-based MSCKF and vanishing point aiding from vertical lines algorithm described in the previous sections was tested using '2011_09_26_drive_0036' dataset from the KITTI benchmark [23]. The synchronized IMU data (10 Hz) and rectified grayscale images (10 Hz, global shutter) were used to verify the algorithm. The biases of gyroscopes and accelerometers are 0.01°/s and 1 mg (see more details in [24]). The reference was the IMU/GPS results (OXTS RT3003) with the L1/L2 RTK positioning accuracy of 0.02 m and pitch/roll accuracy of 0.03°. We used 250 images with traveled distance about 180 m, as shown on Fig. 2.

To illustrate the performance of proposed algorithms, the GPS data was totally blocked of whole trajectory except for the initial value of navigation states. We compared the performance of the following navigation schemes.

- Free INS: Only INS mechanization is performed to calculate the navigation states.
- MSCKF-VINS with minimum 5 tracking poses: minimum 5 poses in the tracking are required to trigger a measurement update.
- MSCKF-VINS with minimum 20 tracking poses: minimum 20 poses in the tracking are required to trigger a measurement update.

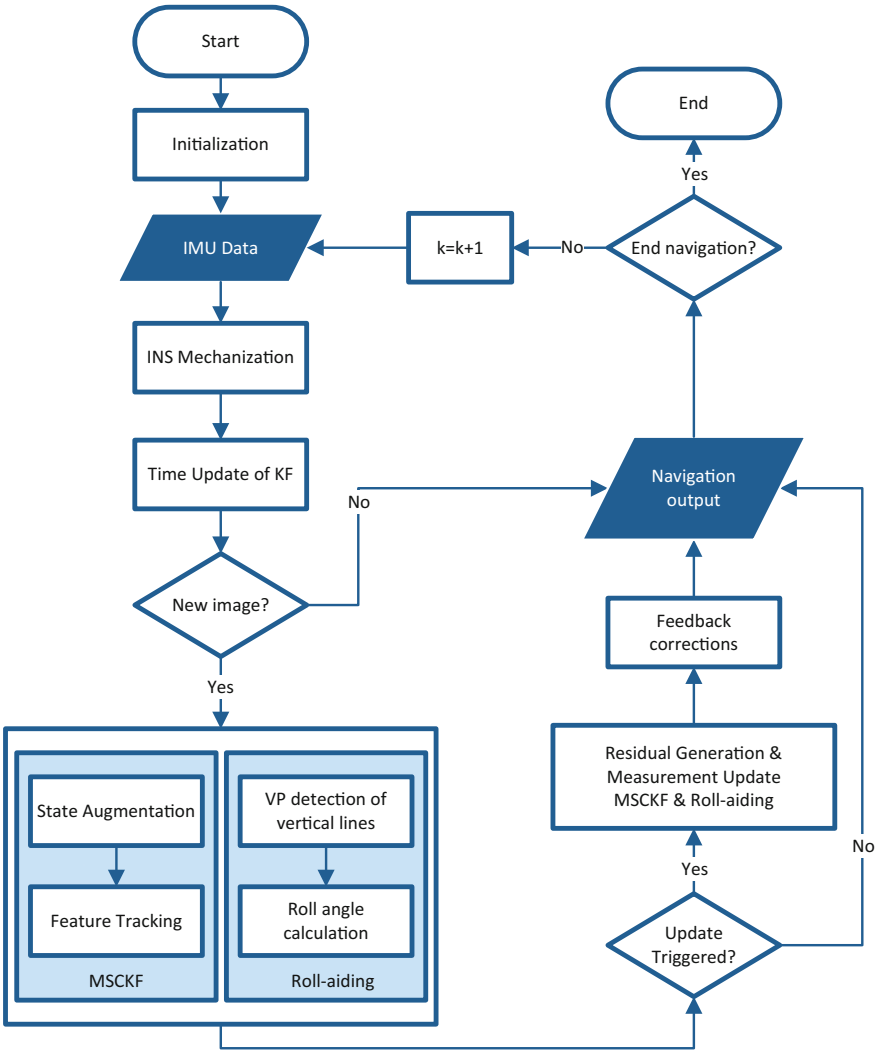


Fig. 1 Navigation flowchart

- MSCKF-VINS with minimum 20 tracking poses and with vertical line aiding: Vanishing point aiding module is added into the system where minimum 20 poses in the tracking are required.

The trajectory results of the 4 navigation schemes are shown in Fig. 2. The horizontal attitude errors and position errors are presented in Figs. 3 and 4. We can see that there is just minor difference between MSCKF-VINS with minimum 5 and 20 tracking poses. A slight better pitch estimation of MSCKF-VINS with 20 tracking poses can be seen in Fig. 3 when there are abundant features (before 14 s).

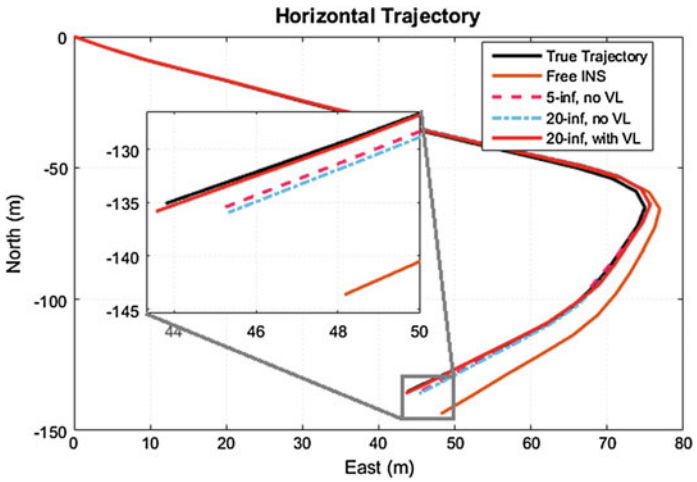


Fig. 2 Reference and calculated trajectory

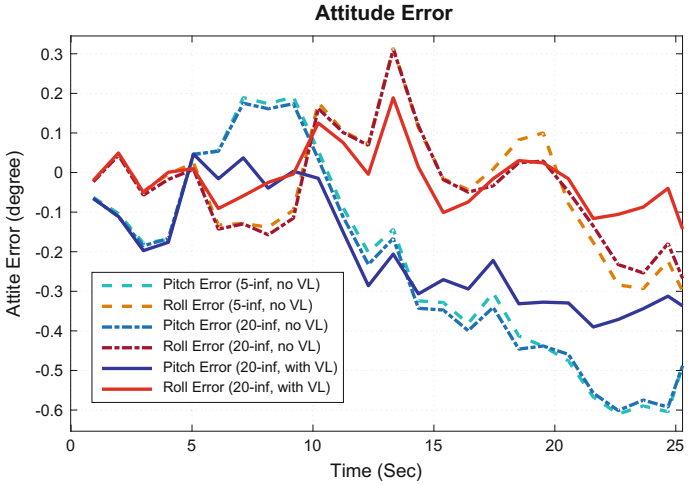


Fig. 3 Attitude errors (pitch and roll)

One reason can be that the longer tracking will observe further distance landmarks, which is beneficial for the pitch angle estimation.

An example of vertical line clustering is shown in Fig. 5, where the green line segments correspond to the vertical VP. When adding the vertical line vanishing point aiding into the system, both pitch and roll estimations are improved to a great extent (up to 0.2°) compared with MSCKF-only methods. The horizontal position error is also less than MSCKF-only methods in general, within 1.2 m. The position estimation improvement results from better attitude estimation, because the attitude

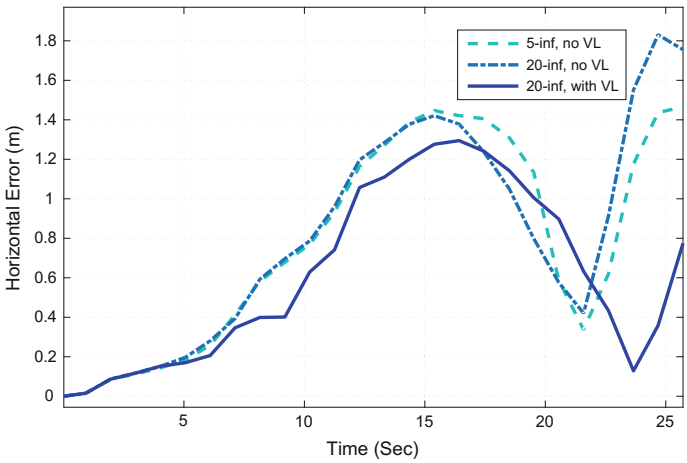


Fig. 4 Horizontal position errors

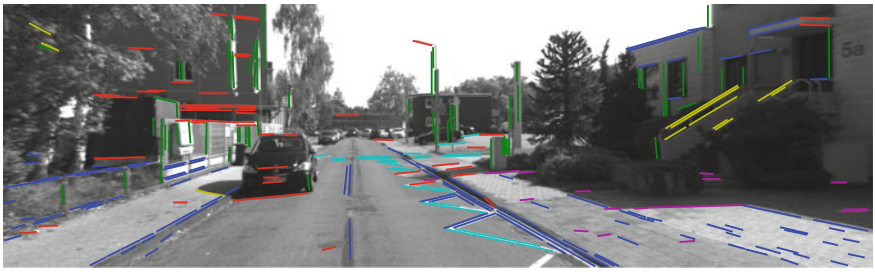


Fig. 5 Line segments groups with different colors (Color figure online)

error will produce horizontal acceleration error by projecting the gravity to the horizontal plane.

5 Conclusion

In this paper, we use MSCKF in the local-level frame and present it in a way that is common to navigation community. To improve the performance of point-based MSCKF, we proposed to add a VP module which observes the vertical lines of building blocks for land vehicle navigation. Based on that, the roll angle measurement model was derived. Therefore, a mixed VINS using points and line observations was developed, and its performance was verified by real data experiments. Horizontal attitude and position estimations are improved by using vertical line vanishing point aiding module.

References

1. Vinet L, Zhedanov A (2012) Handbook of intelligent vehicles. Springer London, London
2. Zolghadri A, Henry D, Cieslak J (2014) Fault diagnosis and fault-tolerant control and guidance for aerospace vehicles. Springer, London
3. Liu Z, El-Sheimy N, Qin Y, Yu C, Zhang J (2016) Partial state feedback correction for smoothing navigational parameters. In: Sun J, Liu J, Fan S, Wang F (eds) China satellite navigation conference (CSNC) 2016 proceedings, vol II. Springer Singapore, Singapore, pp 461–472
4. Ahmed-Zaid F, Bai F, Bai S, Basnayake C, Bellur B, Brovold S, Brown G, Caminiti L, Cunningham D, Elzein H, Hong K, Ivan J, Jiang D, Kenney J, Krishnan H, Lovell J, Maile M, Masselink D, McGlohon E, Mudalige P, Popovic Z, Rai V, Stinnett J, Tellis L, Tirey K, and VanSickle S (2011) Vehicle safety communications–applications (VSC-A)
5. Chowdhary G, Johnson EN, Magree D, Wu A, Shein A (2013) GPS-denied indoor and outdoor monocular vision aided navigation and control of unmanned aircraft. J F Robot 30 (3):415–438
6. Dissanayake G, Sukkarieh S, Nebot E, Durrant-Whyte H (2001) The aiding of a low-cost strapdown inertial measurement unit using vehicle model constraints for land vehicle applications. IEEE Trans Robot Autom 17(5):731–747
7. Niu X, Nassar S, El-Sheimy N (2007) An accurate land-vehicle MEMS IMU/GPS navigation system using 3D auxiliary velocity updates. J Inst Navig 54(3):177–188
8. Nistér D, Naroditsky O, Bergen J (2006) Visual odometry for ground vehicle applications. J F Robot 23(1):3–20
9. Scaramuzza D, Fraundorfer F (2011) Visual odometry [tutorial]. IEEE Robot Autom Mag 18 (4):80–92
10. Veth M, Raquet J (2006) Two-dimensional stochastic projections for tight integration of optical and inertial sensors for navigation, Natl Tech Meet Proc Inst Navig, pp 587–596
11. Veth MJ (2008) Fusion of imaging and inertial sensors for navigation. Air Force Institute of Technology, USA
12. Mourikis AI Roumeliotis SI, (2007) A multi-state constraint Kalman filter for vision-aided inertial navigation, In: Proceedings-IEEE international conference on robotics and automation, pp 3565–3572
13. Leutenegger S, Lynen S, Bosse M, Siegwart R, Furgale P (2015) Keyframe-based visual-inertial odometry using nonlinear optimization. Int J Rob Res 34(3):314–334
14. Veth M, Raquet J, Pachter M (2006) Stochastic constraints for efficient image correspondence search. IEEE Trans Aerosp Electron Syst 42(3):973–982
15. Forster C, Carlone L, Dellaert F, Scaramuzza D (2015) On-manifold pre-integration theory for fast and accurate visual-inertial navigation, IEEE Trans Robot, pp 1–18
16. Kim S-B, Bazin J-C, Lee H-K, Choi K-H, Park S-Y (2011) Ground vehicle navigation in harsh urban conditions by integrating inertial navigation system, global positioning system, odometer and vision data. IET Radar Sonar Navig 5(8):814
17. Camposeco F, Pollefeys M (2015) Using vanishing points to improve visual-inertial odometry. In: 2015 IEEE international conference on robotics and automation (ICRA), pp 5219–5225
18. Clement LE, Peretroukhin V, Lambert J Kelly J (2015) The Battle for filter supremacy: a comparative study of the multi-state constraint Kalman filter and the sliding window filter. In: Proceedings of 2015 12th conference on computer and robot vision, CRV 2015, pp 23–30
19. Hartley R, Zisserman A (2003) Multiple view geometry in computer vision. Cambridge University Press, Cambridge
20. Grompone Von Gioi R, Jakubowicz J, Morel JM, Randall G (2010) LSD: a fast line segment detector with a false detection control. IEEE Trans Pattern Anal Mach Intell 32(4):722–732

21. Toldo R, Fusiello A Informatica D (2008) Robust multiple structures estimation with J-Linkage. In: Computer vision-ECCV: 10th European conference on computer vision, marseille, France, Oct 12–18, Proceedings, Part I, pp 537–547
22. Feng C, Deng F Kamat VR (2010) Semi-automatic 3D reconstruction of piecewise planar. In: 10th international conference construction applications of virtual reality, pp 1–9
23. Geiger A, Lenz P, Stiller C, Urtasun R (2013) Vision meets robotics: the KITTI dataset. Int J Rob Res 32(11):1231–1237
24. RT3000 inertial and GPS navigation system. Available: http://www.oxts.com/Downloads/Support/Brochures/rt3kpres.pdf. Accessed: 17-Jan-2017

An Attitude Determination Algorithm by Integration of Inertial Sensor Added with Vision and Multi-antenna GNSS Data

Fengyang Li, Xuedong Jia, Yanglin Zhou, Ming Dong and Changyuan Chen

Abstract Aim to the question that GNSS cannot directly observe positioning and attitude information using the GNSS/INS integrated system, especially in the bad condition of GNSS signal, positioning and orienting based on satellites are both difficult to use. This chapter establishes the vision/multi-antenna GNSS attitude determination/inertial integrated navigation system, taking inertial error propagation model as the navigation model, the positioning and attitude information of multi-antenna GNSS and visual navigation information as a combination of measurements, achieving high precision integrated navigation using multi-scale extended Kalman filter. The vehicle-mounted integrated navigation test verified that the heading of this system has an accuracy of 0.243°, providing an attitude determination algorithm by integration of inertial sensor added with camera, and multi-antenna GNSS in the denied environment of satellite.

Keywords Integrated navigation · Multi-antenna GNSS attitude determination · Vision navigation · Artificial mark point · Multi-scale Kalman filter

1 Introduction

With the development of navigation demands in different application fields, the integrated navigation technology based on inertial navigation system (INS) has gradually become the most important researching focus in the field of aerospace and missile guidance. For example, in the GNSS (Global Navigation Satellite System)/

F. Li (✉) · X. Jia · Y. Zhou · C. Chen
Institution of Navigation and Aerospace Engineering,
Information Engineering University, Zhengzhou 450001, China
e-mail: FreyrLi@outlook.com

F. Li · M. Dong
State Key Laboratory of Geo-Information Engineering, Xian 710054, China

M. Dong
Satellite Navigation Engineering Center, Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_37

INS integrated navigation system, the inertial instrument can obtain error correction of the position and attitude, and the GNSS can assist the inertial navigation system to achieve high precision navigation in the good observation condition, which can give full play to the advantages of each single navigation system.

Zhang F et al. studied the direct method and least square method on attitude determination using multi antenna GNSS, and the result showed the external accord accuracy of least square method is better when estimating heading [1]; Zhai H implemented the estimation of acceleration using the velocity observation of MEMS/GPS, achieving the attitude determination based on single antenna GNSS, and the simulation results show that the attitude evaluation is better than 0.5° [2]; Dong M carried out the observation of heading through the double difference method using a single antenna GNSS receiver [3], but do not get complete three position observed quantities. However, GNSS cannot observe positioning and attitude information directly, in the denied environment of GNSS, positioning and orienting based on GNSS are both difficult to use. The visual sensor is not affected by the impact of electromagnetic environmental, in the known topological space environment, perspective projection model can be established through the analysis of nonlinear relationship between the image feature points and the real target, and the target motion parameters can be estimated.

This paper established the vision/multi-antenna GNSS attitude determination/inertial integrated navigation system (VMGINS), taking inertial error propagation model as the navigation model, the multi-antenna GNSS pose information and visual navigation positioning information as a combined of measurements, achieving high precision integrated navigation using multi-scale EKF, providing a new method for integrated navigation attitude measurement in a complicate GNSS electromagnetic environment.

2 Strapdown Inertial Navigation Error Model

In the paper, the strapdown inertial navigation mechanization is implemented in the navigation frame (n-frame) for updating of position, velocity and attitude [4]:

$$\begin{cases} \dot{\boldsymbol{r}}^n = \boldsymbol{v}^n - \boldsymbol{\omega}_{en}^n \times \boldsymbol{r}^n \\ \dot{\boldsymbol{v}}^n = \boldsymbol{f}^n - (2\boldsymbol{\omega}_{i.e.}^n + \boldsymbol{\omega}_{en}^n) \times \boldsymbol{v}^n + \boldsymbol{g}^n \\ \dot{\boldsymbol{C}}_b^n = \boldsymbol{C}_b^n(\boldsymbol{\omega}_{ib}^b - \boldsymbol{\omega}_{in}^b) \end{cases}$$

where superscript n refers to the n-frame, $\boldsymbol{r}^n$ refers to the position vector, $\boldsymbol{v}^n$ is the velocity vector, $\boldsymbol{C}_b^n$ is the direction cosine matrix from body frame (body-frame) to navigation frame, $\dot{[]}$ refers to the format of differential coefficient, $\boldsymbol{f}^n$ is specific force vector measured by accelerometers, $\boldsymbol{g}^n$ is the gravity, $\boldsymbol{\omega}_{en}^n$ is the angular rate vector of n-frame with respect to earth-centered earth-fixed frame (e-frame) resolved in n-frame, $\boldsymbol{\omega}_{i.e.}^n$ the angular rate vectors of e-frame with respect to inertial frame

(i-frame), which equals to rotational angular velocity of the earth, $\boldsymbol{\omega}_{ib}^{b}$ is the increment vector of angular velocity measured by gyroscopes, $\boldsymbol{\omega}_{in}^{b}$ refers to the format of $\boldsymbol{C}_{n}^{b}(\boldsymbol{\omega}_{i.e.}^{n}+\boldsymbol{\omega}_{en}^{n})$.

3 CCD Camera Model

Define coordinate systems: ①. Object space frame, defined as the space rectangular coordinate system, has its origin at a point of 3D space, which is also called as world frame (w-frame). ②. Plane frame (p-frame) has its origin at the center of the image, and its x-axis points towards right, y-axis points towards upside. ③. Image space frame, also called as camera frame (c-frame), has its origin at the optical center of lens, and its x-axis is parallel to the u-axis of the plane frame, the z-axis is perpendicular to the image plane. The p coordinate of p-frame expressed as (x, y) is corresponding to that in c-frame expressed as (x, y, −f). ④ Image frame (u-frame) has its origin at the upper left corner of the image, and its u-axis points to right, v-axis points to underside. Each pixel refers to the (u, v) format, which refer to the number of column and row respectively. The relationships between the above three frames are shown in Fig. 1.

Where $\overrightarrow{OP}$ refers to the vector of feature point P in n-frame, denoted by $\boldsymbol{p}^{n}$, the corresponding vector of which in v-frame refers to $\overrightarrow{cP}$, denoted by $\boldsymbol{p}^{c}$. $\overrightarrow{Oc}$ refers to the vector of optical center of lens c in n-frame, denoted by $\boldsymbol{t}^{n}$. The transformation relation of P between of above-mentioned n-frame and v-frame is:

$$\boldsymbol{p}^{c}=\boldsymbol{R}_{n}^{c}\boldsymbol{p}^{n}+\boldsymbol{t}^{c} \tag{1}$$

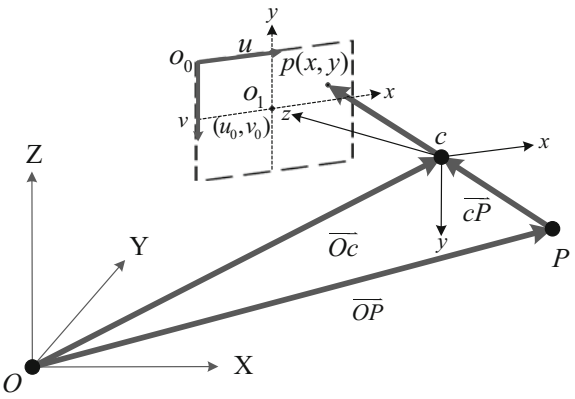


Fig. 1 The relationship of world, camera, plane and image frame

where R_n^c represent the direction cosine matrix from n-frame to v-frame in the Euler angle ordered by $\varphi - \omega - \kappa$, $\boldsymbol{t}^c$ is the translation vector between two frames, equaled to $-\boldsymbol{R}_n^c \boldsymbol{t}^n$.

Define the real plane coordinate of P, which is projected on the image plane where $\hat{z}$ equals to 1, refers to $\boldsymbol{V} = (\hat{x}, \hat{y}, 1)^T$, which is known as normalized plane. The format of homogeneous transformation can be expressed as:

$$\boldsymbol{V} = \frac{1}{\boldsymbol{c}_{3n}^c \boldsymbol{r}_z^n + \boldsymbol{t}_z^c} (\boldsymbol{C}_n^c \boldsymbol{r}^n + \boldsymbol{t}^c) = \frac{\boldsymbol{p}^c}{\boldsymbol{p}_z^c} = m\boldsymbol{p}^c \tag{2}$$

where m is constant, $\boldsymbol{c}_{3n}^c$ and $\boldsymbol{t}_z^c$ are the third row of matrix $\boldsymbol{C}_n^c$ and vector $\boldsymbol{t}^c$ respectively. The image plane coordinates $\boldsymbol{U} = (u, v)^T$ can be expressed as:

$$\boldsymbol{U} = m \begin{bmatrix} f/dx & 0 & u_0 \\ 1 & f/dy & v_0 \end{bmatrix} \boldsymbol{p}^c = m\boldsymbol{K}\boldsymbol{p}^c \tag{3}$$

where, $\boldsymbol{K}$ represents the sub-matrix of camera intrinsic parameter. Due to the deviation between the ideal image plane coordinates $(u, v)^T$ and practice one $(\tilde{u}, \tilde{v})^T$, the nonlinear error Eq. (3) is expressed as follow.

$$\delta\boldsymbol{U} = \boldsymbol{U} - \tilde{\boldsymbol{U}} = m \begin{bmatrix} f_u & 0 & \tilde{u} - u_0 \\ 0 & f_v & \tilde{v} - v_0 \end{bmatrix} \delta\boldsymbol{p}^c = \boldsymbol{H}_{up} \delta\boldsymbol{p}^c \tag{4}$$

4 Multi-antenna GNSS Attitude Determination Model

The installation of antennas is shown as Fig. 2: the antennas of GNSS receiver A, B and C are set in the same plane on vehicle-mounted platform. The GNSS frame (g-frame) has its origin at the center of antenna of GNSS receiver A. The antennas of GNSS receiver B and C are installed along the heading of vehicle and vertical direction of it. The x-axis points from A towards B, the z-axis points towards upside, which is perpendicular to the vehicle-mounted platform, and the y-axis completes a right-handed orthogonal frame, which points towards the right of vehicle [5].

4.1 GNSS Attitude Determination Using Direct Method

At each epoch, the coordinate of antenna of GNSS receiver A respect to the e-frame can obtained by the method of pseudo-range single point positioning, the baseline solution of antennas of GNSS receiver B and C respect to A, denoted by $d\boldsymbol{r}^e$, can be

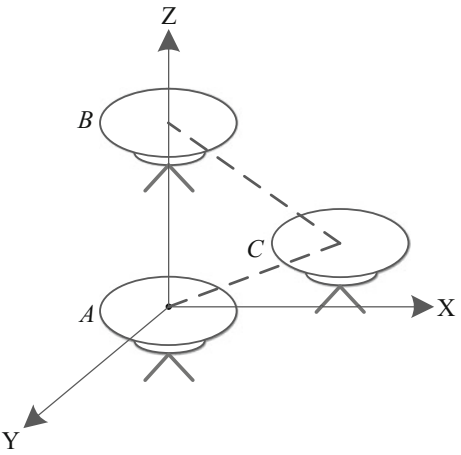


Fig. 2 The installation of multi-antennas GNSS receiver

obtained by carrier phase measurement. Hence, the baseline solution respect to n-frame, denoted by $d\boldsymbol{r}^n$, can be expressed as [6]:

$$d\boldsymbol{r}^n = \begin{bmatrix} dx_{AB}^n & dx_{BC}^n \\ dy_{AB}^n & dy_{BC}^n \\ dz_{AB}^n & dz_{BC}^n \end{bmatrix} = \boldsymbol{R}_e^n d\boldsymbol{r}^e \tag{5}$$

where $\boldsymbol{R}_e^n$ is the direction cosine matrix from e-frame to n-frame.

Define the coordinate of antenna B of GNSS receiver respect to b-frame is $(0, L_{AB}^b, 0)^T$, the transformation relation of B between above two frame is given as:

$$\boldsymbol{r}^n = R_b^n \boldsymbol{r}^b \Rightarrow \begin{bmatrix} dx_{AB}^n \\ dy_{AB}^n \\ dz_{AB}^n \end{bmatrix} = \begin{bmatrix} L_{AB}^b \cos\theta \sin\psi \\ L_{AB}^b \cos\theta \cos\psi \\ L_{AB}^b \sin\theta \end{bmatrix} \tag{6}$$

Hence, the initial yaw ψ and pitch θ can be expressed as:

$$\begin{aligned} \psi &= \arctan\left(dx_{AB}^n / dy_{AB}^n\right) \\ \theta &= \arctan\left(dz_{AB}^n / \sqrt{d(x_{AB}^n)^2 + d(y_{AB}^n)^2}\right) \end{aligned} \tag{7}$$

Define the coordinate of antenna C of GNSS receiver respect to b-frame is $(L_{AC}^b \sin\varepsilon^b, L_{AC}^b \cos\varepsilon^b, 0)^T$, we can get the Eq. (8):

$$\begin{bmatrix} L_{AC}^b \sin \varepsilon^b \\ L_{AC}^b \cos \varepsilon^b \\ 0 \end{bmatrix} = R_y(\gamma)R_x(\theta)R_z(-\psi) \begin{bmatrix} dx_{AC}^n \\ dy_{AC}^n \\ dz_{AC}^n \end{bmatrix} = \begin{bmatrix} \cos\gamma & 0 & -\sin\gamma \\ 0 & 1 & 0 \\ \sin\gamma & 0 & \cos\gamma \end{bmatrix} R_x(\theta)R_z(-\psi) \begin{bmatrix} dx_{AC}^n \\ dy_{AC}^n \\ dz_{AC}^n \end{bmatrix} \tag{8}$$

Hence, initial roll γ of vehicle-mounted platform can be obtained from the third row of Eq. 9:

$$\gamma = -\arctan\left(d\tilde{z}_{AC}^n / d\tilde{x}_{AC}^n\right) \tag{9}$$

Internal accord accuracy of attitude determination using the direct method in heading, pitch and roll is about 0.3–0.5°, 0.3–1.0° and 0.5–1.0° [1].

4.2 GNSS Attitude Determination Using Parameter Adjustment Method

Define at an epoch, the i set of baseline vector between n-frame and b-frame can be expressed as:

$$\boldsymbol{r}_{ij}^n = \boldsymbol{R}_b^n \boldsymbol{r}_{ij}^b, i, j = 1, \ldots, n \tag{10}$$

where the subscript i and j represent combination of antennas composing baseline, n = 3 represents the number of antennas, $\boldsymbol{R}_b^n$ represents the attitude cosine matrix composing the unknown parameters, $\boldsymbol{r}_{ij}^b$ and $\boldsymbol{r}_{ij}^n$ represent the real observation of baseline vector.

Assuming the three approximate attitude is ψ_0, θ_0 and γ_0, and the baseline vector $\boldsymbol{r}_{ij}^b$ observed in the body frame has no error. So the evaluated $\boldsymbol{r}_{ij}^n$ can be calculated according to the initial attitude as shown:

$$\boldsymbol{r}_{c,ij}^n = \tilde{\boldsymbol{R}}_{0,b}^n \boldsymbol{r}_{ij}^b \tag{11}$$

In this case, $\boldsymbol{r}_{c,ij}^n$ represents the calculated baseline vector, $R_{0,b}^n$ represents the attitude cosine matrix composing with the initial attitude. Assuming the correction vector of attitude obtained by GNSS is $\delta\boldsymbol{\varphi} = (\delta\psi, \delta\theta, \delta\gamma)^T$, the linearized attitude parameter adjustment Eq. (12) can be described as:

$$\boldsymbol{v}_{\boldsymbol{r}_{ij}^n} = \boldsymbol{A}_{ij}\delta\boldsymbol{\varphi} - \boldsymbol{L}_{ij} \tag{12}$$

where $A_{ij} = \left(\frac{\partial R}{\partial \psi}\boldsymbol{r}_{ij}^b, \frac{\partial R}{\partial \theta}\boldsymbol{r}_{ij}^b, \frac{\partial R}{\partial \gamma}\boldsymbol{r}_{ij}^b\right)$ represents the observation matrix, $\boldsymbol{v}_{\boldsymbol{r}_{ij}^n}$ represents residual error vector, $\boldsymbol{L}_{ij} = \boldsymbol{r}_{ij}^n - R_{0,b}^n\boldsymbol{r}_{ij}^b$ represents the observation vector. Hence, we can get the least square solution of correction of attitude parameters [7]:

$$\boldsymbol{X} = (\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{L} \tag{13}$$

The attitudes at the current epoch measured by multi-antenna of GNSS receiver is $\boldsymbol{\varphi} = \hat{\boldsymbol{\varphi}} + \delta\boldsymbol{\varphi}$. Internal accord accuracy of attitude determination using the least square method in heading, pitch and roll is about 0.1°, 0.2–0.5° and 0.5–2.0° [1].

5 Integrated Navigation Filter Model

5.1 Inertial Navigation Status Equation

In this paper, the geographic coordinate (ENU-frame) is considered as n-frame, the error of navigation parameters and inertial instrument are considered as state variables, and the strapdown inertial navigation error model was established. In addition, the scale factor error and installation error of gyro angle are without considering the inertial navigation error equation. The state vector can be expressed as:

$$\begin{bmatrix}\delta\boldsymbol{r}^n & \delta\boldsymbol{v}^n & \delta\varphi^n & \boldsymbol{\varepsilon}_b^n & \nabla_b^n\end{bmatrix}^T \tag{14}$$

where, $\delta\boldsymbol{r}^n = \begin{bmatrix}\delta\boldsymbol{\lambda} & \delta\boldsymbol{L} & \delta\boldsymbol{h}\end{bmatrix}$ represents the position error vector, $\delta\boldsymbol{v}^n = \begin{bmatrix}\delta\boldsymbol{v}_E & \delta\boldsymbol{v}_N & \delta\boldsymbol{v}_U\end{bmatrix}$ represents the speed error vector in ENU-frame, $\delta\varphi^n = \begin{bmatrix}\boldsymbol{\varphi}_E & \boldsymbol{\varphi}_N & \boldsymbol{\varphi}_U\end{bmatrix}$ represents the attitude error vector in ENU-frame, $\boldsymbol{\varepsilon}_b^n = \begin{bmatrix}\boldsymbol{\varepsilon}_{bx}^n & \boldsymbol{\varepsilon}_{by}^n & \boldsymbol{\varepsilon}_{bz}^n\end{bmatrix}$ represents the three axis component of gyro drift, $\nabla_b^n = \begin{bmatrix}\nabla_{bx}^n & \nabla_{by}^n & \nabla_{bz}^n\end{bmatrix}$ represents accelerometer bias. Then, the error equation of the strapdown inertial navigation system (SINS) can be expanded as a matrix form [8].

5.1.1 Position Update

$$\begin{aligned}\delta\dot{\boldsymbol{r}}^n &= \begin{bmatrix} 0 & 0 & -\frac{v_N}{(R_M+h)^2} \\ -\frac{v_E \sec L \tan L}{R_N+h} & 0 & -\frac{v_E \sec L}{(R_N+h)^2} \\ 0 & 0 & 0 \end{bmatrix}\delta\boldsymbol{r}^n + \begin{bmatrix} 0 & \frac{1}{R_M+h} & 0 \\ \frac{\sec L}{R_N+h} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\delta\boldsymbol{v}^n \\ &= F_{rr}\delta\boldsymbol{r}^n + F_{rv}\delta\boldsymbol{v}^n \end{aligned} \tag{15}$$

where $\delta \boldsymbol{r}^n$ refers to position error vector, $[\dot{\boldsymbol{r}}]$ refers to the differential coefficient form of $\boldsymbol{r}$, R_M refers to radius of curvature among meridian circle, R_N refers to radius of curvature among prime vertical circle.

5.1.2 Velocity Update

$$\begin{aligned}\delta\dot{\boldsymbol{v}}^n = &\begin{bmatrix} 2\omega_{i.e.}(v_N\cos L + v_U\sin L) + \frac{v_E v_N \sec^2 L}{R_N+h} & 0 & \frac{v_E(v_U - v_N\tan L)}{(R_N+h)^2} \\ -v_E\left(2\omega_{i.e.}\cos L + \frac{v_E\sec^2 L}{R_N+h}\right) & 0 & \frac{v_N v_U}{(R_M+h)^2} + \frac{v_E^2\tan L}{(R_N+h)^2} \\ -2\omega_{i.e.}v_E\sin L & 0 & -\left(\frac{v_E^2}{(R_N+h)^2} + \frac{v_N^2}{(R_M+h)^2}\right) \end{bmatrix}\delta\boldsymbol{r}^n \\ &+\begin{bmatrix} \frac{v_N\tan L - v_U}{R_N+h} & 2\omega_{i.e.}\sin L + \frac{v_E\tan L}{R_N+h} & -\left(2\omega_{i.e.}\cos L + \frac{v_E}{R_N+h}\right) \\ -2\left(\omega_{i.e.}\sin L + \frac{v_E\tan L}{R_N+h}\right) & -\left(\frac{v_U}{R_M+h}\right) & -\frac{v_N}{R_M+h} \\ -2(\omega_{i.e.}\cos L + \frac{v_E}{R_N+h}) & -\left(\frac{2v_U}{R_M+h}\right) & 0 \end{bmatrix}\delta\boldsymbol{v}^n \\ &+(\boldsymbol{f}^n\times)\boldsymbol{\varphi}^n + \nabla^n \\ &= F_{vr}\delta\boldsymbol{r}^n + F_{vv}\delta\boldsymbol{v}^n + (\boldsymbol{f}^n\times)\boldsymbol{\varphi}^n + \nabla^n \end{aligned} \tag{16}$$

where $\delta \boldsymbol{v}^n$ refers to velocity error vector.

5.1.3 Attitude Update

$$\begin{aligned}\dot{\boldsymbol{\varphi}}^n = &\begin{bmatrix} 0 & 0 & \frac{v_N}{(R_M+h)^2} \\ -\omega_{i.e.}\sin L & 0 & -\frac{v_E}{(R_N+h)^2} \\ \omega_{i.e.}\cos L + \frac{v_E\sec^2 L}{R_N+h} & 0 & \frac{-v_E\tan L}{(R_N+h)^2} \end{bmatrix}\delta\boldsymbol{r}^n + \begin{bmatrix} 0 & \frac{-1}{R_M+h} & 0 \\ \frac{1}{R_N+h} & 0 & 0 \\ \frac{\tan L}{R_N+h} & 0 & 0 \end{bmatrix}\delta\boldsymbol{v}^n - (\boldsymbol{\omega}_{in}^n\times)\boldsymbol{\varphi}^n - \boldsymbol{\varepsilon}^n \\ &= F_{er}\delta\boldsymbol{r}^n + F_{ev}\delta\boldsymbol{v}^n - (\boldsymbol{\omega}_{in}^n\times)\boldsymbol{\varphi}^n - \boldsymbol{\varepsilon}^n \end{aligned} \tag{17}$$

Where $\boldsymbol{\varphi}^n$ refers to attitude error vector.

5.1.4 Inertial Instrument Update

$$\dot{\boldsymbol{\varepsilon}}^n = 0, \dot{\nabla}^n = 0 \tag{18}$$

where ε^n refers to gyroscopic drift, ∇^n refers to Acceleration bias.

Hence, the inertial navigation status equation can be expressed as:

$$\delta\dot{\boldsymbol{x}} = F_I\delta\boldsymbol{x} + G_I\boldsymbol{w}_I \tag{19}$$

where subscript I refers to inertial navigation, F_I refers to the inertial navigation system transfer matrix, $\delta\boldsymbol{x}$ refers to the inertial navigation status vector, G_I refers to the noise matrix, $\boldsymbol{w}_I$ refers to the system noise vector, which can be expressed as:

$$F_I = \begin{bmatrix} F_{rr} & F_{rv} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ F_{vr} & F_{vv} & (f^n\times) & 0_{3\times3} & C_b^n \\ F_{er} & F_{ev} & -(\omega_{in}^n\times) & -C_b^n & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} \end{bmatrix}, \quad \delta\boldsymbol{x} = \begin{bmatrix} \delta\boldsymbol{r}^n \\ \delta\boldsymbol{v}^n \\ \delta\boldsymbol{\varphi}^n \\ \boldsymbol{\varepsilon}^n \\ \nabla^n \end{bmatrix},$$

$$G_I = \begin{bmatrix} 0_{3\times3} & C_b^n & 0_{3\times3} & I_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & -C_b^n & 0_{3\times3} & I_{3\times3} \end{bmatrix}^T, \quad \boldsymbol{w}_I = \begin{bmatrix} \boldsymbol{w}_{\varepsilon^b} \\ \boldsymbol{w}_{\nabla^b} \end{bmatrix}.$$

5.2 Vision Navigation Observation Equation

Due to the position difference between inertial sensors I and camera on the vehicle-mounted platform, the sensors outputs cause the navigation error, which is called by the lever arm effects [9]. Hence, we should take the correction from lever arm effects into account when establishing observation equation.

Define the vector from camera to the barycenter of inertial sensors is $\boldsymbol{l}_c^I$, according to the geometrical relationship in Fig. 1, the feature point P of vector $^cP^c$ respect to c-frame can expressed as:

$$\overrightarrow{{}^cP^c} = C_b^cC_n^bC_m^n(\boldsymbol{r}_P^m - \boldsymbol{r}_c^m) = C_b^cC_n^bC_m^n(\boldsymbol{r}_P^m - C_b^m\boldsymbol{l}_c^I - \boldsymbol{r}_I^m) \tag{20}$$

Where, $\boldsymbol{r}_{[\cdot]}^m$ refers to the position $[\cdot]$ respect to user-defined frame (m-frame), C_b^c, C_n^b and C_m^n refers to direction cosine matrix. Due to the difference between inertial sensors practical pose and its real values:

$$\begin{gathered} \boldsymbol{r}_I^m = \tilde{\boldsymbol{r}}_I^m + \delta\boldsymbol{r}_I^m \\ C_b^n = (\boldsymbol{I} - \boldsymbol{\varphi}^n\times)\tilde{C}_b^n \Rightarrow C_n^b = \tilde{C}_n^b(\boldsymbol{I} + \boldsymbol{\varphi}^n\times) \end{gathered} \tag{21}$$

Hence, Substituting Eq. (19) into Eq. (18) and omitting the high-order item, we can easily get:

$$\overrightarrow{{}^cP^c} = \widetilde{{}^cP_i^c} + C_b^c\tilde{C}_n^bC_m^n\delta\boldsymbol{r}_I^m - C_b^c\tilde{C}_n^b\left(C_m^n(\tilde{\boldsymbol{r}}_I^m - C_b^m\boldsymbol{l}_c^I - \boldsymbol{r}_c^m)\right)\times\boldsymbol{\varphi}^n \tag{22}$$

Where $\widetilde{cP^c}$ refers to the practice u-frame according to the backward reasoning of inertial output, $\overrightarrow{cP^c} = \boldsymbol{p}^c$, $\boldsymbol{r}_I^n = C_m^n \boldsymbol{r}_I^m$, and the correction of image coordinate can be obtained easily from Eqs. (4) and (22):

$$\delta \tilde{\boldsymbol{U}}_i = H_{up}\left(H_{pr}\delta \boldsymbol{r}_I^n + H_{pe}\boldsymbol{\varphi}^n\right) \tag{23}$$

Here, $\tilde{\boldsymbol{r}}_I^n$ is the positions measured by inertial sensors, and $\tilde{C}_n^b$ is the rotational matrix composed of attitude measured by inertial sensors, $H_{pr} = C_b^c \tilde{C}_n^b$, $H_{pe} = -C_b^c \tilde{C}_n^b \left(\tilde{\boldsymbol{r}}_I^n + C_b^n \boldsymbol{L}_c^l - \boldsymbol{r}_c^n\right) \times$.

Hence, the vision navigation observation equation can be expressed as:

$$\boldsymbol{z}_V = H_V \delta \boldsymbol{x} + \boldsymbol{v}_V \tag{24}$$

Where subscript V refers to vision navigation, $\boldsymbol{z}_V = [\delta \boldsymbol{U}]$ is the vision navigation observation vector, $H_V = [\, H_{up}H_{pr} \quad 0_{2\times 3} \quad H_{up}H_{pe} \quad 0_{2\times 6} \,]$ is the vision navigation observation matrix.

5.3 Multi-antenna GNSS Navigation Observation Equation

The satellite navigation observed quantities based on multi-antenna GNSS include position, velocity and attitude. The observation equation can be written as:

$$\begin{cases} \boldsymbol{z}_{Gr} = H_r \delta \boldsymbol{x} + \boldsymbol{v}_{Gr} \\ \boldsymbol{z}_{Gv} = H_v \delta \boldsymbol{x} + \boldsymbol{v}_{Gv} \\ \boldsymbol{z}_{G\varphi} = H_\varphi \delta \boldsymbol{x} + \boldsymbol{v}_{G\varphi} \end{cases} \tag{25}$$

Here, $\boldsymbol{z}_{Gr} = \boldsymbol{r}_I^n - \boldsymbol{r}_G^n$ refers to the position difference measured by inertial sensors and GNSS and corrected considering lever arm effects, $\boldsymbol{z}_{Gv} = \boldsymbol{v}_I^n - \boldsymbol{v}_G^n$ refers to its velocity difference, $\boldsymbol{z}_{G\varphi} = \boldsymbol{\varphi}_I^n - \boldsymbol{\varphi}_G^n$, and $\boldsymbol{\varphi}_I^n$ refers to its attitude difference considering lever arm effects, $\boldsymbol{\varphi}_G^n$ refers to the body attitude through adjustment without limited achieved from the baseline solution obtained by multi-antennas GNSS, i.e. Sect. 3.2.

Hence, the GNSS navigation observation equation can be expressed as:

$$\boldsymbol{z}_G = H_G \delta \boldsymbol{x} + \boldsymbol{v}_G \tag{26}$$

Where subscript G refers to GNSS navigation, H_G refers to GNSS navigation observation matrix, $\boldsymbol{z}_G$ refers to GNSS navigation observation vector.

5.4 Multi-Scale Extended Kalman Filter

In practical applications, multi navigation information fusion system usually has the problems of different data sampling rate, unsynchronized observation and the uneven sampling, therefore, the measured quantities cannot achieve real-time synchronization in the system filter [10]. Normally, data sampling rate of INS is approximately 100–200 Hz, and that of GNSS is about 1–10 Hz, and that of vision navigation system is not the same, such as the high sampling rate high speed and low resolution camera (up to 150 Hz), high resolution and low speed camera (about 1 Hz). Therefore, we need to select different observation models for different sampling rate of the system.

In the paper, a visual/multi-antenna GNSS attitude determination/inertial integrated navigation model is established based on the multi-scale extended Kalman filtering algorithm, which is used to describe the measurement of three different data sampling rates:

$$\begin{aligned} \boldsymbol{X}_{N,k+1} &= \boldsymbol{F}_N \boldsymbol{X}_{N,k} + \boldsymbol{\Gamma}_N \boldsymbol{W}_{N,k} \\ \boldsymbol{Z}_{i,k} &= \boldsymbol{H}_i \boldsymbol{X}_{i,k} + \boldsymbol{V}_{i,k} \end{aligned} \quad N = 3, i = 1, 2, 3 \tag{27}$$

where $\boldsymbol{X}_{N,k}$ is a n-dimensional state vector, $\boldsymbol{F}_N$ refers to the system transfer matrix, $\boldsymbol{\Gamma}_N$ refers to the system noise matrix, $\boldsymbol{W}_{N,k}$ refers to the system process noise vector. $\boldsymbol{Z}_{i,k}$ refers to the observed quantities of the i navigation system at the time of k, $\boldsymbol{H}_i$ refers to the observation matrix, $\boldsymbol{V}_{i,k}$ refers to the observation noise vector.

(1) When only existing GNSS measurements, the relative part of visual navigation should be omitted from the observation matrix and observation vector. Hence, the measurement equation is rewritten as:

$$z_V = \boldsymbol{H}_V \delta \boldsymbol{x} + \boldsymbol{v}_V \tag{28}$$

(2) When only existing vision measurements, the relative part of GNSS navigation should be omitted from the observation matrix and observation vector. Hence, the measurement equation is rewritten as:

$$z_G = \boldsymbol{H}_G \delta \boldsymbol{x} + \boldsymbol{v}_G \tag{29}$$

(3) When existing both GNSS and vision measurements, the measurement equation is rewritten as:

$$\begin{bmatrix} z_V \\ z_G \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_V \\ \boldsymbol{H}_G \end{bmatrix} \boldsymbol{\delta x} + \begin{bmatrix} \boldsymbol{v}_V \\ \boldsymbol{v}_G \end{bmatrix} \tag{30}$$

6 Tests and Results

6.1 Test Condition

In order to verify the effectiveness of proposed VMGINS algorithm, the INS/GNSS/CCD test based on vehicle-mounted platform is carried out, which includes three rover GNSS receivers called as M300 Pro [11], Hasselblad camera called as H4D-60, MEMS inertial sensor called as INS-B from Inertial Lab and data synchronization and acquisition device, besides, the based GNSS receiver has also be set up at known control points. The direction of baseline AB is set up among back-up moving direction of the vehicle, The baseline AC is set up among left-right direction and orthogonal with baseline AB. Meanwhile, the results of the laser inertial navigation system are used as attitude reference value.

The initial alignment (15 min) is selected as the reference value of the horizontal attitude angle, and the result of the optical prism northwards is used as the reference value of the azimuth attitude angle. In the test, the baseline vector AB is (−0.0081, −1.8240,0.0458), and the baseline vector AC is (−0.9496, −0.0068, −0.0001). The system error of heading between multi-GNSS and INS is 2°, which has been corrected in the test data. Assuming there is no GNSS installation error including position and attitude in the tests. The accuracy index can be shown as Tables 1 and 2.

6.2 Result Analysis

In order to analyze the performance of the VMGINS algorithm, the vehicle-mounted data of the integrated navigation sampling in the outdoor artificial marking field is processed. The marking field is mounted with about 25 round

Table 1 Camera essential parameters

Focus length	Resolution ratio	Size of CCD	Size of pixel
35.2 mm	8956 × 6708	40.2 × 53.7 mm	6.0 μm

Table 2 Accuracy index of INS

Accuracy index	Laser INS	INS-B
Bias stability of Gyro $(^\circ/h)$	0.02	1
Random walk of Gyro $(^\circ/\sqrt{h})$	0.003	0.004
Bias stability of Acc (mg)	0.06	0.005
Data sampling rate (Hz)	200	200

artificial marks, and a total of four artificial marking fields were arranged. The schematic diagram of VMGINS test on vehicle-mounted platform in artificial mark field is shown in Fig. 3.

When the vehicle-mounted integrated navigation platform drove to the artificial marking field area, the CCD sensor took the picture of known control points on the wall (The effect of pictures was shown in Fig. 4).

800 s' dynamic experimental data is Selected as the measurement of VMGINS (epoch 36,940–37,740 s), in which the vehicle trajectory is plotted as shown in Fig. 5. The estimated position had a span of 535.511 m in the east-west direction and 291.745 m in the south-north direction. The line of attitude determination is plotted in Fig. 6.

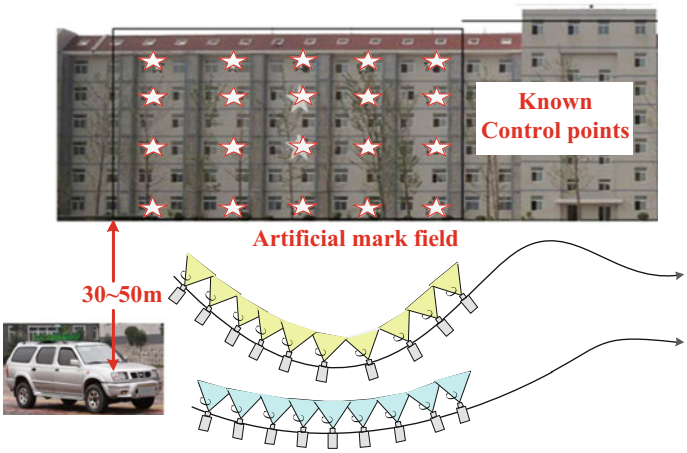


Fig. 3 The schematic diagram of vehicle-mounted integrated navigation test in artificial mark field

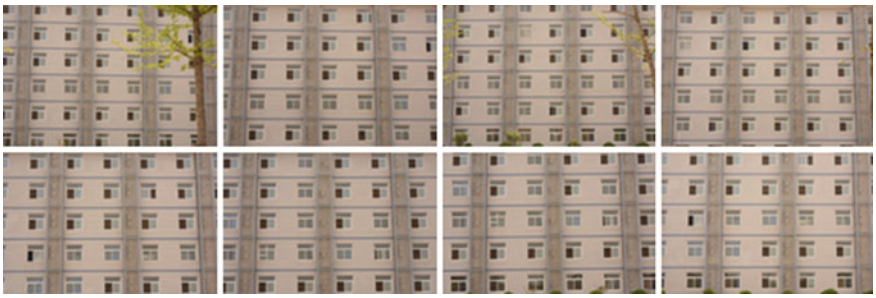


Fig. 4 The pictures of artificial mark field form different time during the test

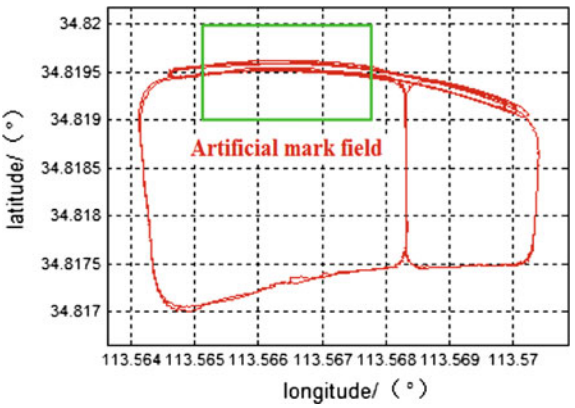


Fig. 5 The trajectory of the vehicle during the test

The data1 (blue) in the left of Fig. 6 refers to the attitude determination of integrated navigation, and the data2 (red) refers to the attitude determination of referenced ILIS. The data (blue) in the right of Fig. 6 refers to the difference of attitude determination between integrated navigation and the referenced ILIS. Besides, the green rectangle area in Fig. 6 indicates that the experimental data is the fused navigation results from VMGINS with vision observation when driving to the artificial marking field area.

According to Fig. 6 and Table 3, the test can draw the conclusion:

During the green rectangle area in Fig. 6, such as the graph of epoch from 36,993 to 37,053 s, the heading of VMGINS with vision observation has an accuracy of 0.243° in Table 3, and the accuracy of heading without vision observation is −0.615°. Similarly, the accuracy of pitch with vision observation is 0.732° comparing with the one without aiding (1.343°). In addition, the roll of VMGINS with aiding and the one without aiding have the similar accuracy (−0.178° comparing with 0.139°), because the GNSS installation error was not taken into account.

The max values of heading, pitch and roll are −4.945, −4.443 and −4.628 m, corresponding to the outliers of graph shown in the right of Fig. 6, the reason of which is the suddenly large angle turning of vehicle-mounted platform. It can be seen that the result of VMGINS with vision observation is much better than the one without aiding.

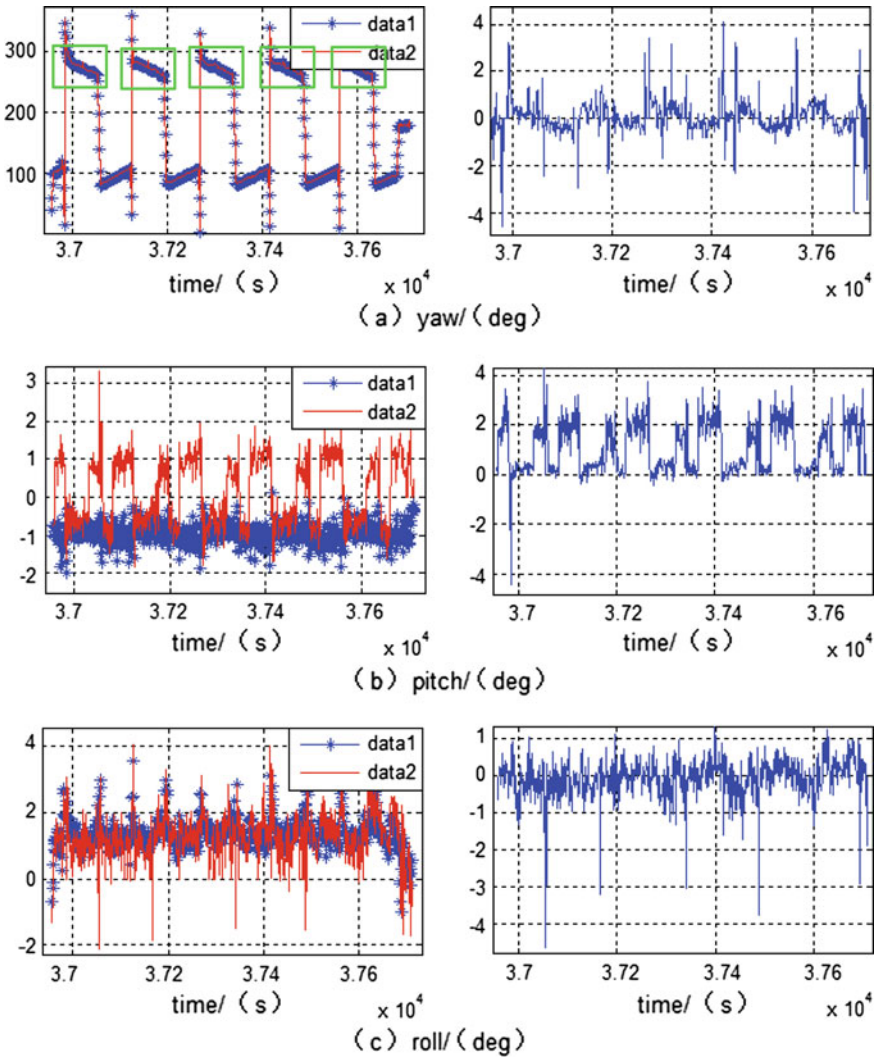


Fig. 6 The comparison of attitude determination between integrated navigation and reference in different time (Color figure online)

Table 3 The difference of between VMGINS and referenced ILIS

Observation type	Epoch/(s)	Heading/(°)	Pitch/(°)	Roll/(°)
Vision/GNSS/INS	36,993–37,053	0.308	0.764	−0.256
	37,130–37,190	0.263	0.603	−0.189
	37,247–37,337	0.040	0.948	−0.119
	37,423–37,483	0.236	0.684	−0.262
	37,571–37,631	0.370	0.661	−0.065
RMSE/(°)		*0.243*	*0.732*	*−0.178*
GNSS/INS	The rest epoch	−0.615	1.343	0.139
All observation	All epoch	−4.345	−4.443	−4.628

7 Conclusion

The paper studies VMGINS algorithm, deducing the tight-coupling observation model of vision/inertial, fusing the measurements of multi-antenna GNSS and vision with inertial measurement using the multi-scale extended Kalman filter, and implementing the vehicle-mounted integrated navigation test of VMGINS. Finally, performance of VMGINS algorithm was verified, the heading outputs of which has an accuracy of 0.243°. The paper provided an integrated navigation method for denied environment of satellite signal.

References

1. Zhang F, Chai Y, Chai H et al (2016) Analysis on precision of two attitude determination methods using GNSS multi-antenna data. J Inert Technol 24(1):30–35
2. Zhai H, Cong L, Qin H et al (2012) Method for MEMS-INS/GPS integrated attitude determination based on single antenna. Comput Eng Des 33(10):3999–4003
3. Dong M (2014) Research on key technology of GNSS/INS/vision integrated navigation information fusion. Information Engineering University, pp 66–73
4. Yong L, Mahmoud E, Andrew D (2012) Attitude determination by integration of MEMS inertial sensors and GPS for autonomous agriculture applications. GPS Solut 16:41–52
5. Cohen CE (1996) Attitude determination. Glob Position Syst Theory Appl (Prog Astronaut Aeronaut) 164:519–538
6. Lu G (1995) Development of a GPS multi-antenna system for attitude determination. The University of Calgary, Calgary, Canada
7. SONG C, HAO J-M, WANG B-H (2015) Real-time high-frequency GNSS attitude determination based on DSP. Glob Position Syst 40(2):53–57
8. Shin EH (2001) Accuracy improvement of low cost INS/GPS for land applications. The University of Calgary, Calgary, Canada, pp 28–35
9. Wang L, Ding X, Zhang Z (2011) Relative pose measurement based on tightly-coupled INS/vision. J Chin Inert Technol 19(6):686–691
10. Fu M, Deng Z, Yan L (2010) Kalman filter theory and the application on navigation system, 2nd edn. Science Press, Beijing, pp 149–159
11. Zhang K (2015) Combined multi-GNSS attitude determination theory and algorithms. Information Engineering University, pp 55–65

An Improved Vector Tracking Loop of Ultra-Tight Integration for Carrier Phase Tracking

Wei Li, Jiahe Xia and Guoda Cheng

Abstract An improved vector tracking loop is proposed to solve the carrier phase tracking problem of typical INS/GNSS ultra-tight integrated system, in which the carrier phase residual cannot be tracked properly and the navigation message demodulation cannot be conducted. Federated filter architecture of ultra-tight integration composed by pre-filters and integration filter is adopted. A non-coherent pre-filter is modeled for carrier phase tracking, of which the gain is adjusted by carrier-to-noise ratio adaptively. The carrier phase, frequency and code phase are vector tracked by pre-filter and navigation filter estimation concurrently. Experimental results indicate that the proposed vector tracking loop can improve the carrier phase tracking performance, eliminate the carrier phase residual of typical vector tracking solution effectively, with a satisfactory dynamic tracking performance.

Keywords Ultra-tight integration · Deep integration · Vector tracking loop · Carrier phase tracking

1 Introduction

Compare to the traditional inertial navigation system and global navigation satellite system (INS/GNSS) integrated system, ultra-tight (or deep) integration system combines the baseband signal processing, signal tracking, navigation solution and integrated navigation with estimating filter. According to the implementation of estimating filter, the system architecture can be classified into centralised structure and federated structure. With different pre-filter design, the federated structure can

W. Li (✉) · J. Xia · G. Cheng
AVIC Xi'an Flight Automatic Control Research Institute, Xi'an 710065, China
e-mail: willli618@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_38

be further divided into coherent and non-coherent system [1]. The significant feature of ultra-tight integration is that replacing the scalar tracking loops of the traditional receivers with vector tracking loop. With the benefits of improving navigation performance in dynamic and interference environment, ultra-tight integration has be known as lead-edge technology of the INS/GNSS integrated system and attracted recent research interest.

The basic idea of vector tracking technology is that estimating the Doppler shift with ephemeris and the position, velocity and acceleration of the user, and then solving the Numerically Controlled Oscillator (NCO) control signal, eventually conducting the carrier and code tracking with information of overall satellite channels. Generally, the navigation systems hardly meet the accuracy requirement of carrier phase vector-tracking which is centimetre-level. Therefore, the traditional vector tracking solutions adopt a vector frequency lock loop (VFLL) and vector delay lock loop (VDLL) for carrier frequency and code phase tracking respectively. However, due to the insufficient tracking accuracy of VFLL, the carrier phase cannot be tracked properly. The carrier phase residuals in the baseband signals lead to the failure of navigation message demodulation, and reduce the navigation accuracy.

To overcome such issue, federated structure is adopted, and the carrier NCO is controlled with estimations of pre-filters [2]. Further, dual-filter structure is proposed to avoid the coupling of states [3]. These solutions essentially conduct scalar carrier tracking with pre-filters, which degrade potential benefits of vector tracking in dynamic and interference environment. Meanwhile, a phase lock loop (PLL) is put forward in a VFLL\VDLL vector tracking structure for message demodulation [4]. This solution overcomes the demodulation issue of vector tracking, while the navigation accuracy is still limited by the tracking accuracy. A new tracking approach is proposed in [5], which is tracking the carrier phase with pre-filters, and tracking the carrier frequency and code phase with navigation filter. This approach implements the carrier phase tracking with an ultra-tight integrated architecture, while the tracking performance of carrier frequency is influenced by the system velocity accuracy.

In this paper, an improved vector tracking loop for ultra-tight integration is proposed based on non-coherent system architecture. Based on software receiver tracking algorithm, estimations of pre-filers and navigation filter are adopted to tracking the carrier frequency, carrier phase and code phase simultaneously for optimal tracking accuracy.

Section 2 describes the system structure and key components of filter modelling. Section 3 discusses the error source of GNSS signal tracking, and describes detail formulations of the proposed tracking algorithm. Section 4 conducts a simulation test with a dynamic scenario. Section 5 presents the conclusions and future research directions.

2 Non-coherent System Structure and Filter Model

2.1 System Structure

Figure 1 shows a block diagram of proposed non-coherent ultra-tight integration structure, which is composed by RF front-end, baseband signal processing, ultra-tight integration filters, INS solution and vector tracking. The ultra-tight integration filters applies federated architecture with multiple pre-filters and 1 master filter.

The measurement vector of the proposed non-coherent pre-filters is composed by outputs of carrier phase discriminator and code phase discriminator. The benefits of this design are as follow: to avoid the nonlinear filter problem of typical coherent pre-filters which applies baseband signal to form the measurements vector; to overcome the tracking accuracy loss of typical non-coherent approach caused by carrier frequency discriminator application; to improve the noise modeling accuracy by applying the various existing research results of discriminators modeling.

The estimation results of pre-filters and INS solution are further processed by integration navigation filter to estimate the INS errors, the clock error and clock drift of receiver. The NCO control signals are calculated with compensated INS results, ephemeris, and the residual estimations of pre-filters which include: carrier phase, frequency, frequency rate and code frequency.

2.2 Filter Modeling

To improve the modelling accuracy and reduce state order, the pre-filter model calculates the measure noise covariance matrix with real-time estimation of

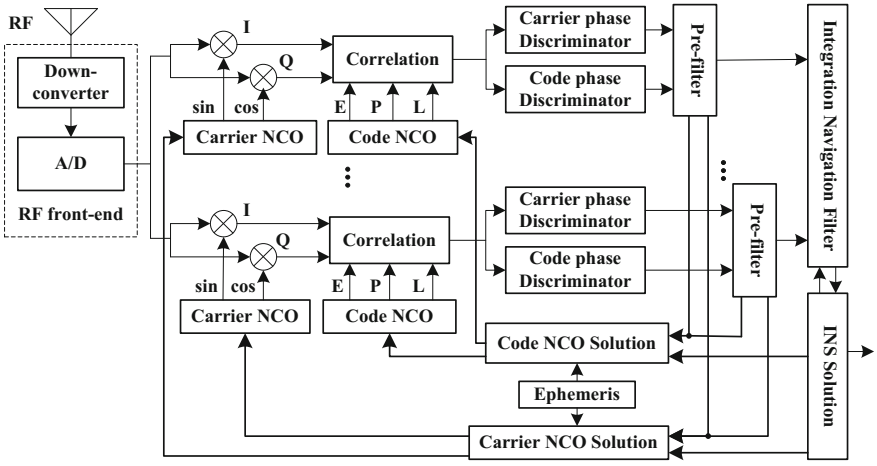


Fig. 1 Non-coherent ultra-tight integration structure

carrier-to-noise ratio (C/N0), instead of estimation of baseband signal amplitude. The state vector of pre-filter is shown as below.

$$X = [\,\delta\ddot{\phi} \quad \delta\dot{\phi} \quad \delta\phi \quad \delta\tau\,]^T \tag{2.1}$$

where the state elements are carrier frequency rate error (rad/s^2), carrier frequency error (rad/s), carrier phase error (rad) and code phase error (chips) respectively.

The state dynamics of pre-filter are given as below.

$$\frac{d}{dt}\begin{bmatrix}\delta\ddot{\phi}\\ \delta\dot{\phi}\\ \delta\phi\\ \delta\tau\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & \beta & 0 & 0\end{bmatrix}\begin{bmatrix}\delta\ddot{\phi}\\ \delta\dot{\phi}\\ \delta\phi\\ \delta\tau\end{bmatrix} + \begin{bmatrix}2\pi/\lambda & 0 & 0 & 0\\ 1 & 2\pi f & 0 & 0\\ 0 & 1 & 2\pi f & 0\\ 0 & \beta & 2\pi f\beta & 1\end{bmatrix}\begin{bmatrix}w_a\\ w_d\\ w_b\\ w_{\delta\tau}\end{bmatrix} \tag{2.2}$$

where β is the conversion from radians to chips, f and λ are carrier frequency and wave length, w_a is the driving noise of line-of sight acceleration, w_d is the driving noise of clock drift, w_b is the driving noise of clock bias, and the $w_{\delta\tau}$ is the driving noise of code tracking error. The spectral densities of w_d and w_b can be solved by Allan Variance analysis of oscillator.

The measurement vector of pre-filter is given as below.

$$Z = \begin{bmatrix}\delta\phi\\ \delta\tau\end{bmatrix} = \begin{bmatrix}\arctan(Q_P/I_P)\\ \frac{\sqrt{I_E^2+Q_E^2}-\sqrt{I_L^2+Q_L^2}}{\sqrt{I_E^2+Q_E^2}+\sqrt{I_L^2+Q_L^2}}\end{bmatrix} \tag{2.3}$$

The arctangent carrier phase discriminator and normalized early-late power code phase discriminator is applied here for better linear performance. In practical, other alternative discriminators can be applied to meet the specific requirements.

The measurement equation takes form as:

$$Z = \begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\delta\ddot{\phi}\\ \delta\dot{\phi}\\ \delta\phi\\ \delta\tau\end{bmatrix} + \begin{bmatrix}v_\phi\\ v_\tau\end{bmatrix} \tag{2.4}$$

where v_ϕ and v_τ are measurement noises of carrier phase discriminator and code phase discriminator respectively, of which the noise variance is the function of C/N0 as below [6]:

$$\begin{cases}\sigma_\phi^2 \approx \frac{1}{2C/N_0T}\left(1+\frac{1}{2C/N_0T}\right)\\ \sigma_\tau^2 \approx \frac{d}{4C/N_0T}\left(1+\frac{2}{(2-d)C/N_0T}\right)\end{cases} \tag{2.5}$$

where T is the coherent integration time, d is code correlator spacing. The narrow to wide power ratio method is applied here to calculate C/N0 for better performance in low signal to noise environments [7].

The model of integration navigation filter is similar as typical tight integration filter design, the measurements and measurement noise covariance can be derived from pre-filters estimations, the specific equations can be found in [5].

3 Improved Vector Tracking Algorithm

The causes for insufficient accuracy of typical vector tracking can be further analysed from the perspective of signal tracking error sources.

For a typical scalar tracking receiver, the error sources of various phase lock loop and frequency lock loop include: dynamic error and phase/frequency jitter error. The later one can be further classified into: thermal noise, jitter by motion and vibration, jitter by Allan Variance [8].

Accounting for these error sources, the parameter design of the scalar tracking loop needs to balance the dynamic error and jitter error. Typical vector tracking applies overall user dynamic information for signal tracking, which accounts for dynamic error above. While, the remaining jitter errors cannot be tracking properly, which lead the carrier phase/frequency residual and influence the accuracy of tracking and navigation.

Based on analysis above, to improve the tracking accuracy, the improvements of vector tracking algorithm are conducted as: to track the dynamic error with compensated INS results and ephemeris; to track residual of carrier phase, carrier frequency and code frequency caused by the jitter error with pre-filters estimations. Corresponding vector tracking NCO equations are derived as:

$$\begin{cases} f_{carr} = f_{IF} + \frac{1}{\lambda_{carr}} \vec{u} \cdot \left(\vec{V}_{INS} - \vec{V}_{Sat} \right) + f_{PFcarr} \\ \phi_{carr} = rem(\phi_{carr}^{-} + \delta\hat{\phi} + \delta\dot{\hat{\phi}} \cdot T + \delta\ddot{\hat{\phi}} \cdot T^2, 2\pi) \\ f_{\text{code}} = f_{\text{codebasis}} + \frac{K}{\lambda_{carr}} \vec{u} \cdot \left(\vec{V}_{INS} - \vec{V}_{Sat} \right) + f_{PFcode} \end{cases} \tag{3.1}$$

where f_{IF} and $f_{\text{codebasis}}$ are nominal carrier intermediate frequency (IF) and code reference frequency respectively, λ_{carr} is the carrier wave length, $\vec{u}$ is the unit line-of-sight vector from receiver to satellite, ϕ_{carr} and ϕ_{carr}^{-} are current and previous carrier phase, $\delta\hat{\phi}$, $\delta\dot{\hat{\phi}}$ and $\delta\ddot{\hat{\phi}}$ are estimations of pre-filter, rem represents remainder operation, T is update period of pre-filter, K is the ratio between code frequency and carrier frequency which is 1/1540 for GPS L1 band, f_{PFcarr} and f_{PFcode} are carrier frequency and code frequency error estimated by pre-filters as below:

$$\begin{cases} f_{PFcarr} = f_{PFcarr}^{-} + \dfrac{1}{2\pi}(\delta\hat{\dot{\phi}} + \delta\hat{\ddot{\phi}} \cdot T) \\ f_{PFcode} = f_{PFcode}^{-} + K \cdot \dfrac{(\delta\hat{\tau} - \delta\hat{\tau}^{-})}{2\pi T^2} \end{cases} \tag{3.2}$$

where f_{PFcarr}^{-} and f_{PFcode}^{-} are estimation of previous update period.

4 Simulation Test

With test environment set up as Fig. 2, the proposed vector tracking algorithm and its dynamic performance are tested.

As shown in Fig. 3, a typical circle trajectory is generated by INS trajectory generator, of which the max velocity is 300 m/s, max acceleration is 5 g. The generated trajectory is further injected into a GNSS signal simulator to simulate the GPS L1 RF signal, the C/N0 level is set to 45–46 dBHz during the simulation. IF signal is sampled by an IF data logger, the sampling frequency is 16.368 MHz, and the IF frequency is 4.124 MHz.

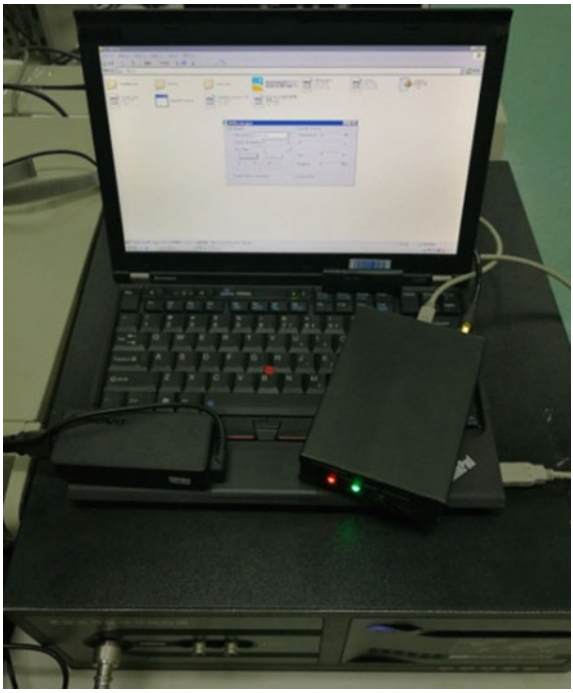


Fig. 2 Test environment

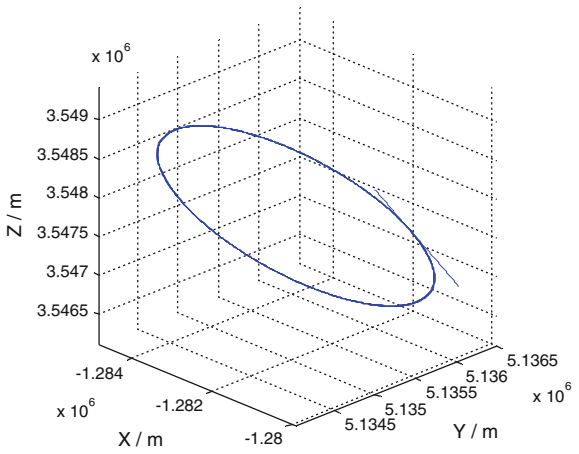


Fig. 3 The track of scenario

Table 1 Error characteristics of inertial sensors

Gyro bias (deg/h)	Gyro random walk (deg/$\sqrt{}$h)	Accelerometer bias (mg)	Accelerometer random walk (m/s/$\sqrt{}$h)
100	0.2	5	0.15

Table 1 shows the inertial error configuration which is to generate measurements of typical low-cost MEMS gyroscope and accelerometer.

With the simulated inertial data and GPS IF data, typical VFLL\VDLL vector tracking loop and the proposed vector tracking loop are adopted for ultra-tight integration navigation solution respectively. Details are compared with 4 s base-band signal at the same time interval.

As shown in Fig. 4, influenced by the carrier phase residual, the wave form overturn and distortion occur in prompt inphase signal (I_P), and signal power leaks into the prompt quadraphase signal (Q_P), and then the message demodulation cannot be conducted. As shown in Fig. 5, the quadrature demodulation of baseband signal is conducted properly, which means the proposed vector tracking loop can tracking the carrier phase residual correctly.

Figure 6 shows the NCO control signals calculated by the integrated navigation filter and pre-filter. Compared to the integrated navigation filter NCO, amplitude of the pre-filter NCO is relevantly small, and its STD is 0.44 Hz. This result indicates that the integrated navigation filter NCO tracking the user dynamic, and the pre-filter NCO only need to track the jitter error with small bandwidth.

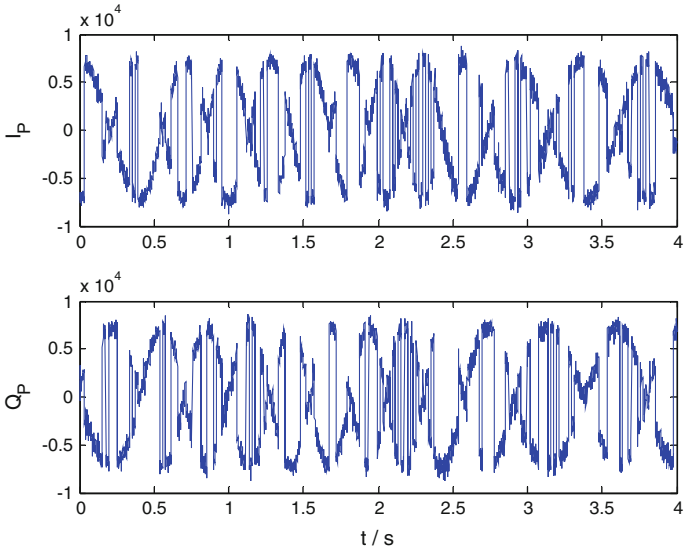


Fig. 4 Baseband signal of VFLL\VDLL

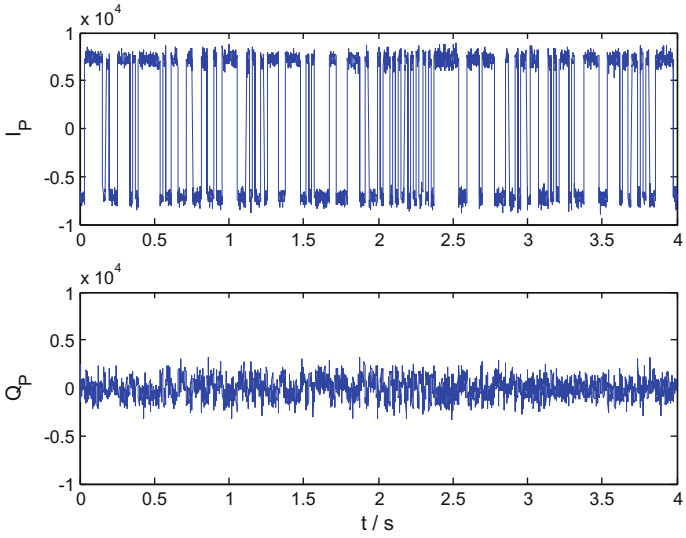


Fig. 5 Baseband signal of vector tracking loop

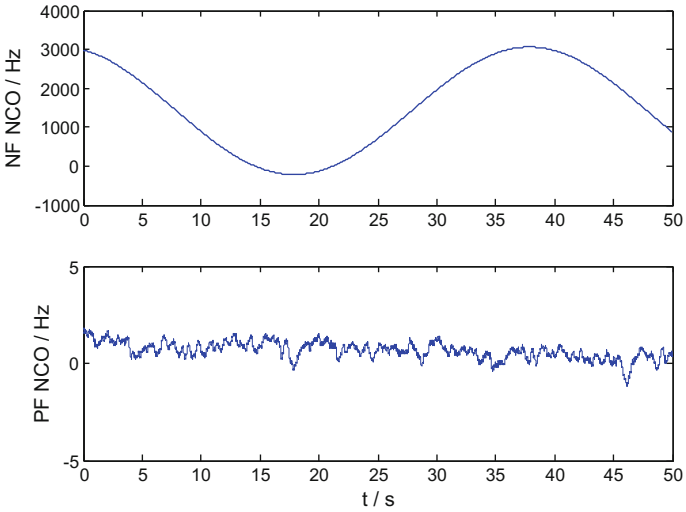


Fig. 6 NCO signal of improved vector tracking loop

5 Summary

A high accuracy system structure and system model is proposed based on non-coherent ultra-tight integration. An improved vector tracking algorithm is proposed to over carrier phase tracking problem of the typical VFLL\VDLL vector tracking loop. The dynamic performance is tested by Semi-physical simulation. The proposed vector tracking loop has benefits with lower complexity, higher modelling and tracking accuracy than typical vector tracking solutions. Further research would be conducted in INS/baseband signal synchronization, performance test in interference environments and integrated system design.

References

1. Groves PD, Mather CJ, Macauly AA (2007) Demonstration of non-coherent deep INS/GPS integration for optimized signal-to-noise performance. In: ION GNSS 20th international technical meeting of the satellite division, Fort Worth TX
2. Petovello MG, Lachapelle G (2006) Comparison of vector-based software receiver implementations with application to ultra-tight GPS/INS integration. In: ION GNSS, Fort Worth TX
3. Luo Y, He X, Wu W et al (2011) Base band signal pre-filter model based on dual-filter structure. J Chin Inertial Technol 19(5): 584–589 (Ch)
4. Zeng G, Shen Q, Li D et al (2015) Modeling and simulation on GNSS vector tracking loop based on extended Kalman filter. AataArmamentarii 36(3):530–538 (Ch)

5. Luo Y, Wu W, Babu R et al (2014) A double-filter-structure based COMPASS/INS deep integrated navigation system implementation and tracking performance evaluation. Sci China Inf Sci 57: P012206:1–14
6. Driscoll CO, Petovello MG, Lachapelle G (2011) Choosing the coherent integration time for Kalman filter-based carrier-phase tracking of GNSS signals. GPS Solut 15:345–356
7. Groves PD (2005) GPS signal to noise measurement in weak signal and high interference environments. In: ION GNSS 18th international technical meeting of the satellite division, Long Beach, CA
8. Xie G (2012) Principles of GPS and receiver design. Publishing House of Electronics Industry, Beijing, pp 292–295

Part IV
PNT New Concept, New Methods and New Technology

Research on Doppler Locating Method of LEO Satellite Backup Navigation System

Zhixin Deng, Guangwei Fan and Chenglong He

Abstract Fast response and low orbit satellites have advantages just like wide coverage, all-weather work, low cost of system construction, and so on, very suitable for positioning service as navigation sources for ground users independent of GNSS, so as to realize the GNSS effective backup navigation means in electromagnetic interference environment. Firstly, low orbit satellite navigation system ground coverage were briefly analyzed, pointed out that in the conditions of global coverage at different time in different locations users may have 1–3 different number of visible satellites. In order to solve the problem of continuous positioning solution under the condition of user different low orbit visible satellite number, and to avoid the system complexity increases due to high precision synchronization required by traditional pseudo-range positioning system, according to the characteristics of the fast speed of the low orbit satellite, a variety of positioning methods were proposed under the condition of different visible satellite number and only using the Doppler satellite observation information, positioning solution iteration formulas were given, and simulation analysis results show that this method can meet the three-dimensional positioning error is less than 20 m in most area. The proposed localization solution method can achieve single point positioning, and can realize continuous observation positioning, using Newton iterative least squares method, with low computational complexity, which is very suitable for fast positioning application for fixed or low dynamic users near the ground.

Keywords LEO satellite · Backup navigation · Doppler locating · Least square

Z. Deng (✉) · G. Fan · C. He
State Key Laboratory of Satellite Navigation System and Equipment Technology, Shijiazhuang, China
e-mail: chinadengzhixin@139.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_39

1 Introduction

A large number of low-cost, no special anti-jamming measures of the GNSS navigation terminal will not work in a complex electromagnetic interference environment. In the North and South polar regions, the GNSS system constellation with low tilt design, resulting in significantly reduced coverage. Therefore, the construction of GNSS backup radio navigation system is a hotspot in recent years.

Low-orbit satellites have the advantages of wide coverage area, all-weather workability and good observability. The micro-satellites and fast response satellites also have advantages such as short construction period and low cost. Therefore, low-orbit satellites and fast-response satellites platform is very suitable as a navigation source for the ground users to provide GNSS-independent navigation and positioning services.

GNSS navigation system users through the observation of at least four GNSS satellite signal pseudorange information for positioning solution. However, there are two problems in the application of this method to the LEO satellite navigation system. One is that, due to the limited number of LEO satellites, ground users in most areas can not observe four or more LEO satellites at the same time. The prerequisite of positioning solution is that many satellites must realize strict time synchronization, which will greatly increase the construction complexity of the system. The low-orbiting satellites have good Doppler-observability because of their fast velocity. Therefore, it is the first choice for low-earth-orbit satellite navigation system to use the Doppler observation information to locate the solution.

Recently, the United States Iridium communications company announced its development of satellite timing and positioning (STL) system already has the capacity as a GPS backup or supplement. The system provides users with less-vulnerable navigation services through 66 low-orbit satellites and Doppler positioning systems, but does not give a specific location solution. In January 2011, Asher et al. [1] proposed the idea of using single LEO satellites to provide stationary means for ground stationary users. However, the algorithm requires users to transmit signals to low-orbit satellites, This is not conducive to concealment of user location. In 2013, Pan Lei of Southwest Jiaotong University analyzed the principle of Doppler localization in his master thesis [2], and got the conclusion that the Doppler positioning accuracy is related to the precision and velocity precision of Doppler measurement. But do not give a solution of Doppler positioning solution method under the conditions of different number of visible stars.

To sum up, how to use a small amount of Doppler observation information to solve the continuous positioning solution problem under the condition of different visible stars is key issues in a low-earth satellite backup navigation system.

2 Coverage Analysis of LEO Satellite Backup Navigation System

In this paper, Iridium GPS backup navigation system is taken as an example to analyze the influence of different parameters on system coverage by adjusting constellation parameters.

Reference to the United States Iridium system [3], orbital altitude is 780 km, inclination of satellite orbit and the equatorial is 86.4°, the constellation consists of six orbital planes, each track surface 11 stars. Its global satellite layout is shown in Fig. 1. Obviously, satellite visibility is better in high latitudes and lowest in equatorial regions, so only the constellation coverage performance in the equatorial region is simulated. Figure 2 shows the typical LEO satellite in the equatorial region. The basic layout unit visibility distribution shown in Fig. 3, in which blue represents visible 1 star, green represents visible 2 stars, red represents visible 3 stars.

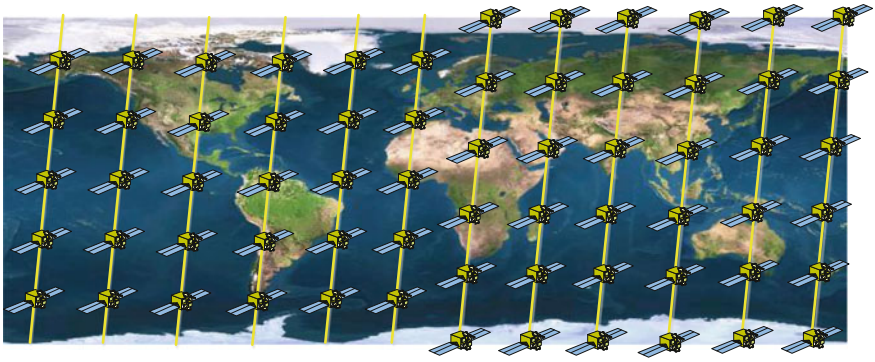


Fig. 1 LEO satellites constellation distribution picture of iridium satellite system

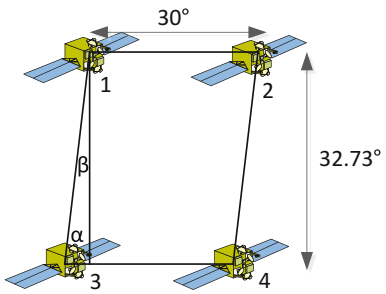


Fig. 2 Typical distribution unit of LEO satellites in equatorial area

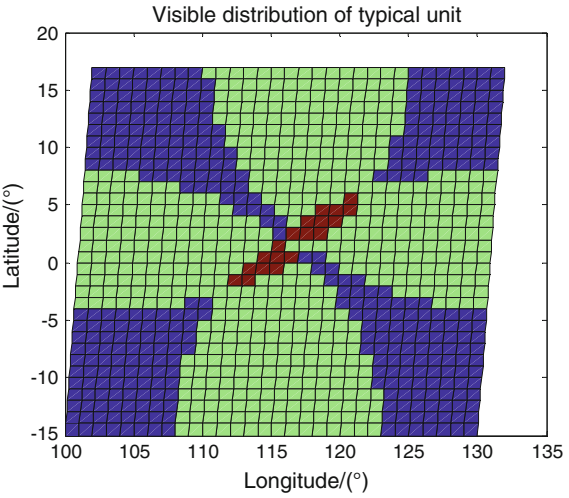


Fig. 3 Visible distribution of typical distribution unit when defilade angle is 5° (Color figure online)

3 Doppler Location Algorithm for Multi-scene

3.1 Newton Least Squares Iterative Doppler Location

From the coverage analysis of LEO satellite navigation system, we can see that the number of visible satellites of terrestrial users is determined by such factors as sheltering angle, orbital inclination, orbital altitude and the number of satellites per orbit. In general, the number of visible satellites at a given site is Ranging from 1 to 4, so must be able to adapt to different number of visible stars positioning solution method to meet the different scenarios of continuous navigation and positioning requirements.

In this paper, we propose a Doppler localization method which can be applied to multi-scene. This method can satisfy the demand of continuous positioning solution with different number of visible stars to the maximum extent. The results of the positioning solution at each epoch are obtained by the Newton least-squares iterative process. The iterative calculation formula is:

$$[x_{i+1}, y_{i+1}, z_{i+1}]^T = [x_i, y_i, z_i]^T + [\Delta x_{i+1}, \Delta y_{i+1}, \Delta z_{i+1}]^T$$

In which $[x_i, y_i, z_i]^T$ is the iteration result for the ith (i is natural number) iteration, $[\Delta x_{i+1}, \Delta y_{i+1}, \Delta z_{i+1}]^T$ is the correction result for the $i+1$ iteration, and there are:

$$[\Delta x_{i+1}, \Delta y_{i+1}, \Delta z_{i+1}]^T = (G^T \times G)^{-1} \times G^T \times Y$$

In which

$$G = \begin{bmatrix} \frac{V_{x1}\cdot f_0}{f_{d1}\cdot C} + \frac{x_i - x^{(1)}}{r^{(1)}} & \frac{V_{y1}\cdot f_0}{f_{d1}\cdot C} + \frac{y_i - y^{(1)}}{r^{(1)}} & \frac{V_{z1}\cdot f_0}{f_{d1}\cdot C} + \frac{z_i - z^{(1)}}{r^{(1)}} \\ \frac{V_{x2}\cdot f_0}{f_{d2}\cdot C} + \frac{x_i - x^{(2)}}{r^{(2)}} & \frac{V_{y2}\cdot f_0}{f_{d2}\cdot C} + \frac{y_i - y^{(2)}}{r^{(2)}} & \frac{V_{z2}\cdot f_0}{f_{d2}\cdot C} + \frac{z_i - z^{(2)}}{r^{(2)}} \\ & \cdots \quad \cdots \quad \cdots & \\ \frac{V_{xn}\cdot f_0}{f_{dn}\cdot C} + \frac{x_i - x^{(n)}}{r^{(n)}} & \frac{V_{yn}\cdot f_0}{f_{dn}\cdot C} + \frac{y_i - y^{(n)}}{r^{(n)}} & \frac{V_{zn}\cdot f_0}{f_{dn}\cdot C} + \frac{z_i - z^{(n)}}{r^{(n)}} \end{bmatrix},$$

$$Y = \begin{bmatrix} \frac{V_{x1}\cdot f_0}{f_{d1}\cdot C}\left(x^{(1)} - x_i\right) + \frac{V_{y1}\cdot f_0}{f_{d1}\cdot C}\left(y^{(1)} - y_i\right) + \frac{V_{z1}\cdot f_0}{f_{d1}\cdot C}\left(z^{(1)} - z_i\right) - r^{(1)} \\ \frac{V_{x2}\cdot f_0}{f_{d2}\cdot C}\left(x^{(2)} - x_i\right) + \frac{V_{y2}\cdot f_0}{f_{d2}\cdot C}\left(y^{(2)} - y_i\right) + \frac{V_{z2}\cdot f_0}{f_{d2}\cdot C}\left(z^{(2)} - z_i\right) - r^{(2)} \\ \cdots \quad \cdots \\ \frac{V_{xn}\cdot f_0}{f_{dn}\cdot C}\left(x^{(n)} - x_i\right) + \frac{V_{yn}\cdot f_0}{f_{dn}\cdot C}\left(y^{(n)} - y_i\right) + \frac{V_{zn}\cdot f_0}{f_{dn}\cdot C}\left(z^{(n)} - z_i\right) - r^{(n)} \end{bmatrix}$$

And $[x^{(j)}, y^{(j)}, z^{(j)}]^T$ (j = 1–n, n $\geq$ 3) is the 3D position coordinate of the of the jth low orbiting satellites or jth observation point, $r^{(j)}$ is the observation distance between the jth low-orbit satellite or LEO satellite jth observation point and the user ith iteration solution result,

$$r^{(j)} = \sqrt{\left(x^{(j)} - x_i\right)^2 + \left(y^{(j)} - y_i\right)^2 + \left(z^{(j)} - z_i\right)^2}$$

$[V_{xj}, V_{yj}, V_{zj}]^T$ is the three-dimensional velocity of the jth low-orbit satellite or the jth observation point at the current time, f_0 is the carrier frequency of the LEO satellite signal, f_{dj} is the observation Doppler frequency of the jth low orbiting satellite or jth observation point. When the 2-norm of Y is less than a fixed constant D_y, the iterative calculation is finished, and the result of positioning solution is obtained.

3.2 Determination of Initial Value of Large Area Grid Search Positioning Solution

Newton Iterative Location Algorithm has the characteristics of small computation, high positioning accuracy and strong adaptability to satellite visibility. However, if the initial position of the user is more than 100 km away from the real value, it may lead to the convergence of the iterative calculation results. In order to solve this problem, a large area grid search method can be used to realize coarse positioning of the user on the order of 100 km. Then, the rough positioning result is taken as the initial value of the Newton iterative positioning solution, so as to realize high precision robustness in large area Positioning solution.

The initial value of the mesh search and positioning solution is determined as follows: Firstly, the large area is divided into lattice according to the latitude and longitude, then the latitude and longitude coordinates of each lattice point are

converted to ECEF coordinates, and then under the same condition of the satellite observation, the residual values in the Newton iterative localization solution are calculated:

$$Y = \begin{bmatrix} \frac{V_{x1}\cdot f_0}{f_{d1}\cdot C}(x^{(1)} - x_k) + \frac{V_{y1}\cdot f_0}{f_{d1}\cdot C}(y^{(1)} - y_k) + \frac{V_{z1}\cdot f_0}{f_{d1}\cdot C}(z^{(1)} - z_k) - r^{(1)} \\ \frac{V_{x2}\cdot f_0}{f_{d2}\cdot C}(x^{(2)} - x_k) + \frac{V_{y2}\cdot f_0}{f_{d2}\cdot C}(y^{(2)} - y_k) + \frac{V_{z2}\cdot f_0}{f_{d2}\cdot C}(z^{(2)} - z_k) - r^{(2)} \\ \cdots \quad \cdots \\ \frac{V_{xn}\cdot f_0}{f_{dn}\cdot C}(x^{(n)} - x_k) + \frac{V_{yn}\cdot f_0}{f_{dn}\cdot C}(y^{(n)} - y_k) + \frac{V_{zn}\cdot f_0}{f_{dn}\cdot C}(z^{(n)} - z_k) - r^{(n)} \end{bmatrix}$$

In which $[x_k, y_k, z_k]^T$ is the ECEF coordinate of the kth grid point, The minimum value of 2-norm of Y is obtained, and the corresponding grid point coordinates can be used as the iterative initial value of the Newton least square Doppler localization solution.

3.3 Positioning Calculation Mode and Accuracy Improvement

There are two application modes of the Doppler localization method, which are single epoch locating mode and continuous observing and locating mode. Single-epoch locating mode refers to the real-time location of a user by single-epoch Doppler observation. Generally speaking, when the number of visible stars is 3 or more, the user can work in single-epoch positioning mode. Continuous observation positioning mode refers to the use of multi-epoch Doppler observation information to achieve the joint location of static users, generally when the number of visible stars is less than 3, then work in continuous observation positioning mode.

The above is a general case of the mode of division, there is a special case is that if the number of visible stars is 2, but also have user elevation observation information, the same can work in a single epoch positioning mode, that is, two low-orbit satellites and the user elevation observation information is used to realize the user single epoch 3D positioning. At this point the positioning solution method and the method described above is similar, but the calculation formula of G and Y is:

$$G = \begin{bmatrix} \frac{V_{x1}\cdot f_0}{f_{d1}\cdot C} + \frac{x_i - x^{(1)}}{r^{(1)}} & \frac{V_{y1}\cdot f_0}{f_{d1}\cdot C} + \frac{y_i - y^{(1)}}{r^{(1)}} & \frac{V_{z1}\cdot f_0}{f_{d1}\cdot C} + \frac{z_i - z^{(1)}}{r^{(1)}} \\ \frac{V_{x2}\cdot f_0}{f_{d2}\cdot C} + \frac{x_i - x^{(2)}}{r^{(2)}} & \frac{V_{y2}\cdot f_0}{f_{d2}\cdot C} + \frac{y_i - y^{(2)}}{r^{(2)}} & \frac{V_{z2}\cdot f_0}{f_{d2}\cdot C} + \frac{z_i - z^{(2)}}{r^{(2)}} \\ \frac{2x_i}{a_h^2} & \frac{2y_i}{a_h^2} & \frac{2z_i}{b_h^2} \end{bmatrix},$$

$$Y = \begin{bmatrix} \frac{V_{x1}\cdot f_0}{f_{d1}\cdot C}(x^{(1)} - x_i) + \frac{V_{y1}\cdot f_0}{f_{d1}\cdot C}(y^{(1)} - y_i) + \frac{V_{z1}\cdot f_0}{f_{d1}\cdot C}(z^{(1)} - z_i) - r^{(1)} \\ \frac{V_{x2}\cdot f_0}{f_{d2}\cdot C}(x^{(2)} - x_i) + \frac{V_{y2}\cdot f_0}{f_{d2}\cdot C}(y^{(2)} - y_i) + \frac{V_{z2}\cdot f_0}{f_{d2}\cdot C}(z^{(2)} - z_i) - r^{(2)} \\ 1 - (x_i/a_h)^2 - (y_i/a_h)^2 - (z_i/b_h)^2 \end{bmatrix}$$

In which $a_h = a + h$, $b_h = b + h$, a, b are the length of the earth's long and short axis, h is the user's altitude estimate (Fig. 4).

If the visible number of LEO satellites is 1, only the single-satellite Doppler information continuous observation method can be used to realize the positioning solution. The principle of single-satellite continuous observation is shown in Fig. 7. Because of the fast speed of LEO satellites, users can observe the time of about a few minutes, in the limited observation time, select 3 and above the Doppler information observation points to achieve the user continuous observation positioning solution. The selection of observation points should meet the principle of optimal DOP. In this time the error introduced by the user's local clock drift is small, and can be ignored.

Whether single-epoch locating mode or continuous observing and locating mode, if the user is in the static state, the positioning solution can be further improved by the multi-group smoothing method, and the formula of the calculation result $[x, y, z]^T$ is:

$$[x, y, z]^T = \frac{1}{N} \sum_{k=1}^{N} [x_k, y_k, z_k]^T$$

In which $[x, y, z]^T$ is the ECEF 3D positioning solution result coordinates, N is the number of smoothing spots of positioning results, $[x_k, y_k, z_k]^T$ is the kth positioning result.

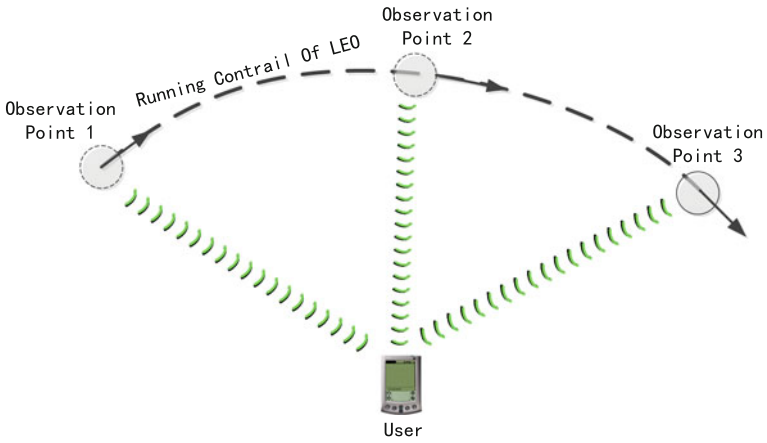


Fig. 4 Sketch map of doppler locating based single continuous observation

4 Simulation Verification of Localization Algorithm Performance

Positioning error simulation condition: The constellation scene is the basic layout unit of the global coverage system of the LEO satellite described in Sect. 1.2. The satellite visibility distribution in the coverage area is shown in Fig. 3. The error of the satellite orbit is 1 m, the Doppler measurement error is 1 Hz, satellite speed propagation error is 0.1 m/s, the user elevation variance is 5 m. The distribution of the Doppler localization errors for the three LEO satellites is shown in Fig. 5. Black in the figure indicates that the positioning error of less than 20 m of the region, about 95% of the region can be targeted, the gray that is greater than 20 m area, white represents non-three satellites positioning area.

The error distribution of the double-satellite Doppler and elevation joint location errors is shown in Fig. 6, where gray indicates that the positioning error is less than 20 m, about 85% of the total positioning area, white is greater than 20 m and black is the non-binary positioning area.

The position error distribution of single-satellite continuous observation for 5 min is shown in Fig. 7. Where gray indicates that the positioning error is less than 20 m, about 93% of the positioning area, white is greater than 20 m, black for non-single-satellite positioning area.

When the Doppler observation is used to locate the solution, there are several locations at each instant which are affected by the Doppler observation, resulting in a low positioning accuracy. However, as the satellite is moving in orbit, the position with poor positioning accuracy changes rapidly, No matter what kind of Doppler positioning method is used, as long as the continuous observation of several epoch, smoothing after removing the outliers can get more accurate positioning accuracy.

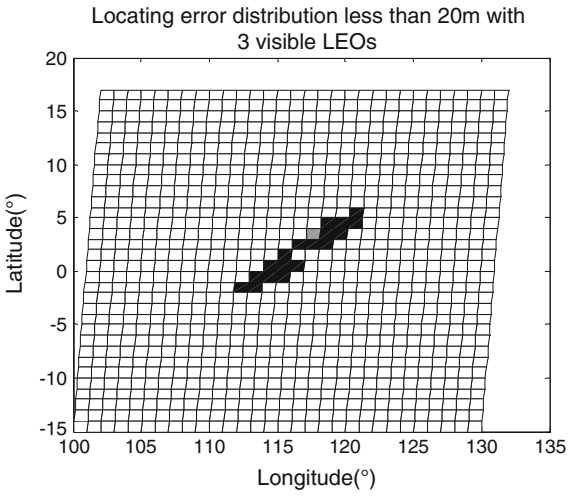


Fig. 5 Doppler locating error distribution with 3 visible LEOs

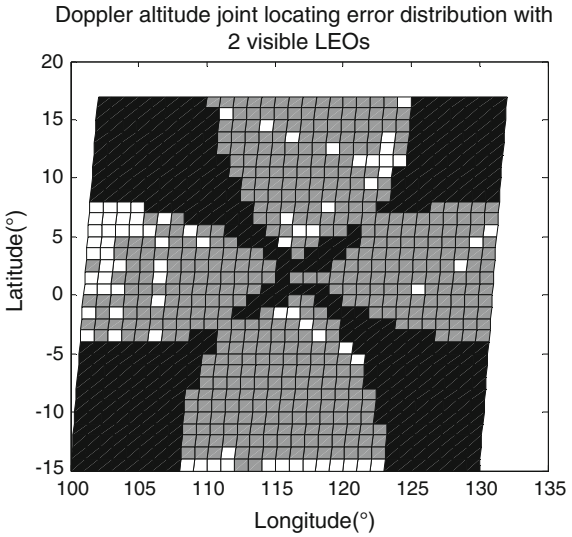


Fig. 6 Doppler altitude joint locating error distribution with 2 visible LEOs

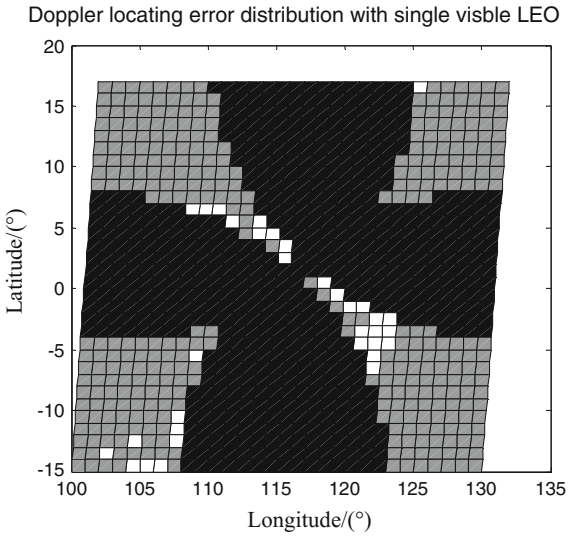


Fig. 7 Doppler locating error distribution with single visible LEO

Figure 8 shows the distribution of the multi-scene joint localization errors for the typical low-orbit constellation after three epoch-smoothing. The white region indicates the positioning error is less than 20 m and the black region is the positioning error greater than 20 m. The location error is better than 20 m region increased to 97.5%, if further increase the number of smoothing epoch can further enhance the positioning accuracy.

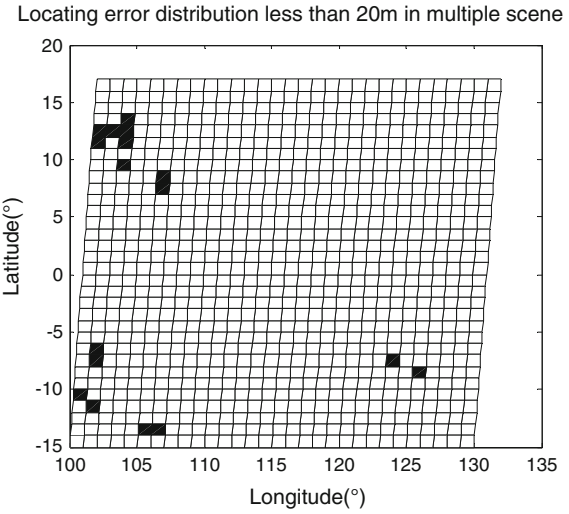


Fig. 8 Doppler locating error distribution in multiple scene

5 Summary

The LEO satellite backup navigation system has the advantages of wide coverage area and all-weather operation. The Doppler positioning system can avoid the complexity of system construction caused by high precision inter-satellite time synchronization [4], and is easy to integrate with communication and remote sensing satellites, which has good application feasibility.

In this paper, a set of Doppler localization methods, Based on Grid Search and Newton Least Squares, which can be applied to any number of visible satellites is proposed, which solves the problem of adaptive positioning algorithm in any low-orbit constellation. The simulation results show that in the typical scene, more than 97% coverage area can be satisfied, and the 3D positioning accuracy is better than 20 m, which can meet the requirements of navigation and positioning applications for most static and low dynamic users.

References

1. Asher MS, Stafford SJ, Bamberger RJ (2011) Radionavigation alternatives for US Army Ground Forces in GPS denied environments. In: International technical meeting of the Institute of Navigation, San Diego, CA, 24–26 January 2011, pp 508–532
2. Pan L (2013) Research on airborne passive location technologies based on doppler frequency diffrence. Southwest Jiaotong University Master Degree Thesis of the degree of Master, May 2013
3. Garrison TP, Ince M, Pizzicaroli J, Swan PA, Keyan P (2000) Iridium satellites constellation dynamics: systems engineering synthesis. Control Eng 1:29–38
4. Xu Y (2009) New techniques for single satellite passive localization based on doppler information. National University of defense technology thesis of the degree of master, Nov 2009

A New Set of Spreading Code Based on Odd Kasami Sequence

Ruxia Wang, Zhongliang Deng and Di Zhu

Abstract In the construction and application of satellite navigation system, the navigation signal structure is in the top design position. The technical basis of signal structure in modern satellite navigation system is spread spectrum, and spreading code with pseudorandom characteristic is one of the elements of spread spectrum signal. With the increasing of application requirements, the lack in quantity and self/cross-correlation performance of C/A code, m sequence has been gradually apparent. Spreading code with better related performance has been a new bright spot. Based on the study of the characteristics of signal structure in satellite navigation system and main evaluation parameters of spreading code, this paper proposes a new set of spreading code based on odd Kasami sequence. This paper analyzes the auto/cross-correlation of this code, and by comparing it with odd Kasami sequence and m sequence in the properties of auto-correlation, cross-correlation, balance performance and code quantity, verifies the excellent characteristics in correlation properties, balance performance and quantity.

Keywords Odd Kasami sequence · A new set of spreading code · Correlation properties · Balance performance · Code quantity

1 Introduction

Global Navigation Satellite System (GNSS) can provide all time, all weather and high accuracy positioning, navigation and timing services to users. GNSS becomes a critical infrastructure of modern country and military power. Signal structure is one of the critical challenges for the success of satellite navigation system. Signal structure with compatibility and interoperability, independent intelligence property, high performance is essential for the development of current navigation satellite

R. Wang (✉) · Z. Deng · D. Zhu
Beijing University of Posts and Telecommunications, Beijing, China
e-mail: wangruxia_go@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_40

system. The high accuracy time transfer and measurement based on the spread spectrum signal is the technical basis for the navigation, positioning and timing service of the satellite navigation system. And spreading code is one of the basic elements of navigation signal structure [1]. In the modern satellite navigation system, spreading code with good self/cross-correlation performance is needed to satisfy the requirements in signal acquisition speed, positioning accuracy and anti-jamming. Besides, it requires sufficient code quantity and balance performance to guarantee the system's user capacity and security. The m sequence and C/A code, which are commonly used in traditional systems, have shortcomings in code quantity and self/cross-correlation performance. The design of new spreading code has become the focus of study.

In this context, this paper chooses the design of spreading code as the study focus, and proposes a new set of spreading code, and studies its correlation, balance performance and quantity. The new spreading code, m sequence and odd Kasami sequence are simulated in the MATLAB environment, and the application value and foreground of the spreading code are comprehensively evaluated.

2 Evaluation Parameters of Spreading Code Performance

In order to compare the performance of different spreading codes, we can test and analyze the performance of the spreading code from the perspective of the distribution uniformity of the sequence and balance performance between different codes, and provide data support for spreading code applications. Correlation is a measure of the degree of signal similarity. The correlation of the spreading code has important effect on signal reception, detection and anti-interference. Auto-correlation and cross-correlation represent the distinguishing ability of signals, multipath signals and multiple access signals. The maximum correlation side lobe plays an important role in the research of spreading code correlation. The balance performance measure is an important part in the randomness test of spreading code. The balance performance is related to the carrier restrains of the system. The communication system will have larger carrier leakage and lower performance if the code is unbalanced. The number of code directly determines the user capacity of the system. With the continuous development of navigation system and the increasing demand of application, code quantity has become a focus of the signal structure research.

In this section, it mainly introduces the related definitions and concepts of correlation and balance performance of the spreading code.

2.1 Correlation of Spreading Code

In the application of spreading code, good sequence correlation is the key to selecting spreading codes, mainly including auto-correlation and cross-correlation.

Good auto-correlation performance can improve the anti-interference ability of the system and facilitate the signal reception and detection. Good cross-correlation can be used in multisite communication and helpful in improving the spectrum utilization.

For sequences of binary spreading codes $X\{x_1, x_2, \ldots x_L\}$, $Y\{y_1, y_2, \ldots y_L\}$ with period L, the definition of auto-correlation function and cross-correlation can be described as follows:

$$\rho_x = \frac{1}{L}\sum_{i=1}^{L} x_i x_{i+\tau} \tag{1}$$

$$\rho_{x,y} = \frac{1}{L}\sum_{i=1}^{L} x_i y_{i+\tau} \tag{2}$$

In above formulas, ρ_x is the auto-correlation value of the spreading sequence $X\{x_1, x_2, \ldots x_L\}$, and $\rho_{x,y}$ is the cross-correlation value of the spreading sequence $X\{x_1, x_2, \ldots x_L\}$ and $Y\{y_1, y_2, \ldots y_L\}$. L represents the length of the spreading sequence.

Among them, the cross-correlation, also known as side lobe, can directly reflect the correlation performance of the spreading sequence.

According to the definition of correlation, sequence with good auto-correlation should be: the main lobe normalized value is 1, and all the side lobe values are 0. And as for sequence with good cross-correlation, all the lobe values should be 0 [2]. But the correlation value of the spreading code in reality is not as good as in ideal, whose side lobes cannot be all zeros. So it's difficult to find out code family with good correlation characteristics. Therefore, this paper not only analyzes the auto-correlation and cross-correlation characteristics, but also compares and analyzes their maximum correlation side lobe power. The definition is:

$$S = 20\log_{10}(\rho'_{\max}) \tag{3}$$

In which, S is the maximum correlation side lobe power with the unit dB, and $\rho'_{\max}$ represents the maximum correlation side lobe coefficient.

2.2 Measurement of Balance Performance

The balance performance of the spreading code is one of the important indexes to judge the security and reliability performance of spreading spectrum system. In spread spectrum communication, information security largely depends on whether the generated sequence has good randomness. Sequence with better randomness is closer to ideal random sequence and has higher security. So it is very meaningful to

test the statistical characteristic of the spreading code. Good balance performance can prevent carrier leakage with high frequency, so as to ensure the security of communication system. The balance performance of the spreading code is determined by the statistical distribution of '0' and '1' [3].

The degree of balance performance E is defined as follows:

$$E = \frac{|P - Q|}{L} \tag{4}$$

Whereas L is the length of spreading code, and P and Q are respectively the total numbers of '1' and '0'. The smaller the E is, the better the balance performance is, and the better performance of the spreading sequence has.

3 The New Set of Spreading Code Based on Odd Kasami Sequence

3.1 The Generation of Odd Kasami Sequence

Kasami sequence set is one of the important types of binary sets which has small cross-correlation. It consists of large ensemble Kasami sequence and small ensemble Kasami sequence. The Kasami sequence used in this paper are based on small ensemble ones.

Kasami sequence is constructed with PN sequence. It is a kind of binary sequence with the period of $2^n - 1$, and n is the number of linear shift register. The so-called odd Kasami sequence represents the Kasami sequence with positive odd number n.

The generation steps of odd Kasami sequence are as follows:

(1) Input the number n of linear shift register and initial state, construct pseudo-random noise sequence m with the length of $N1 = 2^n - 1$.
(2) Construct PN sequence named PN_1 with length of $N_2 = 2^n$. Whereas the m sequence is assigned to the first $2^n - 1$ bits of the sequence and the second 2^n bit is generated by the random function with the random ranging from 0 to 1.
(3) Decimate PN_1 by $2^{(n+1)/2}$ bit and obtain a new sequence $PN_1^{'}$ with the length of $2^{(n-1)/2}$.
(4) Perform all $2^{(n+1)/2} - 1$ cyclic shifts of the bits of $PN_1^{'}$ to get a new sequence PN_2 with the length of N_2.
(5) Perform modulo-2 addition between bits of PN_1 and those of PN_2 to obtain the final odd Kasami sequence with length of N_2 [4].

3.2 The Generation of the New Set of Spreading Code Based on Odd Kasami Sequence

The new set of spreading code proposed in this paper is based on odd Kasami sequence. The specific generation flow chart is as follows [5] (Fig. 1):

3.3 Correlation Performance of the New Spreading Code

This paper simulates and analyzes the correlation performance of the new set of spreading code based on odd Kasami sequence with different sequence lengths in

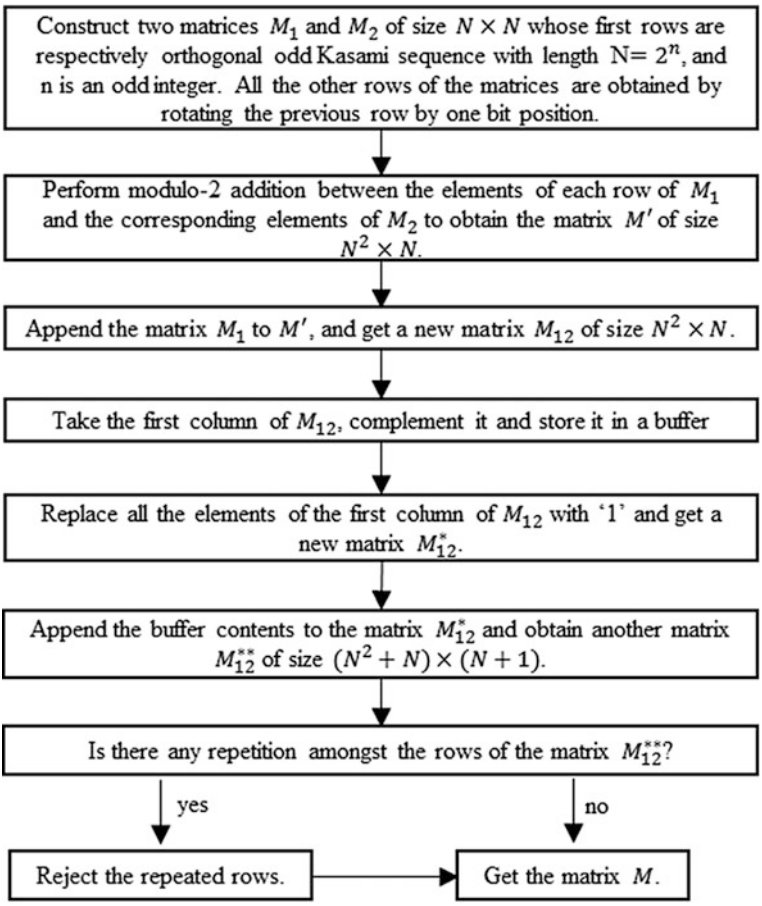


Fig. 1 Flow chart for the new spreading code based on odd Kasami sequence

MATLAB environment. Whereas, the horizontal axis τ represents the time delay, and the vertical axis represents the correlation normalized value.

3.3.1 Auto-correlation Performance of the New Spreading Code

Figures 2, 3, 4 and 5 show the auto-correlation characteristics of the new spreading code. It can be seen from the figures that the new spreading code proposed in this paper has excellent auto-correlation characteristics. Especially with the increase of the length, the auto-correlation function of the sequence approximates the δ function.

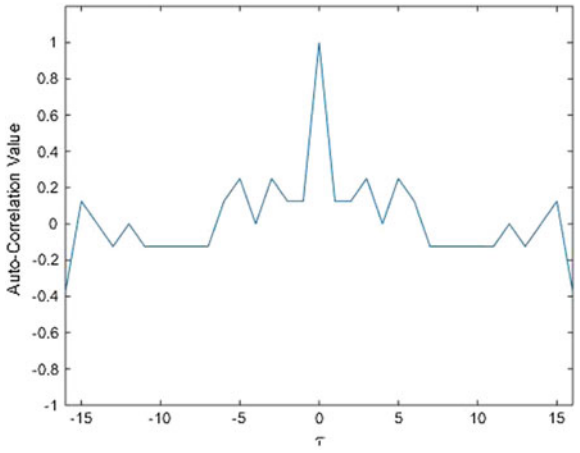


Fig. 2 Auto-correlation simulation of new spreading code with n = 5

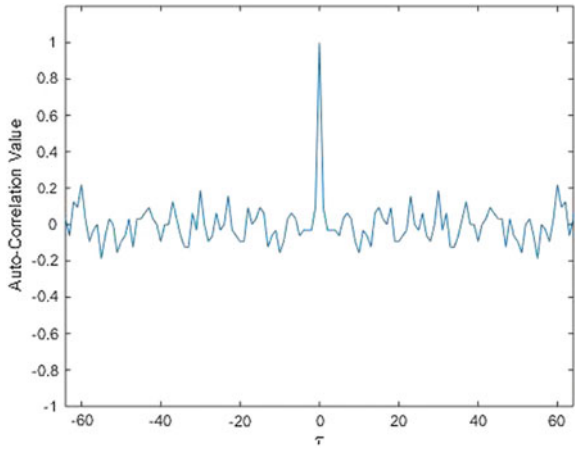


Fig. 3 Auto-correlation simulation of new spreading code with n = 7

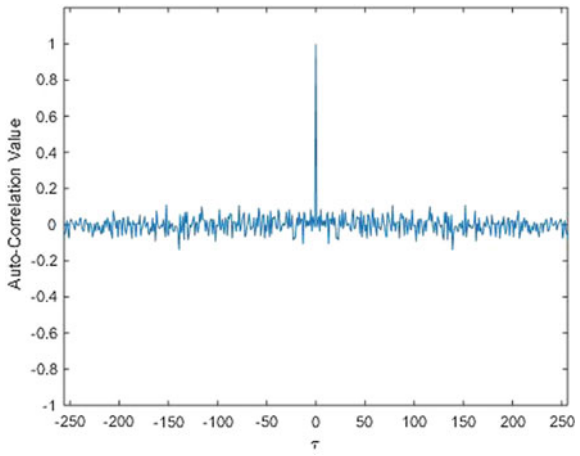


Fig. 4 Auto-correlation simulation of new spreading code with n = 9

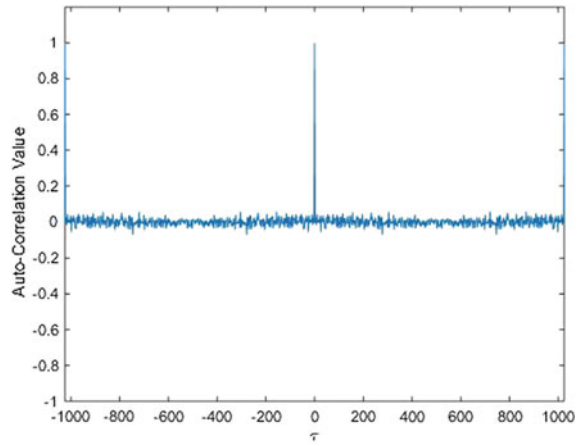


Fig. 5 Auto-correlation simulation of new spreading code with n = 11

3.3.2 Cross-correlation Performance of the New Spreading Code

Figures 6, 7, 8 and 9 show the cross-correlation characteristics of the new spreading code. It can be seen from the figures that the new spreading code proposed in this paper has excellent cross-correlation performance. And with the increasing of length, the cross-correlation function of the sequence tends to be 0.

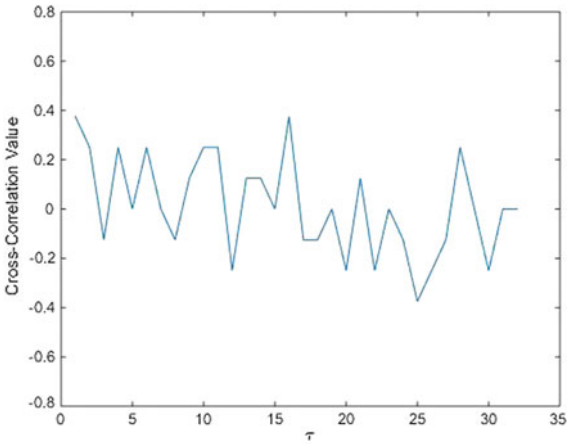


Fig. 6 Cross-correlation simulation of new spreading code with n = 5

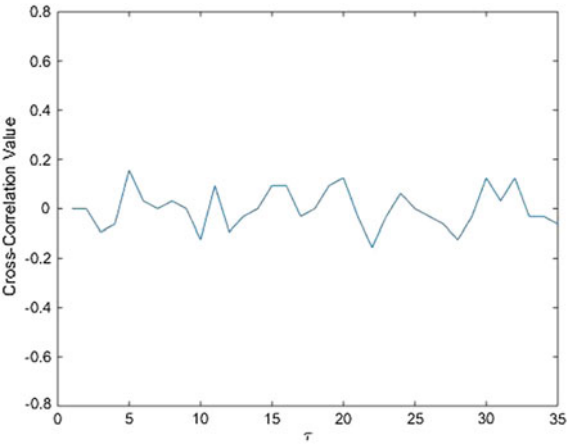


Fig. 7 Cross-correlation simulation of new spreading code with n = 7

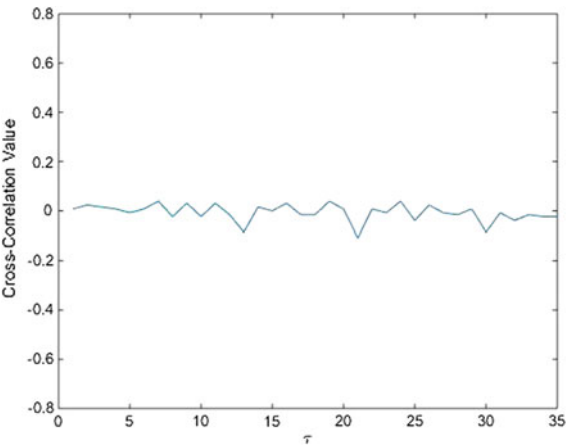


Fig. 8 Cross-correlation simulation of new spreading code with n = 9

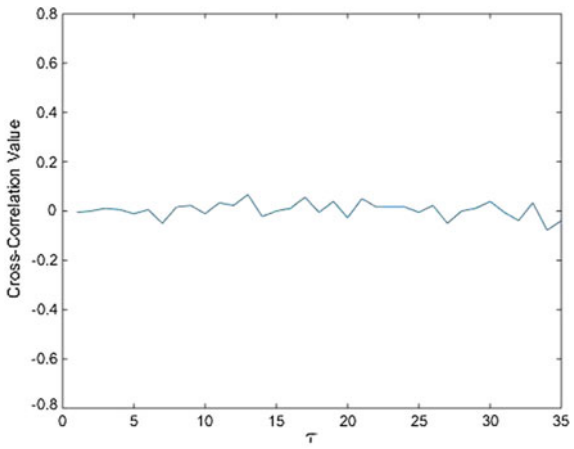


Fig. 9 Cross-correlation simulation of new spreading code with n = 11

Table 1 Maximum in auto-correlation side lobe of spreading codes

Series	New spreading code	Odd Kasami sequence	m sequence
n = 5	−11.2854	−9.8463	−12.0412
n = 7	−17.6028	−13.2018	−18.0618
n = 9	−20.8049	−19.2216	−24.0824
n = 11	−25.4705	−23.0593	−28.6103

4 Performance Comparison of Different Spreading Codes

4.1 Measurement of Correlation Properties

4.1.1 Measurement of Auto-correlation Property

In the MATLAB simulation environment, different spreading codes are generated, including the new set of spreading code proposed in this paper, odd Kasami sequence and m sequence. And by comparing the maximum auto-correlation side lobe of these sequences in different series, results are obtained as Table 1:

By analyzing the data in Table 1, it can be seen that these three kinds of spreading codes have considerable auto-correlation characteristics. With the increasing of length, the auto-correlation side lobes decrease. In addition, it also can be seen that the auto-correlation performance of the new spreading code proposed in this paper is better than odd Kasami sequence, and a little worse than m sequence. But all in all, the new spreading code can reduce the false alarm probability and improve the speed and accuracy in acquisition and tracking of the system.

Table 2 Maximum in cross-correlation side lobe of spreading codes

Series	New spreading code	Odd Kasami sequence	m sequence
n = 5	−8.0314	−8.0986	−8.5194
n = 7	−14.0943	−14.0037	−13.201
n = 9	−17.0412	−18.1973	−16.1236
n = 11	−23.3040	−24.0951	−21.72040

Table 3 Balance performance of spreading codes

Series	New spreading code	Odd Kasami sequence	m sequence
n = 5	0.1516	0.1876	0.1626
n = 7	0.1078	0.1562	0.1136
n = 9	0.0963	0.0904	0.1007
n = 11	0.0205	0.0328	0.0510

4.1.2 Measurement of Cross-correlation Property

In the MATLAB simulation environment, different spreading codes are generated, including the new set of spreading code proposed in this paper, odd Kasami sequence and m sequence. And by comparing the maximum cross-correlation side lobe of these sequences in different series, results are obtained as Table 2:

By analyzing the data in Table 2, it can be seen that these three kinds of spreading codes have considerable cross-correlation characteristics. With the increase of the length, the cross-correlation side lobes decrease. In addition, it also can be seen that the cross-correlation performance of the new spreading code proposed in this paper is close to odd Kasami sequence, and better than m sequence. But all in all, the new spreading code can greatly reduce the multiple access interference and improve the system performance.

4.2 Measurement of Balance Performance

In the MATLAB simulation environment, different spreading codes are generated, including the new set of spreading code proposed in this paper, odd Kasami sequence and m sequence. According to the definition of measurement of balance performance, the degree of balance is calculated with different series. It is shown in Table 3:

By analyzing the data in Table 3, it can be seen that these three kinds of spreading codes all have good balance performance. With the increase of the length, the degree of balance performance E becomes closer to the value 0. In addition, it also can be seen that the balance of performance of the new spreading code

Table 4 Code quantity of spreading codes

Series	New spreading code	Odd Kasami sequence	m sequence
n = 5	916	24	6
n = 7	5432	144	18
n = 9	24,224	768	48
n = 11	158,244	5632	176

proposed in this paper is better than odd Kasami sequence and m sequence. What's more, the higher the series is, the more obvious the superiority of the proposed code in balance performance becomes. This indicates that the new spreading code has less carrier leakage, which is beneficial to improve the system security.

4.3 Analysis of Code Quantity

In the MATLAB simulation environment, different spreading codes are generated, including the new set of spreading code proposed in this paper, odd Kasami sequence and m sequence, and code quantity of these spreading code are calculated in different series. It is shown in Table 4:

By analyzing the data in Table 4, it can be seen that among these three kinds of spreading codes, the proposed new spreading code has absolute advantage in quantity. Moreover, the higher the number of the series, the faster the number of the code grows. It shows that the new spreading code can provide sufficient code quantity for the navigation system, and is more capable for the current development requirement of the system user capacity.

5 Conclusion

In this paper, a new set of spreading code based on odd Kasami sequence is proposed. Its auto-correlation and cross-correlation performance is simulated and compared with odd Kasami sequence and m sequence. From the final results, it can be seen that the distribution of the new spreading sequence is more uniform, and the auto-correlation performance is greatly improved while maintaining the excellent cross-correlation performance inherited form odd Kasami sequence. Furthermore, the number of the code is greatly increased. But, it can also be seen that the construction process of this set of spreading code is a little complicated. So how to reduce the complexity of operation and how to optimize the spreading code will become a further study issue. All in all, this new set of spreading code based on odd Kasami sequence has excellent properties, and is more capable for the requirements of speed and accuracy of acquisition and tracking in modern navigation system, communication security and user capacity.

Acknowledgements This work is supported by the National key Research and Development Program (NO. 2016YFB0502001), National Natural Science Foundation of China (NO. 61372110), National High technology Research and Development Program of China (NO. 2014AA123103).

References

1. Zhu J (2015) Study on spreading codes construction, selection and enhanced reception for satellite navigation signal. A thesis for doctor, Beijing Institute of Technology 2015.6
2. Wang H (2015) Structure and performance analysis of several new spread spectrum codes. A thesis for the degree of master, Xi'an University of Posts and Telecommunications, 2015.6
3. Zuo H (2012) Optimization design of GNSS pseudo-random code. A thesis for the degree of master, Huazhong University of Science and Technology 2012.5
4. Chandra A, Chattopadhyay S (2009) Small set orthogonal Kasami codes for CDMA system. In: International conference on computers and devices for communication, 14–16 Dec 2009
5. Banerjee A, Saif KM, Chattopadhyay S (2015) A new set of spreading codes for synchronous CDMA system. In: International conference on computational intelligence and communication networks, 12–14 Dec 2015

Data Transfer Problem in Navigation Satellite Network Based on Agility Link

Zhenwei Hou, Xianqing Yi, Yue Zhao and Yaohong Zhang

Abstract Autonomous navigation supported by inter-satellite ranging and communication with inter-satellite links (ISL) is the most important direction of development of satellite navigation system. The agility link can achieve more information from other satellites by switching the direction of phased array antenna flexibly. However, the complexity and efficiency of communication are affected at the same time. In this paper, we describe the data transfer problem when the data is split into several datum units whose hops are limited. We investigate the computational complexity of the problem firstly and it is proved to be an NP-complete problem. Then we propose a linear programming model for solving the problem. At last, we show that the problem is polynomial solvable when the number of recipient nodes or the number of datum units is equal to 1. The research in this paper is valuable for solving the data transfer problem in the navigation satellite network based on agility link.

Keywords Agility link · Data transfer problem · Linear programming · NP-complete

1 Introduction

ISLs technology is one of the most important technologies for satellite navigation system to improve the positioning accuracy and to achieve autonomous navigation. Most of the world's satellite navigation systems are actively exploring the technology of ISLs. ISLs can shorten the ephemeris update cycle to improve the accuracy of positioning. By inter-satellite ranging, ISLs can achieve higher precision orbit determination for autonomous navigation. As well, ISLs can forward measurement and control signals to indirectly monitor and control the navigation

Z. Hou (✉) · X. Yi · Y. Zhao · Y. Zhang
Science and Technology on Information Systems Engineering Laboratory,
National University of Defense Technology, Changsha, China
e-mail: houzhenwei@nudt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_41

constellation. So the number of ground stations and maintenance costs of the system can be significantly reduced.

Agility link is a type of ISLs. It is generally based on the phased array antenna equipped on the satellite to communicate and range between satellites. The links switch and reconstruct frequently and its pointer is changing with time. The time is divided into discrete and equal time slots, denoted Δt. The length of time might be so short that the data couldn't be transferred completely. So we subdivide data into several datum units which are sent in any order to recipient satellites. A satellite can only transmit one datum unit in a time slot.

ISLs mainly have two functions in the navigation system: inter-satellite communication and ranging. Agility and quick direction adjustment characteristic of agility link is easy to get a certain number of observations of pseudo range measurements among satellites. We try to obtain as much as possible the observations in a short time to improve the accuracy of the system. So in a cycle period of link assignment, any two satellites establish the link only once. Because of the limitation of energy and computation ability of the satellite, a satellite generally carries only one link in a time slot. So the network can be considered as an intermittently connected network of several satellites from a data transportation point of view. However, with the increase of the forwarding number, errors in some navigation information will also be accumulated. So transmission hops of this kind of information must be limited. In this paper, it is assumed that reliable predictions of the sequence of the links between satellites can be made. During a time slot, satellites connected by the links transfer one stored datum unit to each other and the source satellites initially store some datum units. Via several time slots, the datum units are transferred from one satellite to another satellite until their destination is reached. This routing scheme is based on a store-carry-forward paradigm. So the navigation satellite network based on agility link is a typical delay tolerant network (DTN). Figure 1 shows the topology change of agility links.

As far as we known there is little research on the data transfer problem of navigation satellite network based on agility link currently. Reference [1] proposes and formally analyses least cost journeys in a class of dynamic networks, where the changes of the topology can be predicted in advance. The means of 'cost' here are hop count (shortest journeys), arrival date (foremost journeys), and time span (fastest journeys). In [2] Ronan Bocquillon studied the data transfer problem and dissemination problem in the systems of systems with the assumption that reliable predictions of the contacts sequence can be made and an information item is split into several datum units. The data transfer problem consists in searching for a valid transfer plan allowing the datum units to be transferred from their source systems to the recipient systems. By mapping an evolving graph to an adjacency matrix of an equivalent static graph, Jiahao proved that their generalization of the BFS algorithm correctly accounts for paths that traverse both space and time in [3]. Minsu Huang studied the topology control problem in a predictable DTN using space-time graph [4]. They tried to build a sparse structure from the original space-time graph such that the network is still connected over time and the total cost the structure is minimized. However, the object of these researches is the general satellite network.

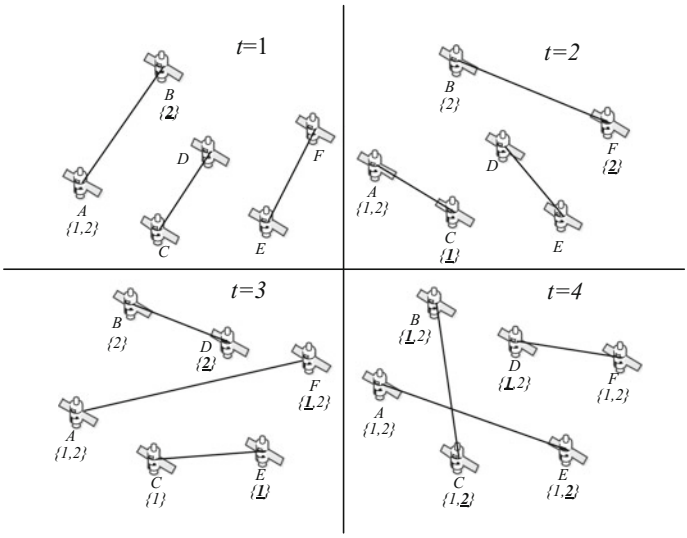


Fig. 1 Topology change of agility link and its data transfer progress. *t* represents the time slot number. In every time slot, the pairing satellites communicate with each other and the datum units are transferred to each other. There are two data units {1, 2} in the network. The *underlined* data units are received in the current time slot

Some of them optimized the data transfer problem considering only one datum unit, while the others were in the case of serial contacts between nodes. All of these researches cannot be directly applied to the data transfer problem in this paper.

In this paper, we fully consider the characteristic of agility link which build the links concurrently. We describe the data transfer problem when the data is split into several datum units whose hops are limited. And then we investigate the computational complexity of the problem. We also propose a linear programming model for solving the problem. At last, we discuss two specific polynomial cases. The remainder of the paper is organized as follows. Section 2 formally describes the data transfer problem. Section 3 investigates the computational complexity of the problem. In Sect. 4 we build the linear programming model to solve the problem and show that two particular cased of the problem are polynomial solvable. We conclude the whole work of this paper in Sect. 5.

2 Description of the Problem

In this paper, buffers are assumed to have an infinite capacity and network failures are disregarded. Reliable predictions of the sequence of the links between satellites can be made. Each satellite establishes at most one two-way link in a time slot and between any two satellites up to one link is established in a chain period.

The information data to be transferred is divided into multiple datum units which may have different source nodes and destination nodes. As well, different datum units may have different maximum survival hops.

The data transfer problem of the navigation satellite network based on agility link is to study how the datum unit can reach the destination node under the premise of given chain relation. How can the transmission be done in the least amount of time when all information can reach the destination? Figure 1 depicts the transmission process of datum units on agility links. Here 'how' mainly refers to 'when to transfer' and 'which to transfer'. For the purpose of further accurate description, the following variables are defined.

We consider a set $V = \{1, \ldots, n\}$ of n nodes (representing the satellites). Set $T = \{1, \ldots, m\}$ represents m time slots. Set $Q = \{1, \ldots, k\}$ represents k datum units to be transferred. Each datum unit $q \in Q$ has two corresponding subsets $S(q) \subseteq V$ indicating the nodes initially store datum q unit and $D(q) \subseteq V$ indicating the destination nodes of datum unit q. So the number of the destination nodes corresponding unit q can be indicated as $r = \left| \cup_{q=1}^{k} D(q) \right|$.

According to the destination nodes of each datum units, we can know the datum units to be received of every node, denoted by $R(i)$. $R(i) = \{q | i \in D(q)\}$. $H(q)$ denotes the maximum survival hops of datum unit $q \in Q$. $O_{i,t} \subseteq Q$ indicates the datum units node i possesses in time slot t. $p(i, t) = j$ represents a link is established between node i and j in time slot t. A transfer plan is denoted by Φ which describes the information transfer progress. In fact it is a mapping from the chain relations between nodes to the datum units. So it can be defined as a function:

$$\Phi: \{p(i,t) | i \in N, t \in T\} \rightarrow \{\emptyset\} \cup Q.$$

Based on the above variables, we will study the data transfer problem of navigation satellite network based on agility link in this paper, which is called *Data Transfer Problem with Limited Hops* (DTLH). The DTLH problem can be described as follows:

INSTANCE: A set $V = \{1, \ldots, n\}$ of n nodes; The links between all the nodes during m time slots $P_{n \times m} = \{p(i, t) | i = 1, \ldots, n; t = 1, \ldots, m\}$; A set $Q = \{1, 2, \ldots, k\}$ of k datum units; For each $q \in Q$, $S(q)$, $D(q)$ and $H(q)$ are also known.

QUESTION: Is there a valid transfer plan Φ such that all the datum units reach the destination nodes within m time slots, that is $\forall i \in V$, $O_{i,m} = R(i)$? If the answer is YES, then we will study how to minimum the cost time? That is to search for the minimum t such that $\forall i$, $O_{i,t} = R(i)$.

In order to describe the spatio-temporal relationship of the inter-satellite link in the navigation satellite network based on agility link, we introduce the space-time graph [5]. The space-time graph depicts the relationship between nodes in different time slots. In the space-time graph, the same node in different time slots is seen as different nodes, and connected with edges, which means that information can be stored in the node. This type of edges is called the temporal edge, such as the edge (v_1^0, v_1^1) in Fig. 2. There is also another type of edges which connect different nodes

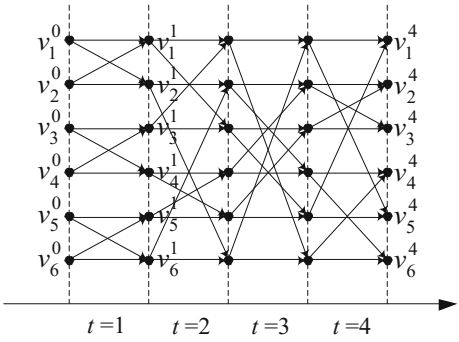


Fig. 2 The agility link depicted by space-time graph. This is the space-time graph corresponding to the agility satellite network in Fig. 1

in the same time. We call them the space edge. Space edge shows the chain relations between nodes in a time slot such as the edge (v_1^0, v_2^1) in Fig. 2.

3 The NP-Complete of the Problem

In this section we demonstrate that the DTLH problem is an NPC problem. Let us consider the decision version of the data transfer problem, i.e. is there a valid Φ such that $\forall i$, $O_{i,m} = R(i)$? DTLH is obviously in NP, since it can be decided in polynomial time whether a given transfer plan Φ is valid and such that $\forall i$, $O_{i,m} = R(i)$. The following shows that the DTLH problem is an NPC problem.

To prove these results, we need to reduce a problem known as being strongly NPC to the study cases. Ronan Bocquilllon in the literature [2] proved Data Transfer problem (DT) is the NPC problem. The DT problem is defined as follows:

INSTANCE: A set $V = \{1, \ldots, n\}$ of n nodes; A subset $R \subseteq V$ of recipient nodes. A sequence of m ordered pairs $(i, j) \in V \times V$ with $i \neq j$. A set $Q = \{1, \ldots, k\}$ of k datum units. For each node $i \in V$, a subset $O_i \subseteq Q$ of units are initially stored by i.

QUSETION: Is there a valid transfer plan $\Phi : \{1, \ldots, m\} \rightarrow \{\emptyset\} \cup Q$ such that $\forall i \in R$, $O_i = Q$ when $t = m$?

Bocquilllon proved that the DT problem is NPC when $k \geq 2$ by reducing the 3SAT problem to the DT problem. DT problem does not restrict the nodes in a chain-building cycle cannot repeat the link. However, from the proof process it can be drawn that not repeating the link clearly does not affect the process and the results of its proof. Therefore, the proof of this paper is based on this issue.

The following describes the progress of converting any instance of a DT problem into a DTLH problem. In fact, if two adjacent contacts concern different nodes, they can occur at the same time slot. Therefore we can create the links of each time slot by the following rules: Following the sequence of contacts, the adjacent contacts concerning different nodes are arranged in the same time slot. If a contact concerning one same node occurs, then it will be arranged in the next time

slot. According to this rule, it is not consistent with the characteristics of agile satellite network that some nodes in some time slots do not have paring nodes. Therefore, we add links for these nodes. Ensure that the link is without duplication. As well, add each new link a new datum units with the corresponding source node, destination node and maximum survival hops. The source and destination nodes are set to be the two pared nodes in the newly added link. The maximum survival hops are set to be 1. This effectively restricts the newly added information to be transmitted to the destination node only through the newly added link. It is worth noting that the new links of each time slot should not repeat, so the number of timeslots within a cycle is the upper limit, the proof of this paper assumes that the number of time slots within this limit. We define the following variables:

Current time slot number *Slot*; The datum units set *Data*; The new added datum units *X*; Constructed Graph *G* is composed of the set of contacts (chain subsequences) in all time slots, e.g. $G(t)$ denotes the contacts in time slot *t*. $G(t).V$ denotes the set of nodes which have been allocated links in time slot *t*. *S* denotes the subset of $G(t).V$ which are source nodes while *R* denotes the subset of that which are the receiver nodes. The process is shown in Table 1.

Table 1 Converting progresses of the instances between DT and DTLH problem

Initialize: $Slot=1$; $Data = Q$,$\forall q \in Q$, $h(q)=m$;
$X=\{\emptyset\}$;$G=\{\emptyset\}$; $S=\{\emptyset\}$; $R=\{\emptyset\}$; $x=1$;

STEP 1:
for $t=1 : m$
  for every $\sigma_t=(s_t,r_t)$
    if $s_t \in S$ or $r_t \in R$ or $\exists s_i \in S\ s_i = s_t\ r_i \neq r_t$
      or $\exists r_i \in R\ r_i = r_t\ s_i \neq s_t$
    $Slot = Slot + 1$;
    $S=\{\emptyset\}$ $R=\{\emptyset\}$
    else
    $G(Slot)=G(Slot)+ \sigma_t$
    endif
end

STEP 2:
while $Slot \neq 0$
 $\forall\ i,j \notin G(Slot).V$ and $(i,j) \notin G$
      $G(Slot)=G(Slot)+ (i,j)$
      $Data=Data+\{x\}$
      $X=X+\{x\}$
      $S(x)=i$; $D(x)=j$;
      $h(x)=1$;
      $x=x+1$;
  $Slot = Slot - 1$;
end

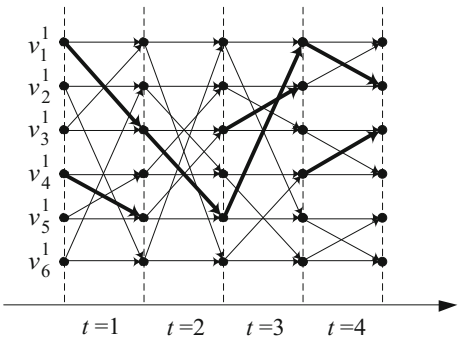


Fig. 3 The space-time graph depicts an instance of DT problem. The instance of the DT problem has 6 contacts σ = ((1, 3), (4, 5), (3, 5), (3, 2), (5, 1), (1, 2))

The above described progress is apparently performed in polynomial time. For example, an instance of the DT problem has 6 contacts σ = ((1, 3), (4, 5), (3, 5), (3, 2), (5, 1), (1, 2)). We can construct an instance of the DTLH problem as shown in Fig. 3. The following shows that it is one-to-one correspondence between the instance of the DTLH problem and DT problem.

If there is a valid transfer plan of a DT instance, this plan can still transmit the original datum units Q in the DTLH instance since the constructed space-time graph doesn't affect the transmission process of the contacts in the DT instance. For the newly added data unit set X, it is known that at least one path can transfer them to the destination nodes within one hop from the construction process of DTLH space-time graph. Therefore, if there is a valid transfer plan of a DT instance, so is the corresponding DTLH instance.

Conversely, if there is a valid transfer plan of a DTLH instance, consider the newly added datum unit set X and the data set Q of the original problem as well. Since there is no repetition of the links and the newly added datum unit $x \in X$, $h(x) = 1$, the datum x can only be transmitted by the newly added links. Furthermore, The number of newly added data units is equal to the number of newly added links, so all of the newly added links are used to transmit the newly added data units. As a result, the original data set Q can only reach the destination nodes through the original contacts of the DT instance, which proving that the DT instance must exist a feasible transfer plan.

4 Model for the Problem

In this section, we first establish a linear programming model based on the constraint and optimized targets of agile satellite networks. Then we discuss two special cases when the number of data units is 1 and the number of receiving nodes is 1.

4.1 Linear Programming Model

Firstly we define following decision variables.

$x_{i,q,t} = 1$ denotes that satellite i transmits datum q to its pairing satellite at time slot t. Otherwise, $x_{i,q,t} = 0$. $y_{i,q,t} = 1$ denotes that satellite i has possess the datum q at time slot t. Otherwise, $y_{i,q,t} = 0$. In the initial state, as constraint (9) shows that for the source node of datum q, $y_{i,q,1} = 1$, otherwise $y_{i,q,1} = 0$. $h_{i,q,t}$ denotes that the hops of satellite i receiving the datum q by the timeslot t. If node i haven't received q, $h_{i,q,t} = 0$. In the initial state, for all the nodes and datums $h_{i,q,1} = 0$, as constraint (10) shows.

According to the characteristics of the agile satellite network and the requirements of the DTLH problem, we get the following constraints:

$$\forall i \in V, \forall t \in T: \sum_{q=1}^{k} x_{i,q,t} \leq 1 \tag{1}$$

$$\begin{aligned} &\forall i \in V, \forall q \in Q, \forall t \in T: \\ &x_{i,q,t} \leq y_{i,q,t} \end{aligned} \tag{2}$$

$$y_{i,q,t} + x_{p_{i,t},q,t} = y_{i,q,t+1} \tag{3}$$

$$h_{i,q,t+1} = h_{i,q,t} + x_{p_{i,t},q,t}(h_{p_{i,t},q,t} - h_{i,q,t}) + x_{p_{i,t},q,t} \tag{4}$$

$$h_{i,q,m+1} \leq H_q \tag{5}$$

Equation (1) states that at most one datum unit can be transferred during each time slot for each node. Equation (2) states that only one node possesses a datum unit can this node transmit this datum unit. Equation (3) contains two meanings: First, once a node receives a datum unit in a time slot, this node possesses this datum forever, i.e. if $x_{p_{i,t},q,t} = 1$, then $y_{i,q,t+1} = 1$. The other meaning is that if a node receives a datum unit in a time slot, this node will not accept this datum unit again in the following time slot. This strategy can guarantee the minimum delay of data transmission.

Equation (4) is the key constraint to solve the problem. It is composed of two cases. One case is that node i doesn't receive datum unit k in time slot t, then the hop number of node i about datum unit k in time slot $t + 1$ is not changed, i.e. if $x_{p_{i,t},q,t} = 0$, then $h_{i,q,t+1} = h_{i,q,t}$. Similarly, another case is that if $x_{p_{i,t},q,t} = 1$ then $h_{i,q,t+1} = h_{p_{i,t},q,t} + 1$. Equation (5) indicates that the datum is to be transmitted to the destination node within the maximum survival hops.

It can be noted that Eq. (4) is the nonlinear constraint. Due to the complexity of the nonlinear programming and the difficulty of getting the optimal solution, we transformed the constraint (4) into the constraints (6), (7) and (8) by introducing a

large value M and the variables $m_{i,q,t}$. The nonlinear constraint is transformed into a linear constraint which guarantees the reliability of the results.

$\forall i \in V, \forall q \in Q, \forall t \in T$:

$$h_{i,q,t+1} = h_{i,q,t} + x_{p_{i,t},q,t} + m_{i,q,t} \tag{6}$$

$$0 \le m_{i,q,t} \le M \cdot x_{p_{i,t},q,t} \tag{7}$$

$$h_{i,q,t+1} - h_{i,q,t} - M \cdot (1 - x_{p_{i,t},q,t}) \le m_{i,q,t} \le h_{i,q,t+1} - h_{i,q,t} + M \cdot (1 - x_{p_{i,t},q,t}) \tag{8}$$

Considering the initial state of the network (t = 1) we get the following constraints:

$$\forall i \in V, \forall q \in Q, y_{i,q,1} = \begin{cases} 1, & i \in S(q) \\ 0, & else \end{cases} \tag{9}$$

$$\forall i \in V, \forall q \in Q \; h_{i,q,1} = 0 \tag{10}$$

The objective function of the problem is:

$$\mathbf{min}\, \lambda \tag{11}$$

where

$$\lambda = \max_{i \in R} \max_{q \in D} \lambda_{i,q} \tag{12}$$

$$\lambda_{i,q} = m - \sum_{t=1}^{m} y_{i,q,t} \tag{13}$$

Equation (13) indicates the time slots used by datum unit q to reach node i. Equation (12) indicates the latest time for all the datum units to reach the destination nodes. Our optimization goal is to make this latest time possible advance, i.e. **min λ**.

4.2 Polynomial Cases

4.2.1 The Case *k* = 1

When the number of the datum units k = 1, it is easy to think of the method of flooding. As long as the datum is within the maximum survival hops, it is allowed to spread and let the units remaining much more hops cover the less ones. In [6] we propose the method of finding the shortest journey and the earliest journey by the

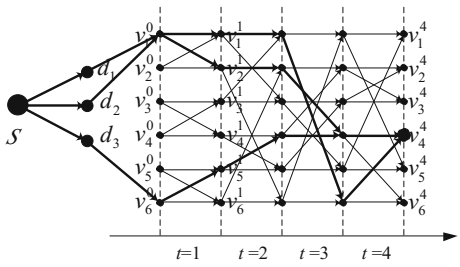


Fig. 4 The DTLH problem is transformed into the problem of minimum cost maximum flow when the number of receiving node is 1. In the figure, there are three datum units {1, 2, 3}, and the source node of datum unit 1, 2 is v_1, the source node of data unit 3 is v_6 and the destination node is v_4

greedy algorithm when the number of data units is 1. The only difference with the DTLH is that in the progress of flooding we must consider the maximum survival hop which is only a slight change on the basis of literature [6]. So we do not repeat it in this paper.

4.2.2 The Case $r = 1$

When the number of receiving nodes $r = 1$ and the maximum survival hops of all the datum units are the same h, the DTLH problem can be transformed into the problem of minimum cost maximum flow. The conversion process is as follows:

On the basis of the space-time graph, add a virtual source node S and k virtual nodes $d_1, \ldots, d_k$, representing the k datum units. k directed edges $(S, d_1), \ldots, (S, d_k)$ are also added to the space-time graph. Connect the datum nodes and the corresponding source nodes in the space-time graph with edges whose capacity is 1 and the cost is 0, i.e. for $\forall q \in Q, i \in S(q)$, add edge (d_q, i) with capacity $c(d_q, i) = 1$ and cost $b(d_q, i) = 0$. The capacity of temporal edge in the space-time graph is $c\left(v_i^t, v_i^{t+1}\right) = \infty$ and the cost is $b\left(v_i^t, v_i^{t+1}\right) = 0$ while the capacity of space edge is $c\left(v_i^t, v_j^{t+1}\right) = 1$ and the cost is $b\left(v_i^t, v_j^{t+1}\right) = 1$. Assuming that the destination node of the DTLH problem is i, the sink node of the flow network will be v_i^m. Figure 4 shows an example where the DTLH problem is transformed into a minimum cost maximum flow problem when the number of receiving nodes is 1.

Theorem 1 *The DTLH problem has a feasible solution if and only if the maximum traffic flow through the node v_i^m is k, and the minimum cost of each data stream is not greater than h.*

Proof When a DTLH problem has a feasible solution, each data unit has a path to the destination node. Since each node can transmit only one data unit per time slot, the path of each data unit to the destination node must not overlap in the space edge. So the maximum flow through the node v_i^m is at least k. Besides, the set of edges

$\{(S, d_1), \ldots, (S, d_k)\}$ is a cut-set of the network and the capacity of the cut-set is k, so the maximum flow of the network is k. Since the hop count of each data unit in the feasible solution must be within its maximum survival hops, the cost of each data stream in the network is no more than h.

When the maximum traffic in the network is k, and the minimum cost of each data stream is not greater than h, since the capacity of the edges in the network except the temporal edge is 1, there are k traffic flows in the network through the nodes $d_1, \ldots, d_k$ and the node v_i^m, and the traffic flow does not overlap in the space edge. Because the cost of each flow is not greater than h, so the k data units reach the destination node within the maximum survival hops, i.e. DTLH problem has a feasible solution.

5 Conclusion

In this paper, we first describe the data transfer problem with limited data hops and analyze the complexity of the problem. Then, according to the characteristics and constraints of the problem, a linear programming model is established. Finally, two special cases when the number of data units $k = 1$ and the number of receiving nodes $r = 1$ are discussed.

References

1. Xuan BB, Ferreira A, Jarry A (2003) Evolving graphs and least cost journeys in dynamic networks. Wiopt03 Modeling & Optimization in Mobile Ad Hoc & Wireless Networks
2. Bocquillon R, Jouglet A, Carlier J (2015) The data transfer problem in a system of systems. Eur J Oper Res 244(2):392–403
3. Chen J, Zhang W (2016) The right way to search evolving graphs. IEEE international parallel and distributed. In: Processing symposium workshops. IEEE Computer Society, pp 867–876
4. Huang M, Chen S, Zhu Y, Xu B (2011) Topology control for time-evolving and predictable delay-tolerant networks. IEEE Trans Comput 62(11):82–91
5. Merugu S, Zegura EW (2004) Routing in space and time in networks with predictable mobility. Georgia Institute of Technology
6. Hou Z, Yi X, Zhao Y, Zhang Y (2016). Information transmission path selection of navigation satellite network based on directional crosslink. In China satellite navigation conference (CSNC) 2016 proceedings, vol III

Research on MIMU/UWB Integrated Indoor Positioning

Yishuai Shi, Ancheng Wang, Jinming Hao and Bo Jiao

Abstract This paper investigates and proposes a low-precision Micro Inertial Measurement Unit (MIMU)/Ultra-Wideband (UWB) integrated positioning scheme, which aims at realizing indoor high-precision and high-stability localization. A loosely-coupled MIMU/UWB integrated positioning system based on position measurements is designed. Considering three possible cases in practical uses—normally working, UWB signal blockage and UWB gross error occurring, we establishes the simulation platform and the experimental environment to analyze and verify the performance of the integrated system. Simulation and indoor kinematic experimental results show that the MIMU/UWB integrated positioning system produces continuous and accurate position within an error of 2 m. It is also proved to have better continuity and robustness than both independent subsystems.

Keywords Indoor positioning · Integrated navigation · Ultra-Wideband · Inertial navigation · Kalman filter

1 Introduction

Indoor precise positioning technology promotes the development of new applications in both military and civil fields, such as intelligent shopping guidance, emergency handling and rapid rescue. Global Navigation Satellite Systems (GNSS) are widely applied outdoors due to their global coverage and reliable accuracy. However, GNSS often suffers from signal attenuation, multipath and blockage introduced by buildings and human motions in indoor environments.

Y. Shi · A. Wang (✉) · J. Hao · B. Jiao
School of Navigation and Aerospace Engineering,
Information Engineering University, Zhengzhou, China
e-mail: wangancheng@ymail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_42

Currently, indoor positioning is based mainly on wireless sensor network (WIFI, ZigBee, Bluetooth, etc.) for its low cost and ease of implement, but the robustness and direction information cannot be provided. Ultra-Wideband (UWB), with high anti-interference ability, excellent multipath resolving capability and low power consumption, is a new promising positioning technology especially for indoor applications. Inevitably, as a kind of radiolocation, UWB signal is also under the influence of blockage and interference. To compensate for the shortcomings of these, extra observations transmitted from other sensors are desired.

Inertial Navigation System (INS) has the advantage of self-contained navigation capacity, as well as adding attitude information synchronously to the positioning system. Yet, it suffers from inherent defect that positioning accuracy decreases as the drift error [7]. By fusing INS and UWB measurements, better accuracy, continuity and stability can be obtained, which is similar to the GNSS/INS integrated navigation system [6]. With the manufacturing technology for inertial components improved, Micro Inertial Measurement Unit (MIMU) based on Micro-electro-mechanical System (MEMS) is widely researched, which makes it possible to decrease the size and power consumption of the MIMU/UWB systems and therefore, be sufficient for indoor applications. Consequently, MIMU/UWB integrated indoor positioning is becoming a research hotspot [1–3, 8, 10]. Researchers in [3] proposed a new method based on INS/UWB for attitude angle and position synchronous tracking of indoor carrier, and the static and dynamic positioning results show that the positioning accuracy meets the requirements of indoor applications. In Ref. [8], a tightly-coupled GPS/UWB/INS cooperative positioning scheme is designed to provide better accuracy and improve the availability in GPS denied environments. The author of [10] investigated the combination of UWB and MEMS-based Pedestrian Dead Reckoning (PDR) technologies, aiming at using the PDR position as a temporary auxiliary result when UWB signal is blocked, and conversely, correcting PDR in a large degree when PDR drifts after a long time working.

2 MIMU/UWB Integrated Positioning System

2.1 Overview

Data provided by MIMU and UWB are combined by a sensor-fusion strategy based on Kalman Filter in order to yield optimal position estimation. A MIMU/UWB integrated positioning system, comprising three modules, is designed, as shown in Fig. 1. On the one hand, the SINS Module is used to carry out INS resolving with the measurements given by the inertial components, and on the other hand, the UWB Module provides the range measurement and outputs the calculated position. Both results are fed into the Fusion Module to realize optimal fusion. Notice that “P” and “A” represent the position and the attitude outputs.

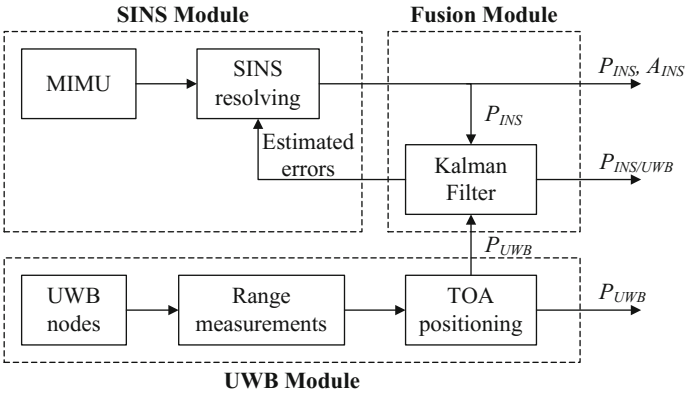


Fig. 1 Architecture of the MIMU/UWB integrated positioning system

2.2 Fusion Strategy Based on Kalman Filter

In the filter model, we use a 15-dimension state vector, which is expressed as follows:

$$\begin{aligned}\mathbf{X} = [\,&\phi_E \quad \phi_N \quad \phi_U \quad \delta V_E \quad \delta V_N \quad \delta V_U \quad \delta\lambda \\ &\delta L \quad \delta h \quad \varepsilon_x \quad \varepsilon_y \quad \varepsilon_z \quad \nabla_x \quad \nabla_y \quad \nabla_z\,]^T\end{aligned} \tag{1}$$

where ϕ_E, ϕ_N and ϕ_U are the misalignment angles of east, north and up direction, respectively, δV_E, δV_N and δV_U are the velocity errors, and $\delta\lambda$, δL and δh are the latitude, longitude and height errors. In addition, the gyro drift error vector $\boldsymbol{\varepsilon}$ and the accelerometer bias error vector ∇ are regarded as the random walk process vectors, which are modelled as follows:

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_b + \mathbf{w}_g \tag{2}$$

$$\nabla = \nabla_b + \mathbf{w}_a \tag{3}$$

where $\boldsymbol{\varepsilon}_b$ and ∇_b denote constant drift errors, $\mathbf{w}_g$ and $\mathbf{w}_a$ are white Gaussian noise vectors.

2.2.1 Dynamic Equation

The system dynamics model can be generalized in a matrix and vector form:

$$\dot{\boldsymbol{X}} = \boldsymbol{F}\boldsymbol{X} + \boldsymbol{W} \tag{4}$$

where $\boldsymbol{F}$ is the system coefficient matrix determined by Eqs. (2) and (3) and the INS error equations [5], $\boldsymbol{W}$ is the system noise vector. The dynamic equation can be represented in a block matrix form as follows:

$$\begin{bmatrix} \dot{\boldsymbol{\phi}}^n \\ \delta\dot{\boldsymbol{V}}^n \\ \delta\dot{\boldsymbol{r}}^n \\ \dot{\boldsymbol{\varepsilon}}^n \\ \dot{\boldsymbol{\nabla}}^n \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_{11} & \boldsymbol{F}_{12} & \boldsymbol{F}_{13} & -\boldsymbol{C}_b^n & \boldsymbol{0}_{3\times3} \\ \boldsymbol{F}_{21} & \boldsymbol{F}_{22} & \boldsymbol{F}_{23} & \boldsymbol{0}_{3\times3} & \boldsymbol{C}_b^n \\ \boldsymbol{0}_{3\times3} & \boldsymbol{F}_{32} & \boldsymbol{F}_{33} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \end{bmatrix} \begin{bmatrix} \boldsymbol{\phi}^n \\ \delta\boldsymbol{V}^n \\ \delta\boldsymbol{r}^n \\ \boldsymbol{\varepsilon}^n \\ \boldsymbol{\nabla}^n \end{bmatrix} + \begin{bmatrix} \boldsymbol{w}_g^n \\ \boldsymbol{w}_a^n \\ \boldsymbol{0}_{3\times1} \\ \boldsymbol{0}_{3\times1} \\ \boldsymbol{0}_{3\times1} \end{bmatrix} \tag{5}$$

2.2.2 Observation Equation

On the basis of the position difference vector between the MIMU and the UWB computation values, the observation equation can be derived:

$$\boldsymbol{Z}_{3\times1}(t) = \left[\boldsymbol{r}_I^n - \boldsymbol{r}_U^n\right] = \begin{bmatrix} \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times3} & \boldsymbol{0}_{3\times3} \end{bmatrix} \begin{bmatrix} \boldsymbol{\phi}^n \\ \delta\boldsymbol{V}^n \\ \delta\boldsymbol{r}^n \\ \boldsymbol{\varepsilon}^n \\ \boldsymbol{\nabla}^n \end{bmatrix} + \left[\boldsymbol{w}_r^n\right] \tag{6}$$

where "I" and "U" denote the observations from MIMU and UWB, "n" represents the navigation coordinate system, and $\boldsymbol{w}_r^n$ is the measurement noise vector. Parameters adopted in this paper to determine the position are latitude (λ), longitude (L), and height (h), in which case the observation vector can be specific:

$$\boldsymbol{r}_I^n - \boldsymbol{r}_U^n = \begin{bmatrix} \lambda_I - \lambda_U \\ L_I - L_U \\ h_I - h_U \end{bmatrix} \tag{7}$$

3 Simulation Researches

3.1 Overview

A simulation platform is built (referring to [9]) to evaluate the performance of the optimal filtering strategy of MIMU/UWB, which contains three modules. The first is the Trajectory Simulation Module, yielding position, velocity and attitude data according to the trajectory designed. At the same time, ideal outputs of MIMU and UWB are provided, that is to say, no errors are introduced into the raw data. The second is the SINS Simulation Module, which adds errors to ideal data mentioned above according to the models in Eqs. (2) and (3), and then, carries out SINS

resolving process. The third, also known as the most important segment, is the Fusion Simulation Module. Information fusion of measurements (with errors) is completed in this module, yielding results to be compared with the real trajectory for accuracy evaluation. The Root Mean Square Error (RMSE) is selected as the evaluation index in our case. Considering that plenty of abnormal cases may occur except system normally working, such as UWB signal blockage and UWB gross error occurring, we should verify the effectiveness in several different cases. The schematic of the proposed simulation platform is shown in Fig. 2.

To reflect different motion states as much as possible, a simulation trajectory is designed comprising static, uniform, accelerated, turning, climbing and other segments. The simulation time is 52.7 s and the output frequency is 100 Hz. The 3-D simulation trajectory is presented in Fig. 3.

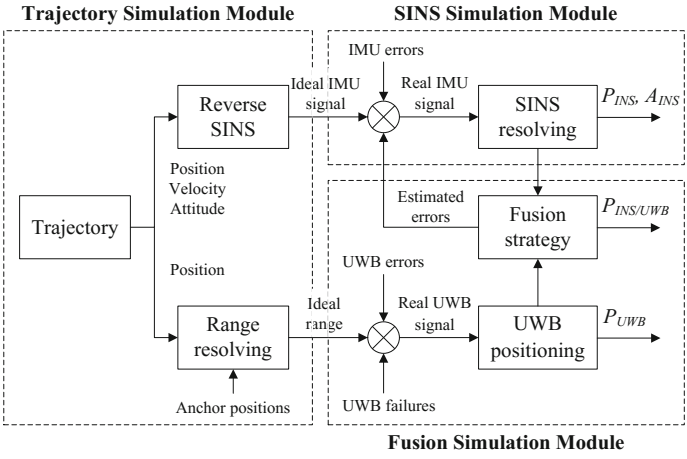


Fig. 2 The schematic of the proposed simulation platform

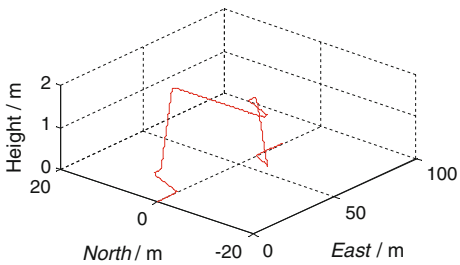


Fig. 3 The 3-D simulation trajectory

3.2 Simulation Research of MIMU/UWB Integrated Positioning

To overall evaluate the performance of the integrated positioning scheme, three cases are considered as follows:

Case 1: Normal case, in which both MIMU and UWB work normally with errors in a reasonable scope;
Case 2: Abnormal case, in which UWB suffers from signal blockages for a certain period of time;
Case 3: Abnormal case, in which a gross error occurs, making UWB position unreliable instantaneously.

We assume UWB positioning noise with a standard deviation of 1 m in east and north direction, 3 m in up direction, which is close to practical uses. Sensor errors are introduced into the MIMU measurements, as presented in Table 1.

3.2.1 Case 1

Through the comparison between the estimated position from the MIMU/UWB system and the computation values from the INS-only and the UWB-only systems, as shown in Fig. 4a, it is illustrated that the trajectory of INS coincides with the true one in a short term but suffers a cumulative error over time, while the UWB-only and the MIMU/UWB systems track the true trajectory smoothly. The partial enlargement of Fig. 4a yields Fig. 4b, which gives a further demonstration that the result of the MIMU/UWB system does not drift a lot compared to the UWB-only one. In addition, a higher output frequency is provided that reinforces the necessity of integrating additional sensors to make full use of more information. Time-varying height outputs from three systems are plotted in Fig. 4c, and the error amplitude of the integrated system is minimal. Particular focus has been given to the position errors of the integrated system (Fig. 4d), and the calculated root mean square (RMS) errors are 0.87, 0.73 and 1.54 m in the east, north and up directions, respectively, which reveals a better estimation. The RMS errors of the other two systems are also listed in Table 2 for comparison.

Table 1 Simulation error conditions of the MIMU

Errors of inertial components		Bias	Bias stability	
	Gyroscope	1°/h	50°/h	
	Accelerometer	50 μg	500 μg	
Errors of initial alignment		East	North	Up
	Attitude	0.5′	0.5′	20′
	Velocity	1 m/s	1 m/s	1 m/s
	Position	1 m	1 m	3 m

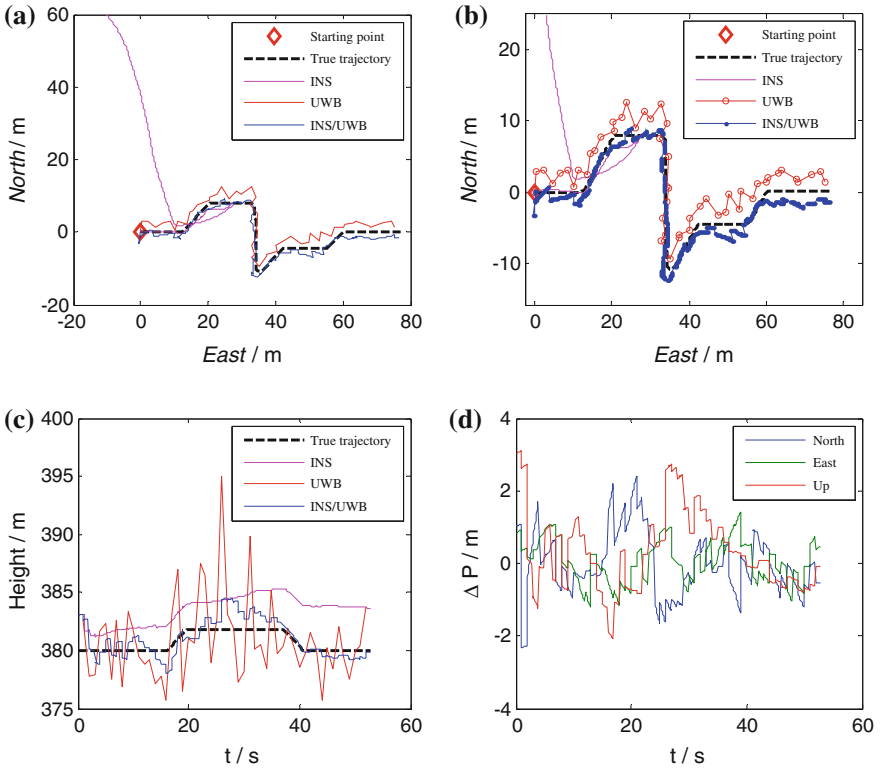


Fig. 4 Positioning performance in Case 1: **a** comparison among the trajectories in the east and north directions; **b** comparison among the trajectories in the east and north directions (partially enlarged); **c** comparison among the trajectories in the up direction; **d** the MIMU/UWB positioning errors

Table 2 Comparison between different systems in terms of RMS errors

	INS	UWB	INS/UWB
East (m)	67.73	1.00	0.87
North (m)	73.31	0.83	0.73
Up (m)	2.98	2.97	1.54

The obtained results show that the filter allows to provide better position estimation stably due to its capability of restraining errors. In addition, with the update rate promoted, the advantage of information fusion is presented.

3.2.2 Case 2

In Case 2, simulation error conditions are set the same as that in Case 1 (Table 1). An abnormal case, where the UWB signal is blocked from 18 to 23 s, is simulated

in view of the drawbacks of radio signal. The integrated system works normally at the very beginning. When the UWB signal is blocked, depending only on the calculation by MIMU accumulates the position errors with time. Subsequently, when the signal recovers, the trajectory is dragged back to the right way and remains relative high accuracy. The position errors in three directions are plotted in Fig. 5a, where the red dashed box represents the blockage phase.

The trajectories (Fig. 5b, c) reveal that continuous outputs are provided by the integrated system, even when the UWB signal suffers a blockage. In a short term, the accuracy does not drift away a lot. Afterwards, corrections can be made along with signal recovery. The obtained results indicate that the integrated system has better accuracy and stability than its subsystems.

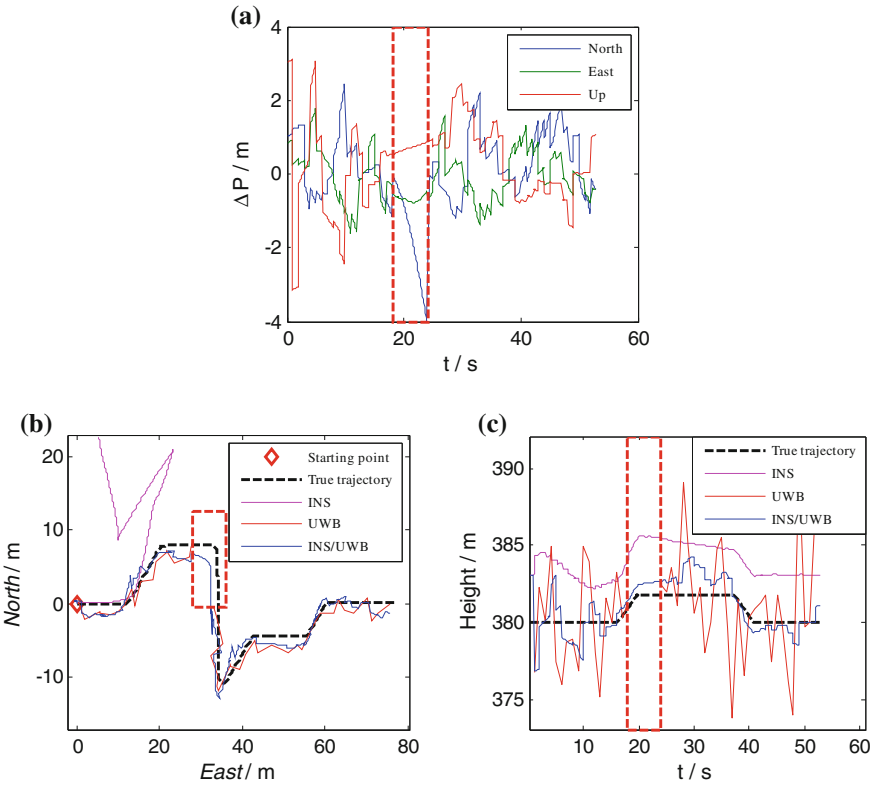


Fig. 5 Positioning performance in Case 2: **a** the MIMU/UWB positioning errors; **b** comparison among the trajectories in the east and north directions; **c** comparison among the trajectories in the up direction

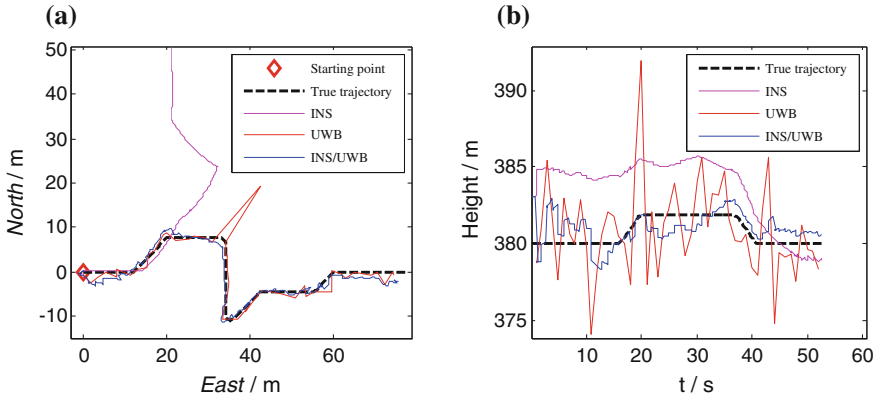


Fig. 6 Positioning performance in Case 3: **a** comparison among the trajectories in the east and north directions; **b** comparison among the trajectories in the up direction

3.2.3 Case 3

In Case 3, simulation error conditions are set the same as that in Case 1 (Table 1). Another abnormal case is simulated for the performance evaluations of the integrated system in presence of gross errors. Simulated gross errors of 10 m in all three directions were added to the UWB observations at the time of 20 s. As shown in Fig. 6a, b, clearly, the effects of gross errors, as well as the acute fluctuation due to directional dependence of the UWB range measurements, have been eliminated or, at least, alleviated. The height trajectory of the integrated system remains relative high accuracy and more stable than its subsystems. The obtained results show that the integrated system has strong robustness which prevents gross errors from disturbing the positioning accuracy to a large extent.

4 Indoor Kinematic Experimental Tests

4.1 Experimental Environment

Field tests were conducted to evaluate the performance of the proposed scheme in an indoor kinematic environment. The deployment of the test platform is presented in Fig. 7. For the MIMU inputs, we used the MTi-G-700 from Xsens. The bias stability of the gyro is 20°/h and that of the accelerometer is 2 mg. UWB units based on the DW1000, providing a ranging accuracy of 20 cm in static and line-of-sight (LOS) condition, were adopted to obtain UWB range measurements. The sensors were connected to a laptop to store the observations and fuse them offline.

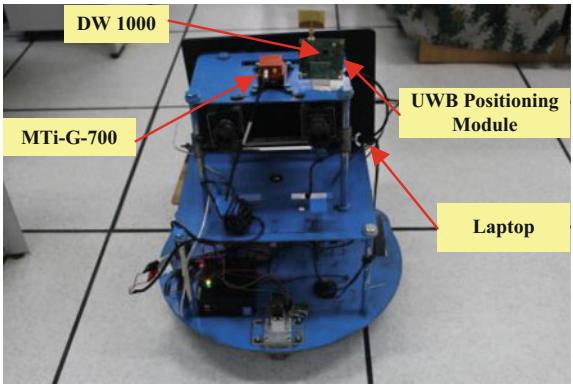


Fig. 7 The deployment of the test platform

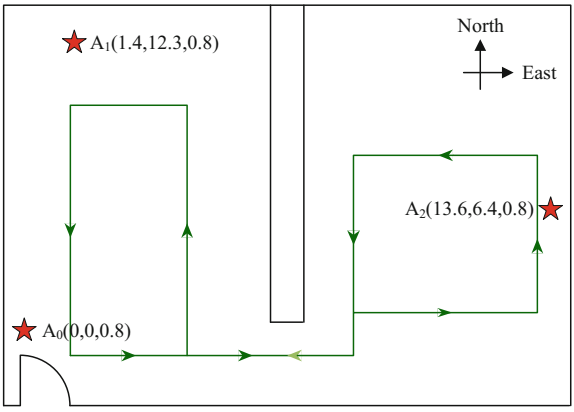


Fig. 8 The approximate kinematic trajectory

In Fig. 8, the test scenario is shown, together with the approximate kinematic trajectory designed previously in two rooms. A_0, A_1 and A_2 are the three UWB anchor nodes for range measuring and positioning. Notice that the two rooms are separated by a 0.12 m-wide wood wall.

4.2 Test Results

Similar to the simulation researches, the performance of the integrated system was verified in three cases. Failures were manually introduced in that no abnormal case occurred in our test.

4.2.1 Case 1

In the normal case (without signal blockages and gross errors), comparison between the UWB position and the MIMU/UWB position is illustrated in Fig. 9. It is noticed that the INS position is seriously affected by the cumulative errors and suffers a large deviation, for which reason INS is not taken into consideration alone. The accuracies are similar, but the lower amplitude of the integrated trajectory is recognized as a sign of better estimation. Owing to the lack of definite reference position, we can't make a quantitative evaluation of the positioning accuracy. However, Fig. 9 reveals that the maximum deviation is no more than 2 m.

4.2.2 Case 2

To test the performance of the integrated system when UWB signal is blocked, a gap of 5 s is manually created at the time of 110 s. The phase described in the red dashed box demonstrates that the integrated system still yields continuous position (Fig. 10).

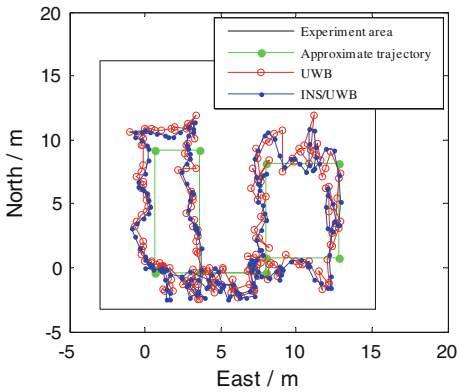


Fig. 9 Comparison among the trajectories in Case 1

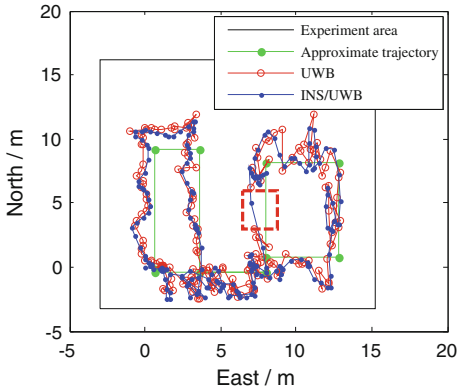


Fig. 10 Comparison among the trajectories in Case 2

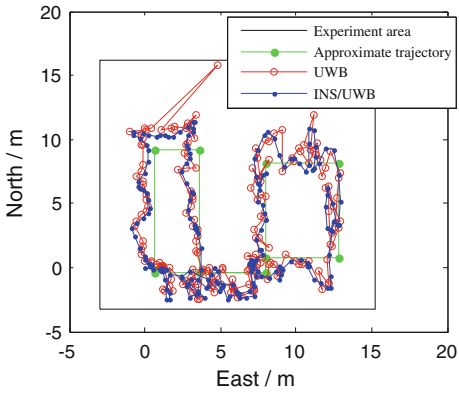


Fig. 11 Comparison among the trajectories in Case 3

4.2.3 Case 3

To test the performance of the integrated system when UWB gross error occurs, gross errors of 5 m in all three directions are manually added to the observation at the time of 170 s. The good shapes of blue trajectories prove that the effect of the added gross errors are eliminated, hence, the integrated system has the advantage of strong robustness (Fig. 11).

5 Conclusion

Indoor positioning technology is becoming very attractive in many fields of applications such as intelligent shopping guidance, emergency handling and rapid rescue. Combining INS and UWB positioning provides higher accuracy and robustness. In this paper, a MIMU/UWB integrated positioning system is presented by investigating the information fusion strategy. To evaluate the performance of the system, simulation researches and indoor kinematic experimental tests are conducted. It is noticed that three possible cases which may occur in practical uses are considered, including normally working, UWB signal blockage and UWB gross error occurring. The obtained results have shown the system capability of continuously and stably estimating the position within an error of 2 m. Particularly, effects of UWB signal blockages and gross errors can be alleviated or eliminated. However, owing to the lack of definite reference position in experimental tests, we can't make a quantitative evaluation of the positioning accuracy. Better experimental environment is expected. MIMU/UWB integrated indoor positioning is a hotspot with desirable value and wide prospect which demands both theoretical and experimental researches. For higher accuracy and stability, improvements of the fusion strategy and investigations on tightly-coupled positioning system are considered as future work.

References

1. Benini A, Mancini A, Marinelli A, Longhi S (2012) A biased extended kalman filter for indoor localization of a mobile agent using low-cost IMU and UWB wireless sensor network. IFAC Proc 45(22):735–740
2. De Angelis A, Nilsson J, Skog I, Händel P, Carbone P (2010) Indoor positioning by ultrawide band radio aided inertial navigation. Metrol Measur Syst 17(3):447–460
3. Fan Q, Wu Y, Hui J, Wu L, Yu Z, Zhou L (2014) Integrated navigation fusion strategy of INS/UWB for indoor carrier attitude angle and position synchronous tracking. Sci World J 215303
4. Li J, Wang J, Chen J, Wang A, Wang L, Ge L (2016) Precise indoor location system based on UWB. In: The 7th China satellite navigation conference (CSNC), Changsha
5. Luo J, Ma W, Yuan J, Yue X (2012) Theory and application of integrated navigation. Northwestern Polytechnical University Press, Xi'an
6. Qin Y, Zhang H, Wang S (2015) Theory of Kalman filter and integrated navigation, 3rd edn. Northwestern Polytechnical University Press, Xi'an
7. Qin Y (2014) Inertial navigation, 2nd edn. Science Press, Beijing
8. Wang J, Gao Y, Li Z, Meng X, Hancock C (2016) A tightly-coupled GPS/INS/UWB cooperative positioning sensors system supported by V2I communication. Sensors 16(7):944
9. Yan G (2004) Research on SINS algorithm and vehicle integrated navigation system. Master thesis, Northwestern Polytechnical University, Xi'an
10. Yang Z (2015) Study on high precision indoor positioning technology based on UWB/MEMS. Master thesis, China University of Mining and Technology, Xuzhou
11. Zhang Y (2014) Research on the key techniques of indoor localization. PhD thesis, University of Science and Technology of China, Hefei

Progress on Novel Atomic Magnetometer and Gyroscope Based on Self-sustaining of Electron Spins

S.G. Wang, C. Xu, Y.Y. Feng and L.J. Wang

Abstract The gyroscope based on atomic technology has the potential to provide the end user a high-performance device in a small package with low-power. Generally, the atomic gyroscope detects the rotation or angular rate of the object by measuring the Larmor precession frequency of spins. However, the precision is limited by the shorter coherence time caused by relaxation effects, and since the nuclear spin precession is often used in detection for atomic gyroscope, longer polarization time limits its application environment. Presented in this paper is a self-sustaining gyroscope based on electron spins. By non-destructively measuring the phase of the Larmor precession and regenerating the coherence via optical pumping, the Larmor precession can persist indefinitely, and the system can quickly regain polarization to environmental variations. Magnetic field measurement has been accomplished by using the self-sustaining technology. The precision of the magnetometer increases with time following a much faster τ^{-1} rule rather than the traditional $\tau^{-1/2}$ rule. The mean sensitivity is close to the shot noise in 300 ms, and the magnetometer has a quick response to sudden magnetic change. The self-sustaining technology can hopefully improve the measurement precision and the response time of the atomic gyroscope.

Keywords Atomic gyroscope · Atomic magnetometer · Electron spin

S.G. Wang (✉) · C. Xu · Y.Y. Feng · L.J. Wang
Joint Institute for Measurement Science, Tsinghua University, Beijing 100084, China
e-mail: wangsg@tsinghua.edu.cn

S.G. Wang · Y.Y. Feng · L.J. Wang
The State Key Lab of Precision Measurement Technology and Instrument,
Department of Precision Instruments, Tsinghua University, Beijing 100084, China

C. Xu · L.J. Wang
Department of Physics, Tsinghua University, Beijing 100084, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_43

1 Introduction

Using coherent superposition of the atomic states, the high-precision and absolute measurement can be realized. By detecting the Larmor precession of atomic spins, one can measure the absolute magnetic field and the angular velocity. However, in most of the coherent quantum measurement, the precision is limited by the coherent time of atomic states. Recently, we have engaged in developing a novel self-sustaining technology. By non-destructively measuring the phase of the Larmor precession and regenerating the coherence via coherent optical pumping, we break through the limit of the short lifetime of the electron spins, and extend the coherent time of the electron spins, making the Larmor precession persist definitely. The self-sustaining technology has the advantage of long coherent time and rapid polarizing speed. It is expected to realize the high-precision magnetometer and gyroscope based on the self-sustaining technology.

In this paper, the self-sustaining technology, the progress of the self-sustaining magnetometer and gyroscope are reported.

2 Self-sustaining Technology

2.1 Theory

In the traditional measurement on the frequency of the Larmor precession, the atom spins are firstly polarized with a circularly polarized pump light. The direction of the spins is along the incident direction of the pump light. When the pump light is turned off, the spins start to precess around the external field B. A damping signal will be detected after a while since the existence of the decoherence and the average value of the spins will relax to zero after a certain lifetime τ_0. We illustrate the method using ^{85}Rb atoms in the experiment. As shown in Fig. 1, setting the $\hat{z}$-axis along the magnetic field, at t = 0 a circularly polarized light along the $\hat{y}$-axis pumps all the atoms into the magnetic sublevel state $|F = 3,\ m_F = 3\rangle$, making the

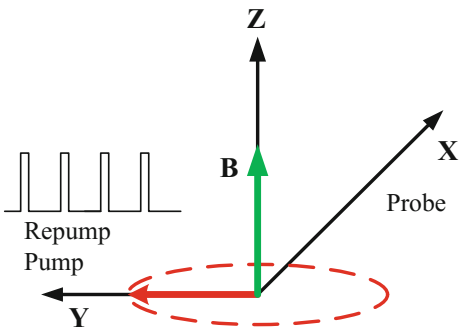


Fig. 1 The diagram of the Larmor precession of atomic spins in self-sustaining setup

direction of the spins along $\hat{y}$ direction. After the preparation of the polarization, the spins start to precess along the magnetic field. The probe light is along the $\hat{x}$-axis proportional to the pump light. A damping signal will be detected after a while since the existence of the decoherence.

At time $t = 2\pi/\omega_L$, where ω_L is the Larmor precession frequency, the population will evolve back to the initial $|F = 3,\ m_F = 3\rangle$ state. If the pumping light pulse is switched on right at this moment again, all atoms in $|F = 3,\ m_F = 3\rangle$ state will remain unaffected while atoms in all other states due to relaxation will be pumped back to $|F = 3,\ m_F = 3\rangle$ state. By coherent pumping, the damping signal is compensated. In this way, the spin is regenerated by the coherent pumping field and can maintain a very long lifetime. Since the time is dependent on the past signal of itself, the technology is called self-sustaining technology.

2.2 Setup

The system is composed of three parts, laser system, atoms with magnetic shielding, and control electronics. The setup is shown as Fig. 2.

A circularly polarized tunable external diode laser and the linearly polarized external diode laser are served as the pumping and repumping light, respectively as shown in Fig. 3. Both of them go through the same AOM (acousto-optic modulator) so that they can be switched on and off simultaneously. The probe laser propagates through the Glan-Taylor polarizer, the cell, and the Wollaston analyzer in sequence to form the Faraday rotation measurement [1]. Its beam size is about 2 mm and its frequency is red-detuned by 4 GHz from $|F = 3\rangle \rightarrow |F'\rangle$ of the D2 line. The three lasers come from three independent diode lasers and they are not phase locked to each other.

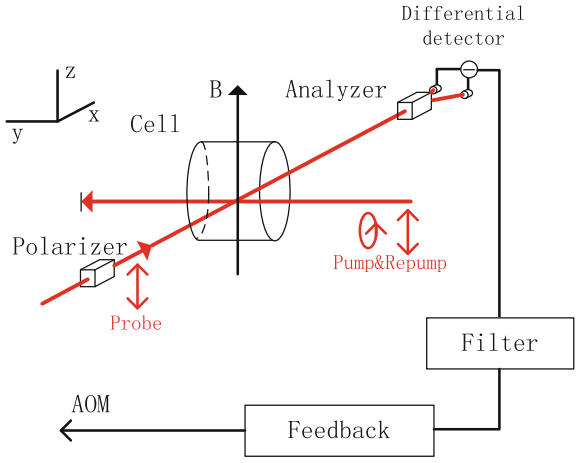


Fig. 2 The schematic diagram of the experimental facility

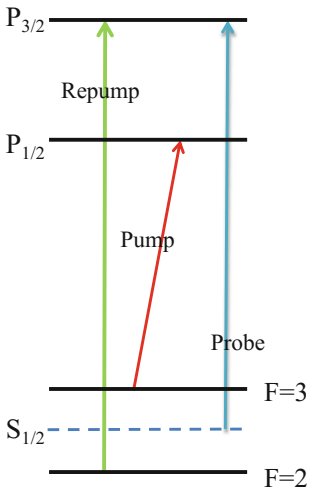


Fig. 3 The schematic diagram of the experimental facility

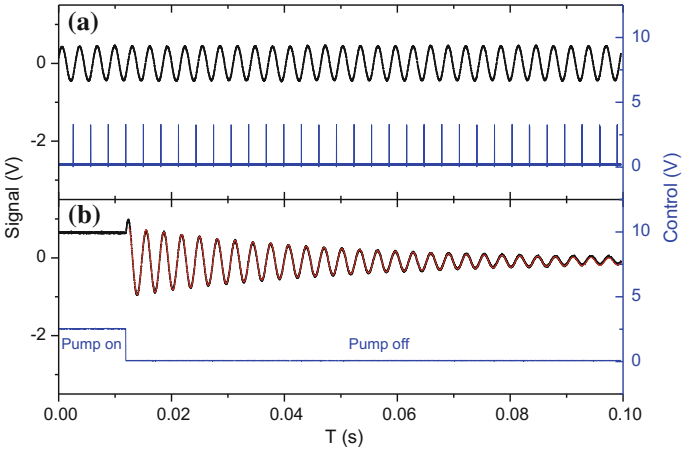


Fig. 4 **a** Spin precession signal of the self-sustaining magnetometer. **b** Signal observed in a single-pump free precession setup. The *black line* is the precession signal, the *blue line* is the control signal and the *red line* is a sine-decay fitting (color figure online)

The Rb cell is cylindrical 20 mm long and 20 mm in diameter. It is a self-made α-olefin coated cell containing natural abundance rubidium atoms. The cell is placed inside a five-layer μ-metal magnetic shield casing. A pair of Helmholtz coils controlled by a stable current source generates the uniform magnetic field B. The experiment is performed at room temperature.

2.3 Results

For comparison, the single pump Larmor precession signal and the self-sustaining signal are both measured. The result is shown in Fig. 4.

For the single pump measurement, the Larmor signal is damped due to the decoherence, and the sine-decay fitting gives a decay time of 30 ms.

3 Self-sustaining Magnetometer

3.1 Theory

For a standard free precession process, the sensitivity is [2]

$$\delta B = \frac{1}{\gamma \cdot SNR \cdot \sqrt{\tau \cdot \tau_0}} \tag{1}$$

where SNR is the signal-to-noise-ratio and γ is the gyromagnetic ratio. Extending the lifetime of atoms is the key point to improve the sensitivity of magnetometer. For a measurement time $\tau \gg \tau_{atom}$, the sensitivity improves as $\tau^{-1/2}$ due to the uncorrelated phase in each repeated measurement. In self-sustaining method, the life time $\tau_0 \approx \tau$, and hence the sensitivity will improve following a much faster τ^{-1} rule, until the accumulating phase error is large enough to destroy the phase coherence.

3.2 Results

As shown in Fig. 5, the Allan deviation [3] is the same as expected. The sensitivity is improved following the τ^{-1} rule. The τ^{-1} rule can extend to about 300 ms. It is followed by then turning into a $\tau^{-1/2}$ rule. The shot noise limit of spin projection is about 130 fT at t = 1 s, and hence, the sensitivity of the self-sustaining method has approached the limit of the shot noise.

The sensitivity of the free-precession repeated method is improved following the traditional $\tau^{-1/2}$ rule.

The FFT (fast Fourier transform) spectrum as shown in Fig. 6 also shows the superior of the self-sustaining method. The average noise level of 1 Hz to 10 Hz is 240 fT/$\sqrt{\text{Hz}}$, while the noise is 2.6 pT/$\sqrt{\text{Hz}}$ for the free precession repeated method [4].

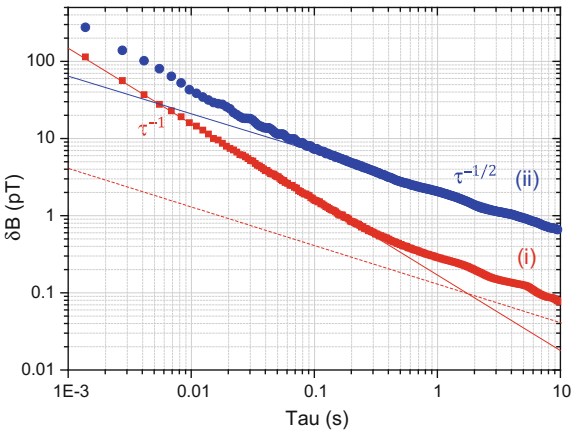


Fig. 5 (*i*) Allan deviation of the self-sustaining mode and (*ii*) the free-precession repeated mode. The *red dash line* represents the spin projection noise limit (Color figure online)

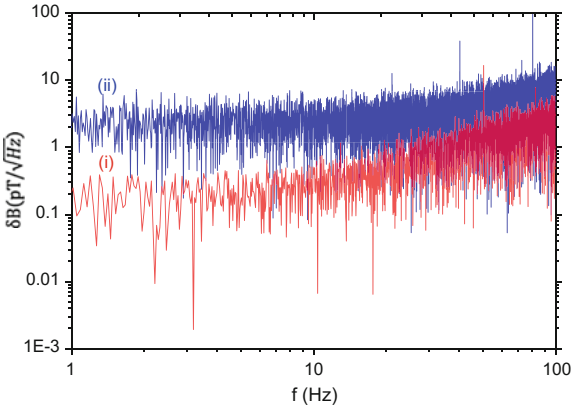


Fig. 6 (*i*) FFT of the Larmor frequency data of the self-sustaining mode and (*ii*) the free-precession repeated mode

4 Self-sustaining Gyroscope

The gyroscope based on the spins also measures the frequency of the Larmor frequency. The longer the lifetime of the spins, the higher the precision is. Hence, the nuclear spins are usually used in the gyroscope [5–7].

However, it will cost a longer time for the polarization of the nuclear spins, which limits the application environment. For electron spins, the polarization time is much less than the nuclear spins, but with a short lifetime. Hence, the self-sustaining gyroscope has the potential to solve the problems of both the nuclear spins and electron spins.

For ^{85}Rb atoms, the gyromagnetic ratios of the two ground states are opposite. Using two cells, one cell is prepared to the upper ground state, and one is prepared to the other state. Both cells are executed with the self-sustaining method, hence

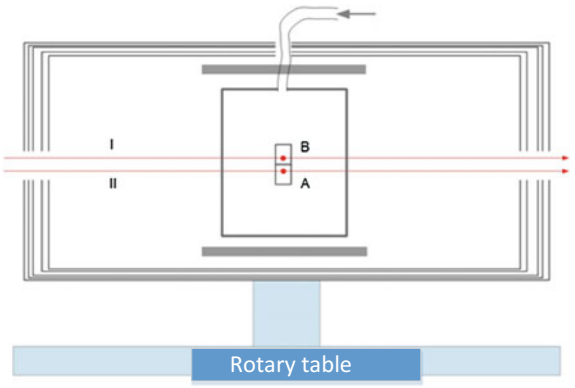


Fig. 7 The schematic diagram of the self-sustaining atomic gyroscopes. The cells are placed in a magnetic field shield case, and pumped by laser *I* and *II*, respectively. Laser *B* and *A* are probe laser, respectively

long coherence time can be accomplished with electron spins. After mixing the frequencies of the two Larmor frequency, the angular velocity Ω is obtained.

$$\omega_A = \gamma B + \Omega \tag{2}$$

$$\omega_B = -\gamma B + \Omega \tag{3}$$

The schematic diagram of the self-sustaining atomic gyroscopes is shown in Fig. 7.

The self-sustaining gyroscope with electron spins has the advantage of both long coherence time and rapid polarization time, hence it has a large potential to improve the performance of the atomic gyroscope.

5 Conclusion

We have introduced a novel self-sustaining method, which largely extends the lifetime of the electron spins. Using the method, we have realized the self-sustaining magnetometer, and it has the potential to realize a gyroscope based on electron spins with high-precision.

References

1. Budker D, Gawlik W, Kimball DF, Rochester SM, Yashchuk VV, Weis A (2002) Resonant nonlinear magneto-optical effects in atoms. Rev Mod Phys 74:1153–1201
2. Budker D, Romalis MV (2007) Optical magnetometry. Nat Phys 3:227–234
3. Allan DW (1966) The statistics of atomic frequency standards. Proc IEEE 54:221–230
4. Larsen M, Bulatowicz M (2012) Nuclear magnetic resonance gyroscope for DARPA's micro-technology for positioning, navigation and timing program. Northrop Grumman, USA

5. Meyer D, Larsen M (2014) Nuclear magnetic resonance gyro for inertial navigation. Gyroscopy Navig 5(2):75–82
6. Kornakc TW, Ghosh RK, Romalis MV (2005) Nuclear spin gyroscope based on an atomic comagnetometer. Phys Rev Lett 95(23):230801
7. Xu C, Wang SG, Feng YY, Zhao L, Wang LJ (2016) A self-sustaining atomic magnetometer with τ^{-1} averaging property. Sci Rep 6:28169

Design of Buoy Array Configuration in the Autonomous Positioning System of Underwater Vehicles

Suyang Liu, Chunjie Qiao and Yangyang Wang

Abstract When underwater vehicles are performing operations in the water, an autonomous positioning system can provide them with their location information in a convert and flexible way. In Wang et al. Design of autonomous underwater vehicle positioning system, 2016 [1], an autonomous positioning system for underwater vehicles is designed. The buoys on the water surface obtain their location with the help of GPS or Beidou Positioning system, and transmit spread spectrum acoustic signal to the underwater vehicle. In this system, the buoy array configuration is a key factor regarding to positioning precision. In this paper, the influence exerted by buoy array configuration on positioning precision is analyzed; to design a better configuration, two principles of configuration design are proposed, which are nonsingularity of matrix G and minimal position dilution of precision (PDOP); and with the limit of maximum working distance of acoustic transducers, a central radiation configuration with minimum elevation angle is proposed. Through simulation, the PDOP values of three different configurations are compared upon the same region. It is concluded that the central radiation configuration has relatively smaller average PDOP value and no singular matrix G in the locating region, which is, therefore, a better configuration for practice.

Keywords Underwater vehicle positioning · Buoy array configuration · PDOP · Central radiation configuration

S. Liu (✉) · C. Qiao · Y. Wang
National University of Defense Technology,
No. 137 Yanwachi Street, Changsha 410073, China
e-mail: lsy_august@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_44

1 Introduction

In recent years, underwater vehicles are used in a variety of applications, such as underwater search and rescue, engineering operations [2, 3], intelligence reconnaissance, underwater weapons deployment and etc. When performing tasks in the water, underwater vehicles usually need their location information. The traditional positioning method is to lay a sonar array on the seabed and measuring the distances or orientations of the underwater vehicle from base array elements [4]. However, a large-scale sonar array has to be laid and calibrated in this method, which is demanding and not flexible. In [1], a scheme of autonomous positioning system for underwater vehicles is proposed. In this method, four buoys are laid on the sea surface, which acquire their location through GPS or Beidou positioning signal and transmit spread spectrum acoustic signal to the underwater vehicle in the meantime. The underwater vehicle measures pseudo-ranges between itself and the buoys through received signal and obtains its absolute location by coordinate transformation. A diagram of the system is shown in Fig. 1. However, the buoy configuration was not discussed in depth in [1].

The buoy configuration plays an important role in the accuracy of positioning results. In most cases, Dilution of Precision (DOP) is used to represent the magnifying power of measurement error into the final position results. Therefore, DOP can be used to evaluate buoy configurations [5]. In [6], it was pointed out that when buoys and the underwater vehicle forms a regular tetrahedron, the DOP can be minimized. However, in the autonomous positioning system of underwater vehicle, buoys float above the underwater vehicle, resulting in elevation angles which are greater than 0. Based on the above research, this paper proposes a design of buoy configuration, which is called central radiation configuration with minimum elevation angle. The simulation results show that the proposed method can guarantee better geometric configuration under same conditions, thus ensuring smaller positioning error.

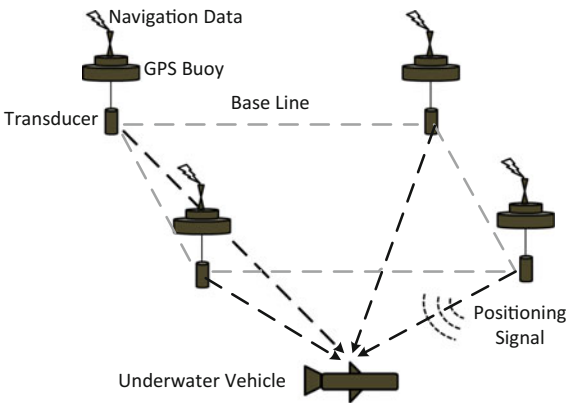


Fig. 1 A diagram of the autonomous positioning system of underwater vehicles

2 Design of Buoy Array Configuration

2.1 Positioning Precision and Position Dilution of Precision

If there is a clock bias between buoys and the vehicle, it would be necessary to use four buoys to complete vehicle localization. In this case, the locating equations are

$$\begin{cases} \sqrt{(x_1-x)^2+(y_1-y)^2+(z_1-z)^2}+t_u=\rho_1-\varepsilon_1 \\ \sqrt{(x_2-x)^2+(y_2-y)^2+(z_2-z)^2}+t_u=\rho_2-\varepsilon_2 \\ \sqrt{(x_3-x)^2+(y_3-y)^2+(z_3-z)^2}+t_u=\rho_3-\varepsilon_3 \\ \sqrt{(x_4-x)^2+(y_4-y)^2+(z_4-z)^2}+t_u=\rho_4-\varepsilon_4 \end{cases} \tag{1}$$

In the formula, (x_i, y_i, z_i) stands for the three-dimensional position of the ith buoy, which is included in the locating message transmitted by this buoy. ρ_i stands for the pseudo-range between the ith buoy and the vehicle, which is acquired by spread spectrum receiver on the vehicle. (x, y, z) stands for the unknown location of the vehicle. t_u stands for the spatial distance corresponding to the clock difference. ε stands for measurement error.

In [5], the following assumptions are made:

1. Measurement error of each buoy obeys the same normal distribution, whose mean value is zero and variance is σ^2_{URE}.
2. The measurement error of different buoys is independent.

When the above assumptions are satisfied, the positioning error covariance matrix can be expressed as

$$Cov\left(\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta t_u \end{bmatrix}\right) = \left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\sigma^2_{URE} = \mathbf{H}\sigma^2_{URE} \tag{2}$$

The coefficient matrix G is

$$\mathbf{G} = \begin{bmatrix} -\cos\alpha_1 & -\cos\beta_1 & -\cos\gamma_1 & 1 \\ -\cos\alpha_2 & -\cos\beta_2 & -\cos\gamma_2 & 1 \\ -\cos\alpha_3 & -\cos\beta_3 & -\cos\gamma_3 & 1 \\ -\cos\alpha_4 & -\cos\beta_4 & -\cos\gamma_4 & 1 \end{bmatrix} \tag{3}$$

where $\alpha_i, \beta_i, \gamma_i$ stands for the angle between the ith buoy and the x-axis, y-axis, and z-axis respectively.

Equation (2) shows that the accuracy of the static positioning precision of the vehicle is mainly determined by the measurement error and the weight matrix H.

The measurement error consists of two main part, which are buoy position error and pseudo-range measurement error. The buoy position error is the error of buoy location obtained from GPS or Beidou service. The pseudo-range measurement error is produced during the signal process in the receiver. There are many sources of pseudo-range measurement error, including bending transmitting path of acoustic signal, measurement error caused by multipath and background noise in the ocean [7]. The pseudo-range measurement error is affected by the ocean environment to a large extent.

While the pseudo-range measurement error is difficult to be controlled, the weight coefficient matrix H can be designed to minimize the positioning error. It is often the case to use position dilution of precision (PDOP) to represent amplifying power of the measurement error into the positioning error, which is defined as [5]

$$PDOP = \sqrt{h_{11} + h_{22} + h_{33}} \tag{4}$$

where h_{11}, h_{22}, h_{33} are the first three diagonal elements, which are the amplifying power of error variance on x axis, y axis, z axis respectively. With the same measurement error, smaller PDOP value contributes to smaller error in the positioning results. Actually, PDOP matrix is the Cramer-Rao lower bound on estimates of position given that the pseudorange errors are Gaussian distributed [8]. Meanwhile, the geometric matrix G is only related to geometric configuration of buoys and the vehicle. Therefore, the PDOP value can be used as an indicator of buoy configuration performance.

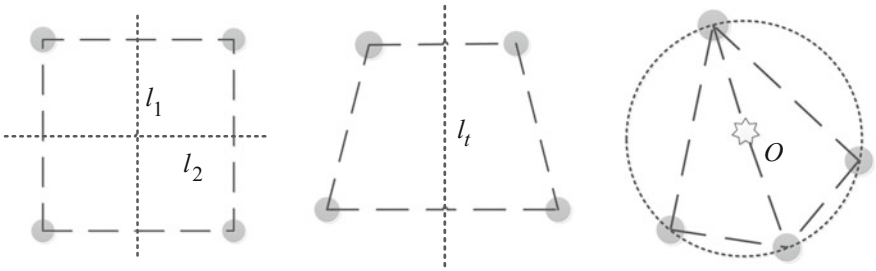
2.2 *Central Radiation Configuration with Minimum Elevation Angle*

In this paper, we mainly consider two aspects when designing the buoy configuration: one is the nonsingularity of the matrix G, which determines whether the positioning system can work stably in the region of interest; the other one is smaller PDOP value to minimize the amplifying power of the measurement error into the positioning error.

1. *Singularity of Matrix G*

Three types of buoy configurations are shown in Fig. 2, all of which has singularity points in the locating area. The first one is a square configuration. In the two vertical planes containing two axes of symmetry l_1, l_2, G is singular [9]. The second ones is an isosceles trapezoid configuration. In the vertical plane passing its axis of symmetry l_t, G is singular [9]. The last one is a circle configuration. On the vertical line through the center O, G is singular.

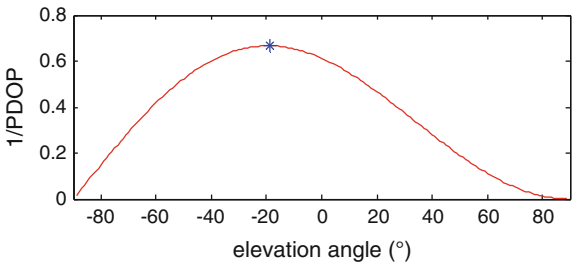
If matrix G is singular, then $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ is also singular. Therefore, there is no inverse matrix and it is impossible to compute PDOP. In fact, if matrix G is singular, there



(a) square configuration (b) isosceles trapezoid configuration (c) circle configuration

Fig. 2 Three buoy configuration of singular matrix G

Fig. 3 The variation of 1/PDOP at different elevation angles



would be more than one solution to the locating equations [6]. Thus, when designing the buoy configuration, we should avoid presence of singular points in vehicle working area. Otherwise, the positioning system cannot work stably in the area.

When four buoys are laid in central radiation configuration, which is one buoy float above the vehicle while the other three locate on the same circle around the central buoy, there is no singular points in the working area, ensuring the system work stably.

2. *PDOP Value*

In [6], it has been proved that the PDOP value is the smallest when the four floats constitute the tetrahedron and the vehicle is located in the center of the tetrahedron. However, the buoys are located on the water surface, thus, elevation angles are greater than 0. Figiure 3 shows the variation of PDOP at different elevation angles. In order to highlight the difference, the ordinate is the inverse of PDOP. It can be seen that PDOP is the smallest at −19.48° (the elevation angle corresponding to tetrahedron), and that as the elevation angle increases from that point, the PDOP gradually increases. Therefore, in order to maintain a small PDOP, elevation angle should be reduced as much as possible.

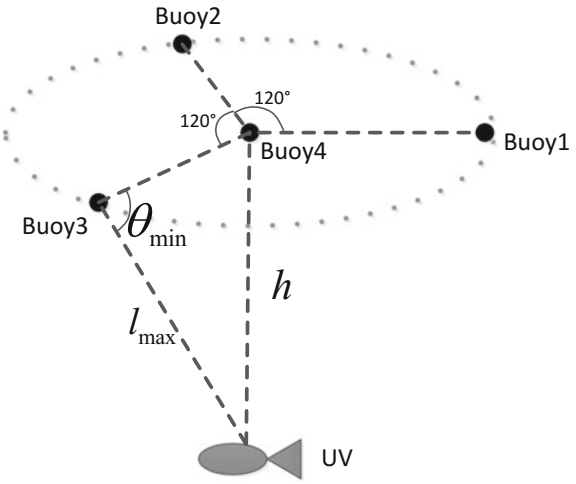


Fig. 4 Central radiation configuration with minimum elevation angle

However, in real-world practice, elevation angles cannot be infinitely close to 0°. First of all, the vehicle needs to navigate in a certain depth of water in order not to be exposed. Secondly, due to the limited power of transducers, the working propagation distance of acoustic wave is limited. The maximal working distance can be used to represent the propagation distance. It defined as the maximal sound path distance, which satisfies the pseudo-range measurement precision of spread spectrum signal when confidence probability and power limit of transducers are given [7].

These two constraints confine the elevation angle in a certain range $[\theta_{\min}, 90°]$. The minimum elevation angle is greater than 0, which can be approximately calculated from the navigation depth of the vehicle, and the maximum working distance of the transducer as follows

$$\theta_{\min} \approx \arcsin \frac{h}{l_{\max}} \tag{5}$$

Based on the above two design requirements, this paper uses the central radiation configuration with the minimum elevation angle, as shown in Fig. 4.

3 Comparison of Three Buoy Configurations

In order to evaluate the central radiation configuration with the minimum elevation angle presented in the previous section, PDOP values of three configurations with the same minimum elevation angle are calculated and compared in this section. In the simulation, the depth of underwater vehicle is assumed to be the same, which is 100 m; and the maximum working distance of transducers are assumed to be the same, which are 400 m. Therefore, the minimum elevation angle of the buoy

relative to the vehicle is 14.48°. The locating region where vehicle may appear is $D = \{(x, y, -100)|x \in [-100, 100], y \in [-100, 100]\}$.

3.1 Square Configuration

The coordinates of the four buoys are shown in Fig. 5. The PDOP values in the locating region are shown in Fig. 6 which is sampled every 5 m along x axis and y axis. On the two symmetric axes of the square, the matrix G is singular; near the two axes, especially near the point (0, 0, −100), the matrix G is nearly singular, and the PDOP is extremely large. The mean value of PDOP over the locating region except singular points is 885.54.

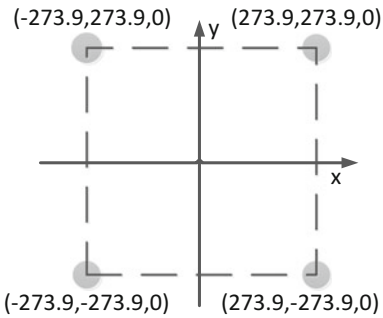


Fig. 5 Coordinates of four buoys in a square configuration

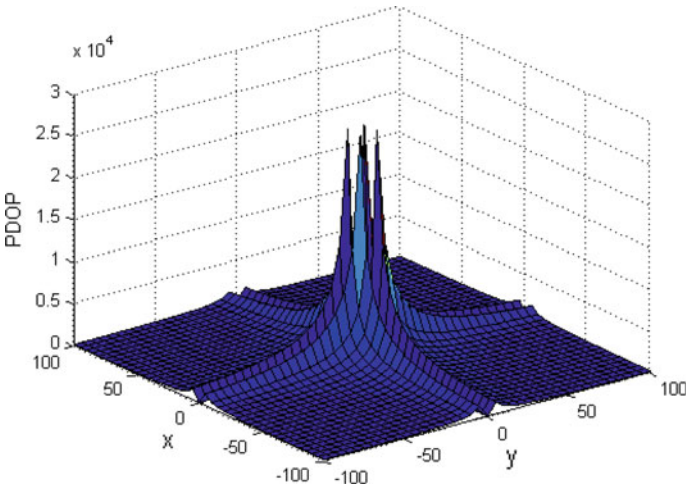


Fig. 6 PDOP in locating region of the square configuration

3.2 Diamond Configuration

The coordinates of the four buoys are shown in Fig. 7. The PDOP values in the locating region are shown in Fig. 8. The shape of PDOP surf is like a 'saddle'. The PDOP value decreases as the locating point moves away from the origin to each side of the x axis; on the contrary, it increases as the locating point moves away from the origin to each side of the y axis. There is no singular point in the locating region. And the mean value of PDOP is 5.03.

3.3 Central Radiation Configuration

The coordinates of the four buoys are shown in Fig. 9. The PDOP value in the locating region is shown in Fig. 10. The PDOP value is lower in the center and

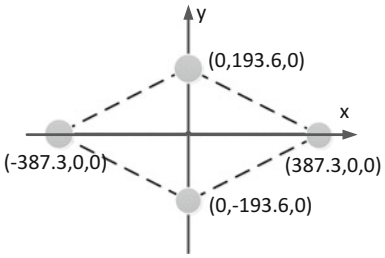


Fig. 7 Coordinates of four buoys in a diamond configuration

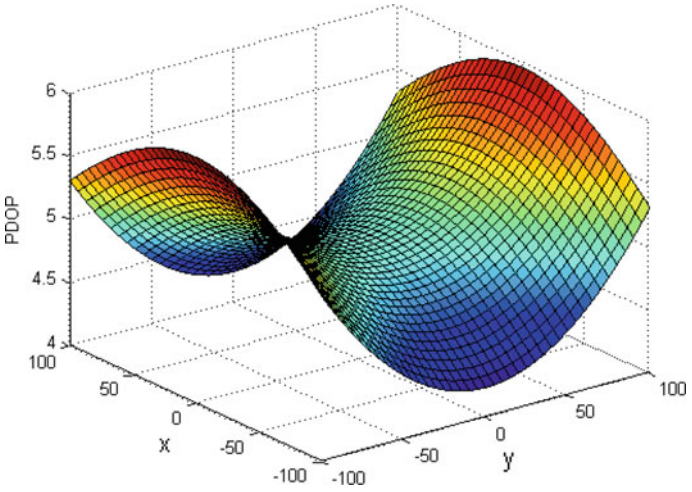


Fig. 8 PDOP in locating region of the diamond configuration

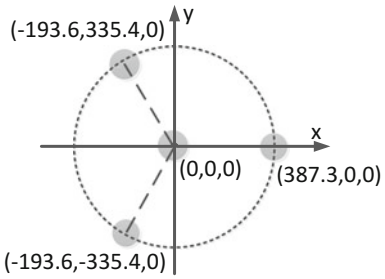


Fig. 9 Coordinates of four buoys in a central radiation configuration

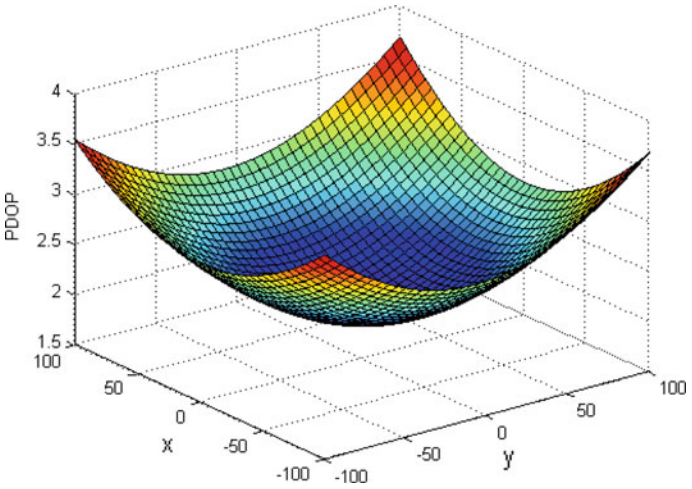


Fig. 10 PDOP in locating region of the central radiation configuration

higher on the edge and has the same gradient at any direction from the origin. There is no singular point in the locating region. The mean value of PDOP is 2.57, which is relatively lower than that of diamond configuration.

4 Conclusion

Based on the autonomous positioning system of underwater vehicle, this paper focuses on the design of buoy configuration for higher positioning precision. In this paper, the influence of buoy array configuration on the positioning precision is analyzed, and a central radiation configuration with the minimum elevation angle is proposed. Through simulation, this paper compared the PDOP value and the singularity of matrix G of three different configurations in the same locating region.

Conclusions are drawn as follows: square configuration has singular points and nearly singular points in the locating region; the diamond configuration and the central radiation configuration do not have singular points; and the average PDOP of the central radiation configuration is lower than that of the diamond configuration. Therefore, it is reasonable to consider central radiation configuration with the minimum elevation angle as a better geometric configuration, which can be implemented in the autonomous positioning system.

References

1. Tian T (2007) Underwater positioning and navigation technology. National Defense Industry Press, Beijing
2. Wang Y, Qiao C, Liu S (2016) Design of autonomous underwater vehicle positioning system. In: 2016 advanced information management, communicates, electronic and automation control conference IMCEC 2016
3. Sun D, Deng C, Qiao H et al (2012) The application of underwater acoustic positioning systems in ocean engineering. Tech Acoust 31(2):125–132
4. Ning J, Wu Y, Sun D (2014) The development of LBL acoustic positioning system and its application. Hydrogr Surveying Charting 34(1):72–75
5. Xie G (2009) Principles of GPS and receiver design. Publishing House of Electronic Industry, Beijing
6. Xue S, Yang Y (2006) Nested cones for single-point-positioning configuration with minimal GDOP. Wuhan Univ J Inf Sci 31(4):1370–1373
7. Chaffee J, Abel J (1994) GDOP and the Cramer-Rao bound. In: Position location and navigation symposium IEEE Xplore, pp 663–668
8. Xue S, Dang Y, Zhang C (2006) Research on setting 3D network of underwater DGPS. Sci Surveying Mapp 31(4):23–24
9. Wang Y (2016) A research on underwater maneuvering platform positioning system based on GPS buoys. National University of Defense Technology, Changsha

Hardware In-Loop System for X-ray Pulsar-Based Navigation and Experiments

Dapeng Zhang, Wei Zheng, Lizhi Sheng, Yidi Wang and Neng Xu

Abstract X-ray pulsar-based navigation uses natural objects, the neutron star, in space as the navigation signal source. The advantages of the method are navigation information is complete, and the reliability and autonomy are high. It is a research hot spot at present both at home and abroad. As a result of the X-ray signal from the pulsars is very weak, it cannot penetrate the thickset atmosphere. In order to validate the pulsar navigation algorithms closer to the real conditions on ground, the special Hardware in-Loop System should be used to do the experiments. This paper adopted the system "Tianshu-II" which is developed by National University of Defense Technology and Xi'an Institute of Optics and Precision Mechanics research institute. A series of X-ray pulsar-based navigation experiments are carried out. Experimental results show that the algorithms are reliable. They are verified to be effective in the hardware-in-the-loop simulation.

Keywords X-ray pulsar-based navigation · Simulation and verification · Hardware in-loop system · Experiments

1 Introduction

The pulsar is a kind of neutron star rotating quickly. The rotating axis and the magnetic axis don't coincide and the two magnetic poles transmit beam. When the beams sweep on the detector, the detector receives a pulse signal [1]. Since the first pulsar is discovered, scientists consider the rotating period of the pulsar is so stable that it could be used as natural clock. X-ray pulsar-based navigation is proposed in

D. Zhang · W. Zheng (✉) · Y. Wang
College of Aerospace Science and Engineering,
National University of Defense Technology, Changsha 410073, China
e-mail: zhengwei@nudt.edu.cn

L. Sheng · N. Xu
Xi'an Institute of Optics and Precision Mechanics of CAS, Xi'an 710119, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_45

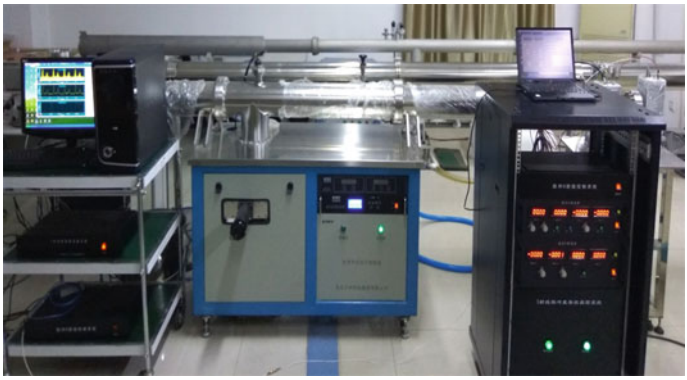


Fig. 1 Hardware in-loop system for X-ray pulsar-based navigation

1970s. From the end of 20th century to the turn of the century, a whole navigation scheme is formed gradually [2, 3]. According to energy range of the radiation, the pulsars can be divided as X-ray pulsar, radio pulsar and gamma-ray pulsar and so on. In order to reduce the mass and power consumption, the X-ray pulsars are selected in the navigation [4–6] (Fig. 1).

Hardware in-Loop System for X-ray Pulsar-based Navigation is used for verifying the X-ray pulsar navigation theory, evaluating the performance of the algorithm. Based on the input parameter from users, the system can output the single photon signal which the flux is controllable, the period is variable and the profile is settable arbitrarily. The system is competent for static signal processing, dynamic signal processing, energy spectrum construction, navigation algorithm verification and performance evaluation of the detectors.

2 Components of "Tianshu II"

"Tianshu II" is a kind of hardware in-loop system for X-ray pulsar-based navigation [7]. It is developed by National University of Defense Technology and Xi'an institute of optics and precision mechanics of CAS. It is consist of central control equipment cabinet, vacuum environment simulation system, X-ray generator, X-ray detector, software system and soon (Fig. 2).

2.1 X-ray Generator

The X-ray generator is equipped by the grid-controlled X-ray tube. If fixing the filament's current and anode's voltage, the grid-controlled X-ray tube changes the output flux of X-ray by changing the voltage of the grid. Transform the pulsar

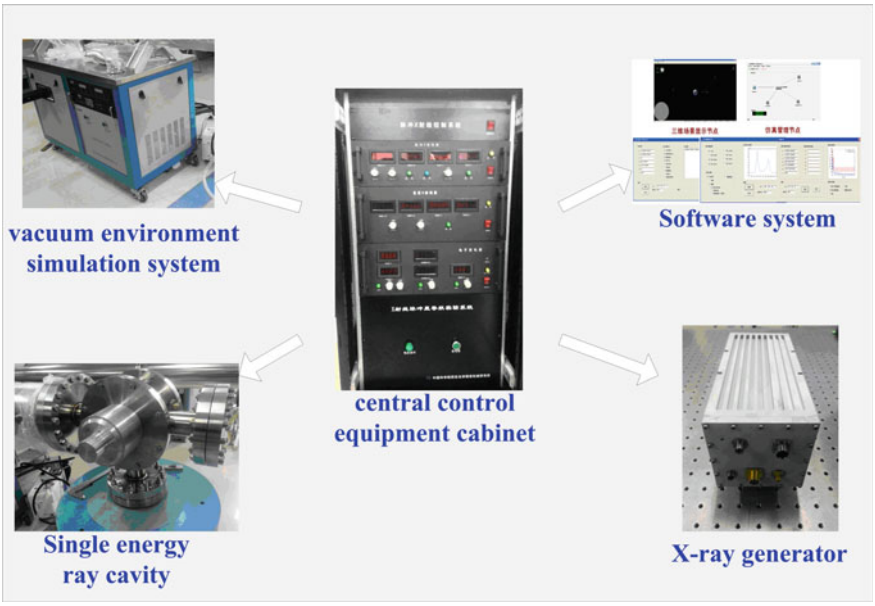


Fig. 2 Components of "Tianshu II"

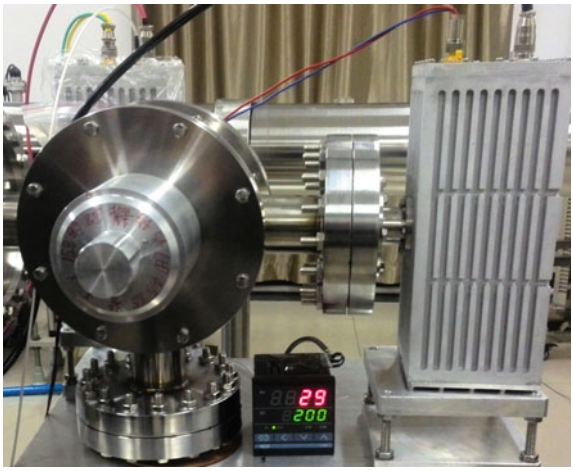


Fig. 3 The X-ray generator of the system

profile data to the voltage value, and drive the grid-controlled X-ray tube. So the flux of the X-ray is modulated by the time. The X-ray tube has no window structure, so the low energy X-ray is reserved (Fig. 3).

2.2 Vacuum Environment Simulation System

Owing to the energy of most of the photons sent by the X-ray generator is below 2 keV, these low energy photons can only be transmitted in the vacuum environment in respect that they will decay quickly in the atmosphere or any other medium. Simultaneity, the open tube in the X-ray generator needs to work in the vacuum environment too (Fig. 4).

"Tianshu II" system uses high-performance compound molecular pump and rotary vane mechanical pump to bleed the air. The speed is 600 L/s. It will take one hour to reach the vacuum degree of 10^{-4} Pa.

2.3 The Time Synchronization System for X-Ray Source and Detector

The traditional rotating-plate modulation scheme is hard to implement the synchronization between X-ray source and detector. "Tianshu II" uses electronic modulation method. This method can use the principal computer to trigger the

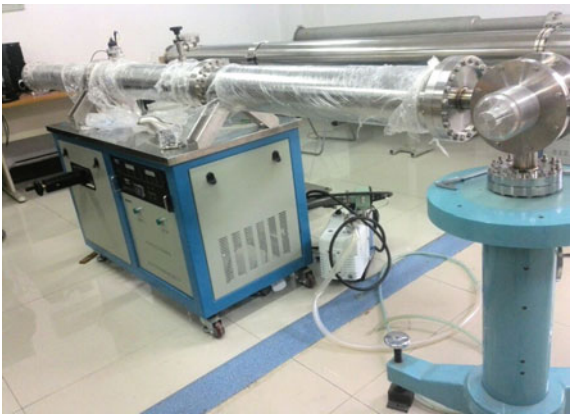


Fig. 4 Vacuum environment simulation system

Fig. 5 High-performance Rb atom clock

X-ray source and detector synchronously. Furthermore, use the atom clock to synchronize the interval of seconds (Fig. 5).

3 The Experiments Based on the "Tianshu II" System

3.1 The Dynamic Signal Simulation

When the spacecraft moves as its orbit, the signal received by the detector contains the Doppler frequency shift. To simulate this kind of signal requires the precision of the period tracking and the flux controlling is very high.

1. **Experiments of dynamic signal for low earth orbit**

Select PSR B0531+21 as the observation object. Set the flux is 276.4793 ph/s. The simulation time interval is 90 min. The orbit is low earth orbit. The six roots are, $[x_0, y_0, z_0, v_{x0}, v_{y0}, v_{z0}] = [0, 6578000, 0, -7784.3384, 0, 0]$, the unit of position is meter and the unit of velocity is meter per second.

Search the periods every 1 min [8] and the period tracking result is shown in Fig. 6. The recovery profile is shown Fig. 7.

2. **Experiments of dynamic signal for high earth orbit**

Select PSR B1937+21 as the observation object. Set the flux is 431.5924 ph/s The simulation time interval is 90 min. The orbit is low earth orbit. The six roots are, $[x_0, y_0, z_0, v_{x0}, v_{y0}, v_{z0}] = [-5654543.43, 41617618.12, 0, 1526.30, -207.37, -2667.93]$, the unit of position is meter and the unit of velocity is meter per second.

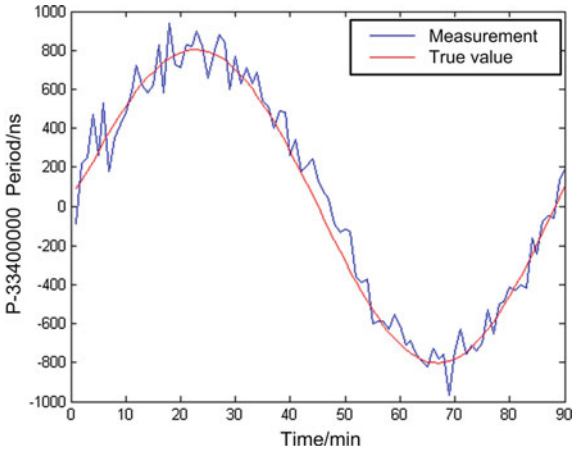


Fig. 6 The period tracking result

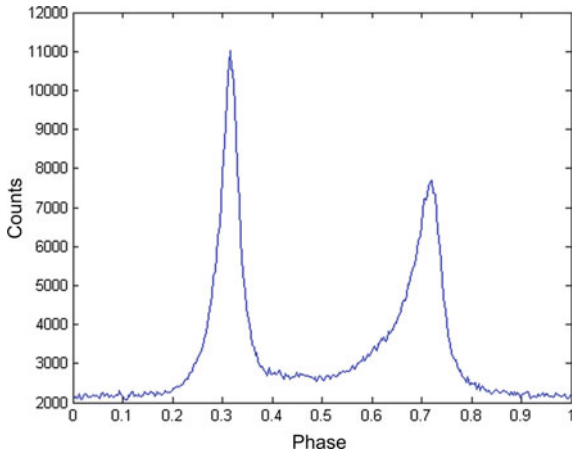


Fig. 7 The recovery profile using orbit information

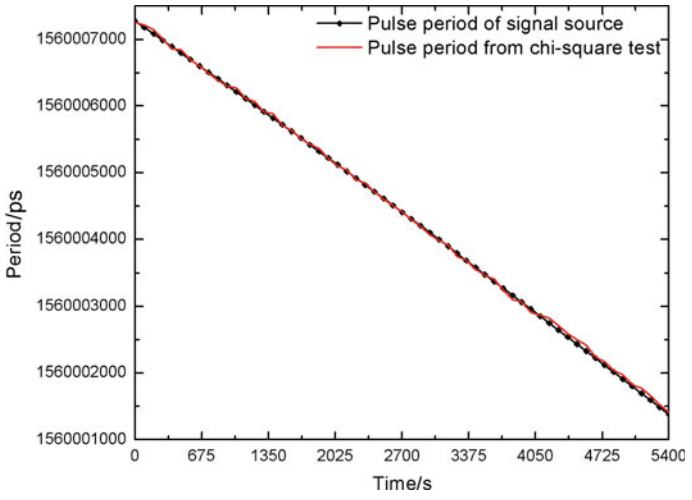


Fig. 8 The period tracking result

Search the periods every 1 min and the period tracking result is shown in Fig. 8. The recovery profile is shown Fig. 9.

Use standard deviation as the error evaluation criterion, the experiment results show the precision of period tracking can reach nanosecond level.

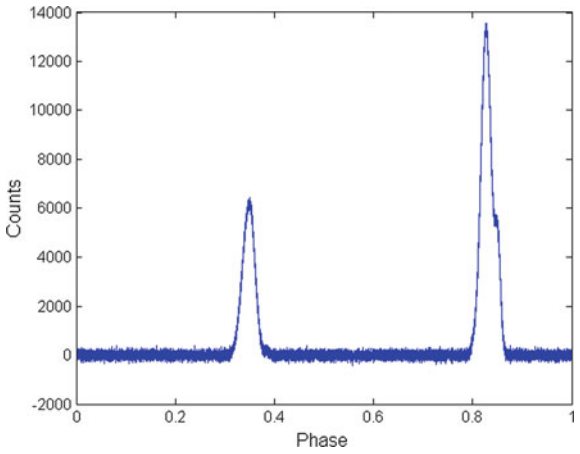


Fig. 9 The recovery profile using orbit information

3.2 The Static Signal Simulation

1. **Experiments of profile similarity**

"Tianshu II" system can simulate arbitrary pulsar's profile. In this experiment, we use SDD detector to receive the photon signal. Recover the signal by epoch folding method. Compare this recovery profile with the standard profile. According to the Pearson's linear correlation coefficient, the comparability is obtained.

Pearson's linear correlation coefficient is shown as following,

$$r = \frac{\sum_{i=1}^{n}(A_i - \overline{A})(B_i - \overline{B})}{\sqrt{\sum_{i=1}^{n}(A_i - \overline{A})^2}\sqrt{\sum_{i=1}^{n}(B_i - \overline{B})^2}} \tag{1}$$

where A_i and B_i are the amplitude of the recovery profile and standard profile. $\overline{A}$ and $\overline{B}$ are the mean value of the amplitude of the standard profile [9] and recovery profile [10].

Set the anode high voltage of the X-ray source as 25 kV and the filament current as 2.2A. Select PSR B0531+21's profile. Use SDD detector to record the photon arrival information. The experiment time interval is 7200 s. Use epoch folding method to recover profile, and compute the comparability according Eq. (1) (Fig. 10).

From the results, the system can simulate the profile and the similarity is high. When the collection time is 7200 s, the similarity is as high as 99.98%.

2. **Detector time delay experiment**

Because of the principle of the detector and the transmission of the electronic circuits, the detector's response exit time delay. In order to get the calibration

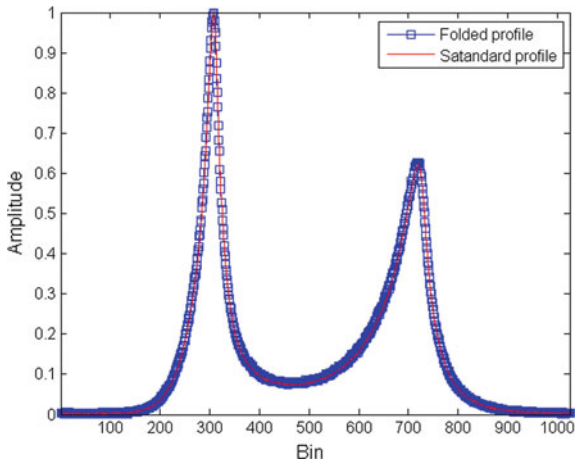


Fig. 10 Similarity experiment of static signal

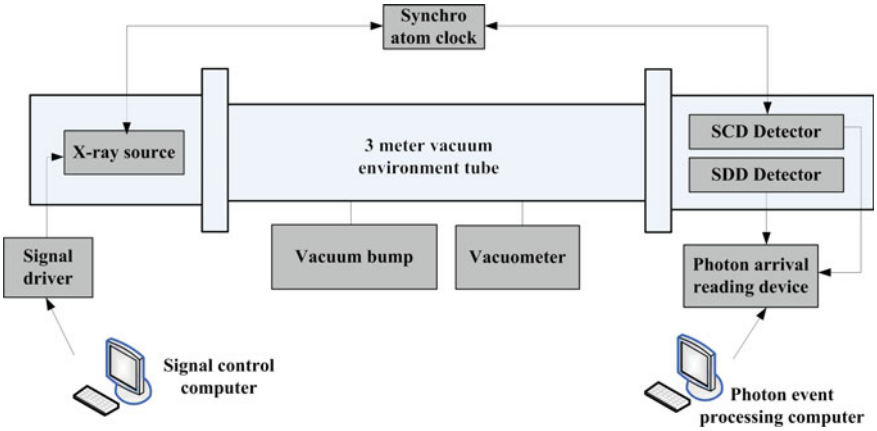


Fig. 11 Setup of the time delay calibration testbed

parameter, this time delay should be calibrated. The setup of the testbed is shown in Fig. 11.

Set the output profile as PSR B0531+21, SCD and SDD Detector collect photon data at one time for one hour. Obtain two profile by epoch folding, as Fig. 12 shows.

By comparison of the two profiles, the time delay of SCD detector is 11.3525 ms.

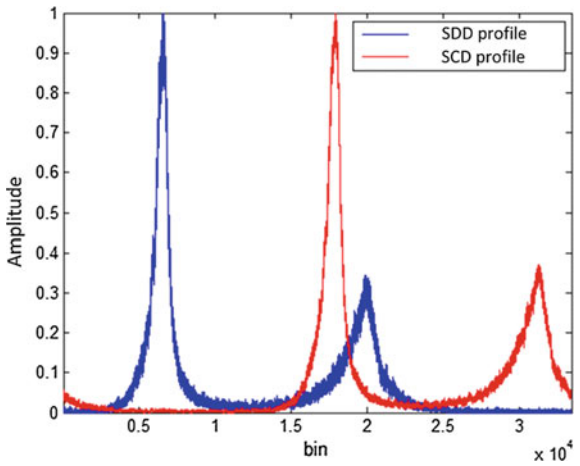


Fig. 12 Time delay of the SCD detector

Table 1 Test results of single energy experiments

Character energy (keV)	FWHM (eV)
1.49	148
4.51	190
5.41	166
6.40	158
8.05	187
9.89	178
15.77	192

3.3 Energy Spectrum Experiment

1. **Single energy X-ray experiment**

Set the anode high voltage of the X-ray source as 25 kV and the filament current as 2.2A and the grid as 0 V. Use SDD Detector to collect photons. Get the energy spectrum, character energy (Kα) and FWHM by the data processing software (Table 1).

2. **Complex energy X-ray experiment**

Set the anode high voltage of the X-ray source as 25 kV and the filament current as 2.2A and the grid as 0 V. Use Mo target to generate X-ray. The working time of the system is one hour.

The complex energy spectrum of Mo target is shown in Fig. 13. By comparison with the PSR B0531+21's energy spectrum curve, the complex energy spectrum is similar to it. Therefore, the complex energy spectrum can support the research of energy spectrum construction of pulsars, test of detector's energy resolution and so on.

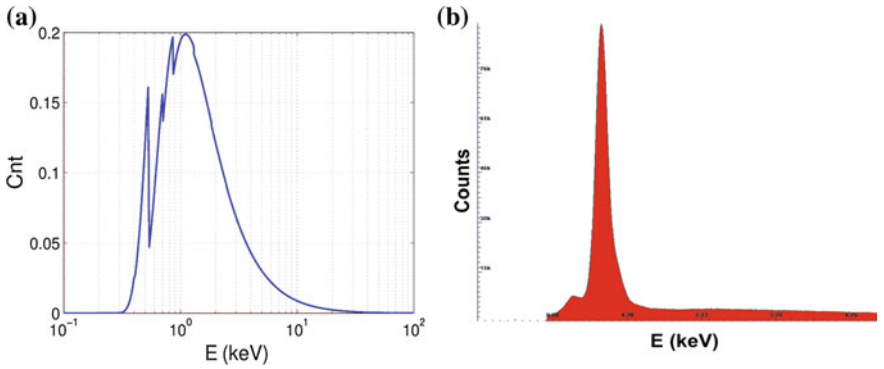


Fig. 13 Simulation of pulsar's energy spectrum. **a** Energy spectrum of PSR B0531+21. **b** The simulation energy spectrum

4 Conclusions

At present, X-ray Pulsar-based Navigation is on the step of technology verification. It is very expensive to use the spacecraft to verify. However, the software simulation cannot reflect the device response accurately. Use the hardware in-loop system can play to their strengths. This paper use "Tianshu II" developed by National University of Defense Technology and Xi'an institute of optics and precision mechanics of CAS to do part experiments. The experiments test the dynamic signal and static signal simulation, calibration of X-ray detector and energy spectrum simulation. The results show that the simulation precision of "Tianshu II" system is high, the system can support the research of X-ray Pulsar-based Navigation technology.

References

1. Shuai P, Ming LI, Chen Shaolong, Huang Z (2009) X ray pulse principle and method of satellite navigation system. China Aerospace Press, Beijing
2. Wang Y, Sun S, Li L (2014) Adaptively robust unscented Kalman filter for tracking a maneuvering vehicle. J Guidance Control Dyn 37(5):1696–1701
3. Wang Y, Zheng W, An X (2013) XNAV/CNS integrated navigation based on improved kinematic and static filter. J Navig 66:899–918
4. Lowe BG, Holland AD, Hutchinson IB (2001) The swept charge device, a novel CCD-based EDX detector: first results. Nucl Instrum Methods Phys Res A 458:568–579
5. Fedotv MG (2000) CCD detectors for X-ray synchrotron radiation application. Nucl Instrum Method Phys Res A 448:192–195
6. Huijun HU (2011) Research for the X ray photon counting detector on pulsar navigation. Chinese Academy of Sciences
7. Lizhi S, Baosheng Z, Jianjun W (2012) Research of X-ray pulsar navigation simulation source. Acta Phys Sinica 62(12)

8. Ge M (2011) The X-ray emission of pulsars. Chinese Academy of Sciences
9. Ge M, Lu F, Qu JL, Zheng S, Chen Y, Han D (2012) X-ray phase-resolved spectroscopy of PSRs B0531 + 21, B1509-58, and B0540-69 with RXTE. Astrophys J Suppl Ser 199(2):8–10
10. Emadzadeh A, Speyer J (2012) X ray pulsar navigation. National Defence Industry Press, Beijing

Magnetic Field Based Indoor Pedestrian Positioning Using Self-contained Sensors

Xiang Mu, Jiuchao Qian, Changqing Xu, Ling Pei, Peilin Liu and Wenxian Yu

Abstract The geomagnetic field exists everywhere and has been demonstrated to be local disturbed and stable over time. In this paper, we present an approach for indoor positioning which does not rely upon any additional support equipment, using magnetic fingerprint collected by magnetic field sensors in smartphone and walking information of user measured by self-contained inertial sensors. In order to improve the practicability of our positioning method, we choose the Z-axis magnetic of world coordinate system as feature information, which can reach no restrain of phone's gesture and orientation when pedestrians walk. Moreover, to improve the discernibility of geomagnetic information, we measure the geomagnetic signals as a sequence, and we match the geomagnetic signals with the established fingerprint using dynamic time warping (DTW) algorithm. Meanwhile, a hidden Markov model (HMM) based positioning process is adopted to strengthen the robustness of the algorithm. Extensive field tests have been conducted in an office building to verify the performance of proposed algorithm. Test results show that the proposed algorithm can achieve less than 1.11 ms of mean error and 2.13 ms of 95th percentile error.

Keywords Indoor pedestrian positioning · Magnetic fingerprint · Hidden markov model · Self-contained sensors

1 Introduction

With rapid development of Global Navigation Satellite System (GNSS), Location-Based Services (LBS) has been widely applied in outdoor environment. However, GNSS signals are degraded or denied in indoor areas [1]. In recent years positioning

X. Mu · J. Qian (✉) · C. Xu · L. Pei · P. Liu · W. Yu
Shanghai Key Laboratory of Navigation and Location-Based Services, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: jcqian@sjtu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_46

in indoor environment has become an essential requirement for LBS applications, such as emergency rescue, health care and commercial information push services etc. Therefore, the research on positioning systems for indoor environment has been one of the most important research hotspots.

Recently indoor position techniques based on wireless signals such as WIFI [2], UWB [3], ZigBee [4] and Frequency Modulation (FM) [5] have been investigated. But when the ambient signals are subject to occasional change, especially in emergency rescue missions where electric power is failed these approaches will run out. Therefore, an indoor positioning method that is infrastructure free is demanded. As we know, the geomagnetic field exists everywhere and has been demonstrated to be stable over time [6]. Stable feature of geomagnetic field and the geomagnetic information can be used to provide accurate indoor positioning without depending on any equipment.

In this paper, we propose an approach for indoor positioning which does not rely upon any additional support equipment, using magnetic fingerprint collected by self-contained magnetic sensors.

To ameliorate the accuracy of positioning, we adopted a sequence measured by magnetic sensor to match with established fingerprint rather than single measurement vectors [7, 8]. The positioning process can be considered as a hidden Markov process, and at the same time matching the sequences with the constructed fingerprint enables to get more precise transition probabilities in HMM, which improves accuracy of positioning.

For purpose of verifying the high performance of our proposed methods, we carried out a series of field tests in our office building. Upon the test results, we found that our indoor positioning system has high performance without any extra infrastructure, and it is convenient to deploy the system on building at relatively high speed.

2 Positioning Method

In this work, we provide a positioning method using smartphones in indoor environments. The positioning method has two phases: offline phase and online phase. In the offline phase, we collect the geomagnetic data to construct the fingerprint database. In the online phase, we match the geomagnetic sequence measured by smartphone online with the geomagnetic fingerprint database to locate the user's position.

2.1 In the Offline Phase

Many researches have demonstrated the fact that the geomagnetic field is stable in indoor environments [9, 10], but single geomagnetic information provides low

discernible information for positioning. Hence, we choose the sequence of geomagnetic as positioning information, rather than acquire the information of single point, which improves discernible for effective positioning.

The magnetometer in smartphone can obtain three-axis data of the raw geomagnetic signal m_x, m_y and m_z in the smartphone's coordinate system, which is represented by the $\{x, y, z\}$, thus the intensity of raw geomagnetic signal measured by smartphone in three axis related with the gesture and orientation of the smartphone. However, the orientation of smartphone should not be added with any restriction when pedestrian walks naturally in daily life. Therefore, it is significant to estimate the direction of the smartphone and realize the transformation from the raw geomagnetic signal sourced from the magnetometer to the world's coordinate system, which represented by the $\{X, Y, Z\}$, where the X axis is pointing to the geomagnetic north pole, the Z axis is pointing to the opposite direction of gravity and the Y axis is orthogonal to both the X axis and the Z axis.

In practice, the gesture information of the smartphone including the roll, pitch, and yaw angles of the body frame with respect to the navigation frame are unreliable, due to the user's motion and the quickly diverging errors in the low cost inertial sensors of smartphone [11, 12]. In indoor area it is hard to estimate the north reliably by the magnetic sensors, since this process needs the measurement of undistorted geomagnetic information. However, the three-axis accelerometer is reliable to figure out the direction of gravity [13, 14], then having known the direction of gravity we can calculate the intensity of geomagnetic on the Z axis of the world reference through a transform formula as [15]:

$$m_Z = \mathbf{m} \odot (-\mathbf{g}/|\mathbf{g}|) \tag{1}$$

where $\odot$ is the vector inner produce operation, m_Z is the geomagnetic intensity of the Z axis of world coordinate system. The measurements of geomagnetic along the Z axis are independent of user's heading.

To demonstrate the stability of m_Z, we collected the geomagnetic information along a path in our office building at different time shown as Fig. 1. Although, the magnitudes of three curves are different, they have similar variation trend. Hence we adopt the Z axis geomagnetic as the feature of positioning.

For the purpose of constructing the database, we divide the path of office room into segments by the cross of path net. We walked straight along every segment at a constant speed with a smartphone held on hand. In order to establish more accurate database of geomagnetic field, we collected the geomagnetic information in one path dozens of times.

Magnetometers in different smartphones, even one magnetometer of the same device in different time also leads to different offsets. To reduce the influence of these offsets, we calculate the mean of each geomagnetic sequence independently and deduct the mean from points in each sequence.

In the online phase, we deduct the mean of geomagnetic sequence in each point as well, thus sequences measurements from the magnetometer all have zero mean.

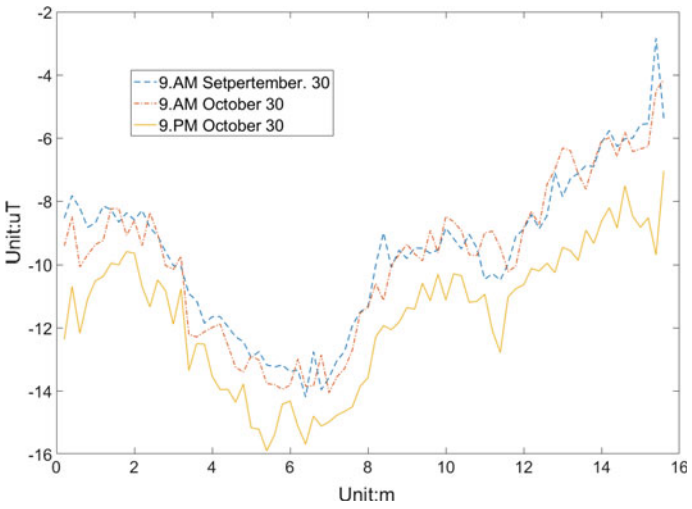


Fig. 1 The geomagnetic intensity of Z-axis measured in different time

Due to we subtract the mean value from the collected geomagnetic intensity online, the uniqueness of geomagnetic information is reduced.

After getting sequences information of geomagnetic, which is read from magnetometer and transformed along Z axis of the world, removing the mean of the sequence as well, for one segment, we divide this sequence into smaller segment averagely by the counts of steps on this path, then we save the geomagnetic signal as $\mathbf{DT} = \{\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_w\}$, where w is the number of total paths in the office room, and $\mathbf{D}_i = \{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_M\}$ $i = 1, 2, \ldots, w$, where M is the number of dataset in one path and the $\mathbf{d}_M$ is one sequence in $\mathbf{D}_i$, $\mathbf{d}_i = \{d_1, d_2, \ldots, d_n\}$, where d is the sequence of one step of ith path. Constant speed is designed in the process of geomagnetic signal collection, so that we can obtain the position of every small sequence of geomagnetic.

2.2 *In the Online Phase*

In the online phase, we can detect pedestrian steps using accelerometer at first, and then calculate the time interval between two neighboring steps. In the time interval, smartphone will get the sequence of geomagnetic by recording the data of magnetometer. The sequence can be compared with the collected data which were saved in the database. As the existence of imparity in step length and speed of different people, we use the dynamic time warping (DTW) algorithm for comparing the database with online gained data.

Since the vertical acceleration fluctuates periodically due to the human motion, we use the detection algorithm [16] to detect the walking steps. Moreover, as the

positioning process has an observed state and a hidden state. The observed state is the observation of geomagnetic intensity. The hidden state is the position of the geomagnetic. Hence, we can consider the positioning process as hidden Markov model (HMM).

In the tracking phase, the mathematical model of HMM can be represented as $\lambda = \{L, R, A, B\}$, with the following variables.

1. $L = \{l_1, l_2, \ldots, l_i, \ldots\}$ is the set of hidden states, where the $l_i = (x_i, y_i)$ responds the true position of the ith point.
2. $\mathrm{O}_i = \{o_{i-N+1}, o_{i-N+2}, \ldots, o_i\}$ is the set of observations, where o_i is the sequence of geomagnetic signal on the ith point, and N is the length of observations.
3. $A = \{a_{i,1}, a_{i,2}, \ldots a_{i,N}\}$ is the set of transition probabilities from state q_i to state q_j where the $a_{i,j}$=$P(q_i = l_{i,j} | q_{i-1} = l_{i-1})$, q_i is the ith hidden state of HMM, N is the total number of possible positions of l_i.
4. $B = \{b_{i,1}, b_{i,2}, \ldots, b_{i,N}\}$ is the set of emission probabilities of l_i, where the $b_{i,j}$ is the emission probabilities of $o_{i,j}$ at $l_{i,j}$.

2.2.1 Online Matching Using DTW

As the step length and speed of the pedestrian are different in different time, the measurement of distance needs to be defined between the database and online reading. The measured distance should be invariant to different sequence time shifts and time scales.

The algorithm of dynamic time warping is to compress or stretch the time axis of sequences to achieve a better alignment. In general, the geomagnetic sequence of online reading for one segment $o_k = \{m'_{Z1}, m'_{Z2}, \ldots m'_{Zs}\}$ and the geomagnetic sequence of database for one segment $d_k = \{m_{Z1}, m_{Z2}, \ldots m_{Zt}\}$ have different length.

The warping path can be represented as $w = w(1), w(2), \ldots w(n)$, where $w(n) = [m'_{Zi}, m_{Zj}]$ is the set of matched samples, where i and j is the corresponding point of time axis of two geomagnetic sequences. The function of warping is to minimize the overall cost, the function given by

$$Dis(m'_Z(n), m_Z(n)) = \sum_{n=1}^{N} \delta(w(n)) \tag{2}$$

where $\delta(w(n))$ is the squared distance between the sample points given by

$$\delta(w(n)) = (m'_Z(n) - m_Z(n)) \tag{3}$$

By the dynamic time warping algorithm, we can calculate the distance between the recent geomagnetic sequence read from magnetometer with the sequence of geomagnetic in the database [17].

2.2.2 Positioning Using HMM

The process of positioning contains two procedures, the first one is to estimate the path that a pedestrian is walking on when the pedestrian starts walking or she (or he) turns the orientation, and the other one is to calculate the point she is standing at. Hence, in the first procedure, the question transform to estimate the geomagnetic sequence with the sequence in the database which has the best matching.

$$\widehat{\mathbf{D}}_i = \arg\max_{\mathbf{D}_i} P(\mathbf{O}_i|\mathbf{D}_i) \tag{4}$$

By the Bayes formula, in Eq. 4 $P(\mathbf{O}_i|\mathbf{D}_i) = P(\mathbf{O}_i, \mathbf{D}_i)$, where $\mathbf{D}_i$ is the geomagnetic sequence for ith path built in the geomagnetic database. As the steps of pedestrian can be considered as consecutive, we can calculate the $P(\mathbf{O}_i, \mathbf{D}_i)$ using the principle of HMM that is defined as:

$$P(\mathbf{O}_i, \mathbf{D}_i) = P(d_1) \prod_{j=2}^{N} P(d_j|d_{j-1}) \prod_{j=1}^{N} P(o_j|d_j) \tag{5}$$

where d_j is the geomagnetic sequence of jth step of the ith path in the database, and o_j is the reading of magnetometer for the jth step. Taking the logarithm for Eq. 5, we can rewrite Eq. 4 as:

$$\widehat{\mathbf{D}}_i = \arg\max_{\mathbf{D}_i} \left[\log(P(d_1)) + \sum_{j=2}^{N} P(d_j|d_{j-1}) + \sum_{j=1}^{N} P(o_j|d_j) \right] \tag{6}$$

For the consecutive steps, we assume the transition probabilities $P(d_j|d_{j-1})=1$, and the observation probabilities are the DTW distance between observed geomagnetic sequence o_j and the geomagnetic sequence d_j in the database. Therefore, the emission probabilities are given by:

$$P(o_j|d_j) = NOR \times \exp(-Dis(o_j, d_j)) \tag{7}$$

where NOR is a normalized constant. Then the $\widehat{\mathbf{D}}_i$ can be calculated by using the Viterbi algorithm, where we set a trellis including all possible states of the path. We calculate the DTW distance of the ith step and the database with all possible states, record the possibility value and the position of best matching sequence in database

of every step. Then we find the most possible sequence of steps, and the position of the steps sequence.

In order to acquire more accurate position, when the position of path has been determined, we use the neighbor steps sequence to localize the pedestrian.

Assuming the l_i is the position of the ith step, for each l_i, we match the geomagnetic sequence in the database with the observation sequence to compute the emission probabilities by:

$$\begin{aligned} b_{i,j} &= p(o_i|l_i, d_j) \\ &= NOR \times \exp(-Dis(o_i, d_j)) \end{aligned} \tag{8}$$

To gain the transition probability $a_{i,j}$, we compute the transition probability of the position of $(i-1)$th step to the ith step.

$$P(l_j|l_{i-1}) = \begin{cases} 1 & \text{if } D(l_j, l_{i-1}) < thrs \\ 0 & otherwise \end{cases} \tag{9}$$

where $D(l_i, l_j)$ is the Euclidean distance of the ith step and the jth step. And transition probability, $a_{i,j}$ can be calculated by weighting $b_{i,j}$ and the probability $P(l_j|l_{i-1})$

$$a_{i,j} = b_{i,j} \times P(l_j|l_{i-1}) \tag{10}$$

Then the position of ith step can be calculated as:

$$l_i = \max_{l_{i,j}} \arg(l_{i,j} \times a_{i,j}) \tag{11}$$

3 Experiment and Result

In this section, we will provide details in the experimental realization. The deployment of experimental system would be stated and the positioning performance of our system would be demonstrated in different conditions.

In order to demonstrate the accuracy and robustness of our positioning method, we conduct experiments in two different paths set up in our office room. We can verify that the method is able to provide accurate results on the long distance condition by walking on the circular path. To validate this method is adapted to the complex condition, we perform experiments in a zigzag path. The robustness of the method was demonstrated by walking on two different paths with swinging the smartphone casually. As Figs. 2 and 3 show, route in Fig. 2 is a circle path round the office room, and route in Fig. 3 is a zigzag path in the office room.

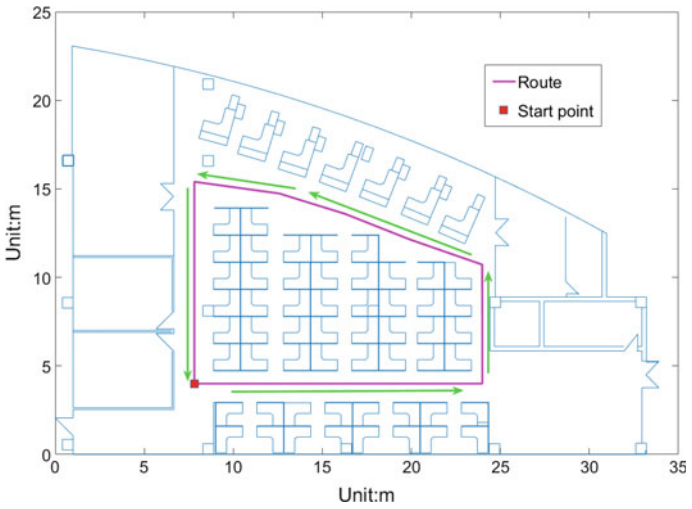


Fig. 2 The route of circular path round the office room, and the arrow is orientation of the pedestrian in testing

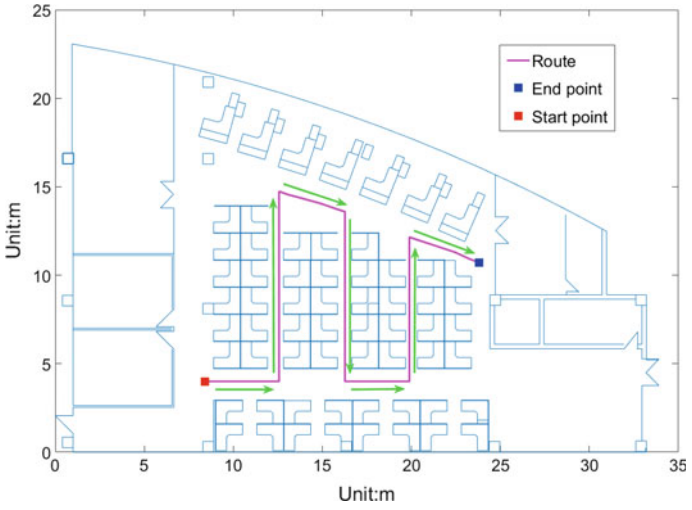


Fig. 3 The zigzag path round the office room, and the *arrow* is orientation of the pedestrian in testing

Then we compared geomagnetic based positioning method with Pedestrian Dead Reckoning (PDR) [18], which is another positioning method basing on self-contained sensor [19, 20].

In addition, all experiments were conducted in a typical office building. We developed an android application to measure the geomagnetic intensity in 3-axis

and recorded the measurements on the smartphone. The smartphone we used is Huawei Honor 8 with Android 6.0 operating system.

3.1 Geomagnetic Database Construction

In the construction process of geomagnetic database, we divided the path of the office room into small segments firstly, as Fig. 4 shows the segment of each path is about 4 m.

We trained the database on the path segments marked on Fig. 4. The geomagnetic signal of every path segment had been measured 30 times, and saved in the database with the tags that are marked beside the path.

As for the ground truth, we set up many landmarks on the floor, and acquired their true position in advance. We recorded the true position when pedestrian walk passing the landmark. Therefore, the error can be defined as the Euclidean distance between the result of location based on geomagnetic positioning and the ground truth.

3.2 Positioning Performance

On the circular route, we walked around the circular route three times, and total length of the trajectory is about 150 m. On the zigzag route, total length of the trajectory is about 48 m.

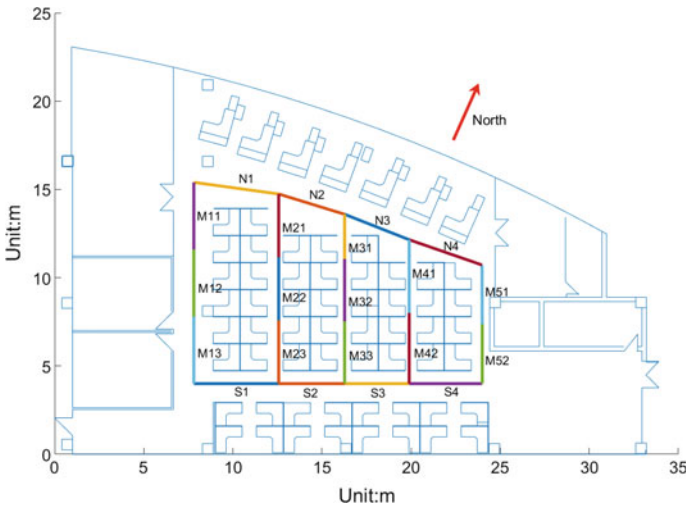


Fig. 4 All of paths are divided into 21 segments, and every segment has its own tag. In the train phase, the user holds the smartphone to measure the geomagnetic signal repeatedly

When we walked holding the smartphone stably, the results of positioning by two different methods are shown in Figs. 5 and 6.

The red lines represent pedestrian's trajectory and blue points represent ground truth that is demarcated by using laser rangefinder.

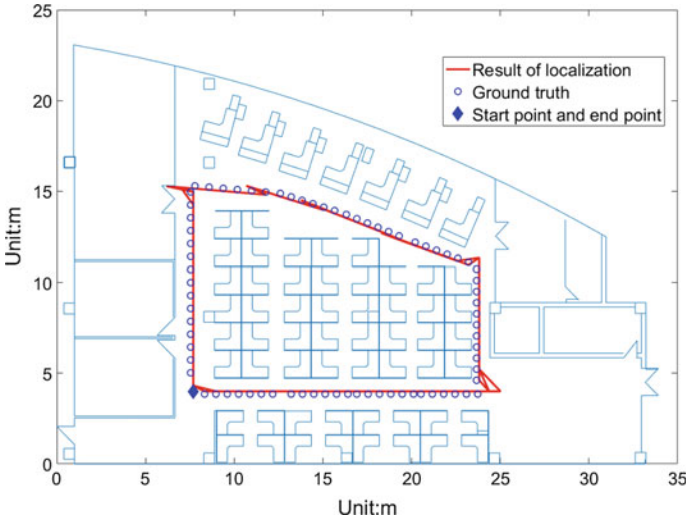


Fig. 5 The result of positioning for the route of circular path round the office room

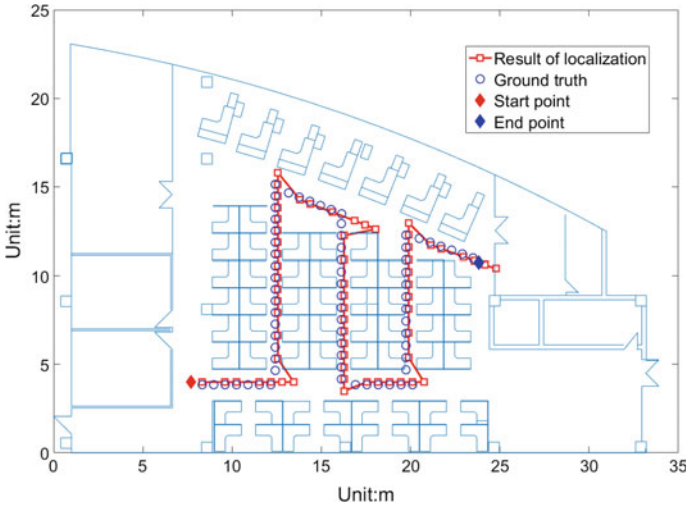


Fig. 6 The result of positioning for the zigzag path in the office room

In the circle path, start point and the end point are at the same point, denoted by a blue diamond. In the zigzag path, red diamond and blue diamond denote start point and end point respectively.

The result indicates that the positioning of our algorithm is reliable and accurate.

For the verification of the stability of our positioning method without orientation restraint of smartphone, we walked with swinging the smartphone casually.

As shown in Fig. 7, the user was swinging the smartphone when he walked on the path.

Fig. 7 The swinging gesture during walking

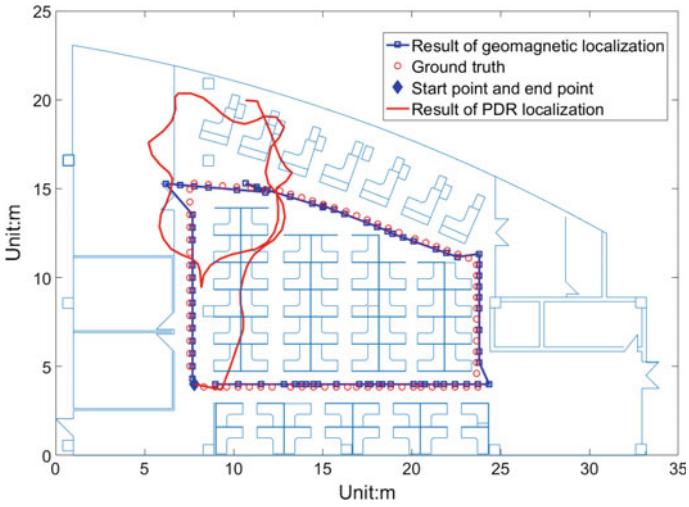


Fig. 8 The result of positioning for the circular path round the office room with swinging the smartphone continuously

Comparing with PDR, we made a series of experiments with swinging the smartphone. We used the geomagnetic positioning method and PDR method respectively on the two paths at the same time. In order to demonstrate the difference of performance between geomagnetic based positioning method and PDR method, the circular route is changed to one circle, as the positioning result shown in Figs. 8 and 9.

The red lines indicate the estimated trajectory using the PDR algorithm. The blue lines represent pedestrian's trajectory estimated by the geomagnetic based positioning method, and red points represent ground truth that is demarcated by using laser rangefinder.

From Figs. 8 and 9, we can see that method by geomagnetic made a much better performance than PDR. As the geomagnetic based positioning can update the position, which relies on the stable geomagnetic information, the geomagnetic based positioning method is able to provide accurate and stable positioning, while the PDR method accumulates errors leading to failed positioning.

The uniqueness of geomagnetic information is reduced due to the online subtraction of the mean value from the collected geomagnetic intensity. Hence, in neighboring steps, the geomagnetic information exists similarity, which causes the mismatching. But when the pedestrian walks through a path, the positioning method will judge the next path by matching the geomagnetic information with the neighbor path. Hence, the positioning can be amended.

The detailed positioning results are listed in Table 1. From the results, we can find that the mean error is no more than 1.1 m, and the 95th percentile error is less than 2.13 m. In Table 1, we also draw a conclusion that the error of positioning is not related to swinging smartphone when the pedestrian walking.

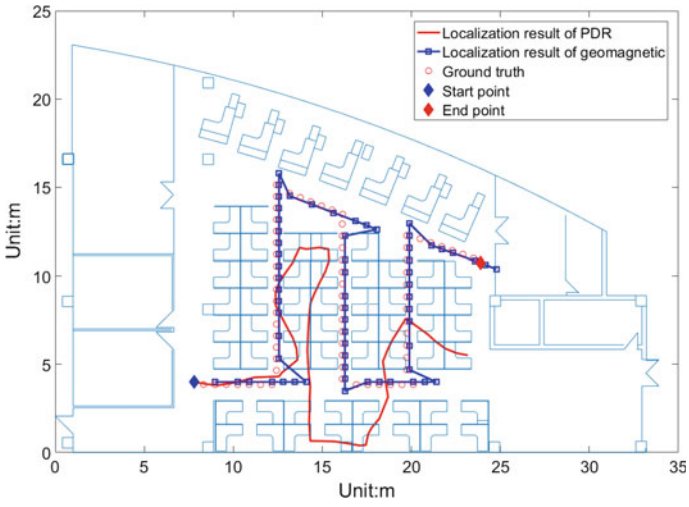


Fig. 9 The result of positioning for the zigzag path in the office room using two different positioning methods

Table 1 The positioning error of proposed method

Scenario	Mean error (m)	95th percentile error (m)
Hold the smartphone stably on circular route	0.91	2.13
Hold the smartphone stably on zigzag route	1.06	1.90
Swinging smartphone casually on circular route	0.99	2.13
Swinging smartphone casually on zigzag route	1.10	1.96

4 Conclusion

In this paper, we present an indoor positioning algorithm based on the geomagnetic trajectory matching methods. It leverages the stable geomagnetic information without deploying any additional infrastructure. Experiments in two different indoor conditions, including a circular route and a zigzag route, demonstrate that the proposed self-positioning algorithm can give high-accuracy positioning performance.

References

1. Liu D, Pei L, Qian J, et al (2016) Simplified ellipsoid fitting-based magnetometer calibration for pedestrian dead reckoning. In: China satellite navigation conference (CSNC) 2016 proceedings: Volume II
2. Liu H, Darabi H, Banerjee P, Liu J (2007) Survey of wireless indoor positioning techniques and systems. IEEE Trans Syst Man Cybern C Appl Rev 37(6):1067–1080
3. Kok M, Hol JD, Schön TB (2015) Indoor positioning using ultrawideband and inertial measurements. IEEE Trans Veh. Technol 64(4):1293–1303
4. Tomic S, Beko M, Dinis R (2015) RSS-based positioning in wireless sensor networks using convex relaxation: noncooperative and cooperative schemes. IEEE Trans Veh Technol 64 (5):2037–2050
5. Chen Y, Lymberopoulos D, Liu J et al (2013) Indoor positioning using FM signals. IEEE Trans Mobile Comput 12(12):1502–1517
6. Haverinen J, Kemppainen A (2009) Global indoor self-positioning based on the ambient magnetic field. Robot Auton Syst 57(10):1028–1035
7. Shahidi S, Valaee S (2015) GIPSy: geomagnetic indoor positioning system for smartphones. International conference on indoor positioning and indoor navigation. IEEE
8. Ma Y, Dou Z, Jiang Q et al (2016) Basmag: an optimized HMM based positioning system using backward sequences matching algorithm exploiting geomagnetic information. 16(20):7472–7482
9. Wang Q, Luo H, Zhao F, Shao W (2016) An indoor self-localization algorithm using the calibration of the online magnetic fingerprints and indoor landmarks. In: 2016 international conference on indoor positioning and indoor navigation (IPIN)
10. Haverinen J, Kemppainen A (2009) Global indoor self-localization based on the ambient magnetic field. Robot Auton Syst 57(10):1028–1035
11. Qian J, Pei L, Ma J, Ying R, Liu P (2015) Vector graph assisted pedestrian dead reckoning using an unconstrained smartphone. Sensors 15(3):5032–5057

12. Qian J, Ma J, Ying R, Liu P (2013) RPNOS: reliable pedestrian navigation on a smartphone. In: Geo-informatics in resource management and sustainable ecosystem. Springer, Berlin Heidelberg
13. Li B, Gallagher T, Dempster AG, Rizos C (2012) How feasible is the use of magnetic field alone for indoor positioning. In: International conference on indoor positioning and indoor navigation, pp 1–9
14. Yun X, Bachmann ER, Mcghee RB (2008) A simplified quaternion-based algorithm for orientation estimation from earth gravity and magnetic field measurements. IEEE Trans Instrum Meas 57(3):638–650
15. Kim B, Kong SH (2016) A novel indoor positioning technique using magnetic fingerprint difference. IEEE Trans Instrum Meas 1–11
16. Qian J, Ma J, Ying R, Liu P, Pei L (2013) An improved indoor localizaion method using smartphone inertial sensors. In: IPIN, pp 1–7
17. Subbu KP, Gozick B, Dantu R (2013) Locateme: magnetic-fields-based indoor localization using smartphones. ACM Trans Intell Syst Technol 4(4):99–103
18. Chen R, Pei L, Chen Y (2011) A smart phone based PDR solution for indoor navigation. In: Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2011), pp 1404–1408
19. Qian J, Pei L, Ying R, Chen X, Zou D, Liu P et al (2014) Continuous motion recognition for natural pedestrian dead reckoning using smartphone sensors. In: Proceedings of international technical meeting of the satellite division of the Institute of Navigation
20. Qian J, Pei L, Zou D, Liu P (2015) Optical flow-based gait modeling algorithm for pedestrian navigation using smartphone sensors. Sens J IEEE 15(12)6797–6804. doi:10.1109/JSEN.2015.2464696

3D Indoor Layered Localization of Multi-information Fusion Via Intelligent Terminal

Leinan He and Hu Lu

Abstract Aiming at the problem that that pedestrian can only get low accuracy and plane coordinate of current indoor localization from intelligent terminal, a layered indoor 3D positioning technology based on WiFi positioning, pedestrian dead reckoning (PDR) and architecture information is proposed. In order to estimate height changes of pedestrian, the two pedestrian behavior patterns of horizontal walking and up-down stairs can be distinguished by detecting the pitch-rate mutation. Height changes are got by using the effective steps and stairs information. The equations of the two different behavior patterns are constructed and three-dimensional PDR is realized. A particle filter algorithm combining multi navigation information is proposed to eliminate the cumulative errors of PDR. Fusing the architecture information and behavior patterns into the weight calculation and the sampling of the particle, the localization precision is effectively improved. The results show that the average error of the horizontal location is 2.1 m and the average error of the height estimation is 0.8 m. The accuracy could satisfy the demand of pedestrian location service using individual intelligent terminal.

Keywords 3D indoor localization · 3D PDR · WiFi · Pitch rate · Particle filter

1 Introduction

Intelligent terminal has become an indispensable tool in people's daily life and travelling. The location technology of mobile terminal plays a key role in location-based services [1]. At present, the researches and implementations of intelligent terminal location are mostly focused on two-dimensional plane. However, in the indoor environment, applications such as equipment management, fire detection, logistics management, store location of large-scale shopping malls

L. He · H. Lu (✉)
School of Information and Navigation, Air Force Engineering University, Xi'an 710077, China
e-mail: 10956087@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_47

often need to know the horizontal position and the knowledge about the location of the floor information, so the positioning results have meaning.

There are a large number of positioning technologies on the intelligent terminals which can be divided into three categories: the GPS navigation as the representative of the satellite positioning technology, dead reckoning (DR) based on self-contained sensor and wireless local area network (WLAN) positioning technology. In the indoor environment, GPS only has 4.5% of the time to obtain a reliable GPS signals because of the building block [2]. It can't be applied to indoor positioning scene. WLAN-based positioning technology is the research hotspot of indoor positioning at present, because the WLAN access point is widely deployed. It achieved good positioning effects in the two-dimensional positioning. It can be divided into distance-dependent positioning technology (based on RSSI, TOA and AOA) and distance-independent (such as centroid algorithm, DV-Hop algorithm positioning technology) [3]. Researchers further proposed indoor three-dimensional positioning technology based on the mature two-dimensional positioning technologies. However, the inherent characteristic of WLAN localization is not sensitive to the three-dimensional position information [4, 5], which can't meet the requirements of the application. In view of this, some scholars put forward three-dimensional positioning method based on UWB, signal receiving phase and vision [6–8], and achieved high positioning accuracy. However, these methods need to increase the corresponding intelligent terminal hardware, and also bring an additional burden to the wireless network and higher cost. PDR positioning technology refers to the use of intelligent terminals embedded micro-electro-mechanical (MEMS) sensors (usually including inertial sensors, magnetic compass, barometer, etc.). Through off-line training, stride detection, filtering and determine direction, one can achieve the basic pedestrian location calculation, with high accuracy in a short time and the advantage of a continuous track. But due to the error drifts of the sensor, the accumulated errors will result in the persistence of the positioning result. It needs to be corrected by other positioning techniques [9], and traditional PDR can only achieve two-dimensional level of positioning. Currently the researches of three-dimensional PDR is relatively limited. In [10], based on the two-dimensional PDR, three-dimensional position estimation is achieved using altitude conversion relationship between pressure and the height. But the indoor atmospheric pressure is unstable, the temperature and the environment have a greater impact on it. Collected data at the same location and different moments may be very different. In [11], a three-axis MARG sensor equipped with a smart phone is used to filter the sensor data and reduce the estimation error to detect pedestrian frequency, finally, the three-dimensional track estimation and location is achieved based on the EKF, but the height estimation error is large. Zheng et al. [12] used low-power MEMS device to achieve a three-dimensional pedestrian walking three-dimensional positioning system based on KF, requiring the sensor in accordance with the human ankle. In [13], the EKF fusion barometric altimeter and accelerometer information are used to improve the accuracy of altitude estimation, but it is susceptible to environmental and cumulative errors. In [14], the UKF-based WiFi/PDR localization algorithm is proposed to achieve high-precision two-dimensional localization

for 3D PDR error accumulation. 2D positioning results are displayed using a visual platform which has been inputted with three-dimensional geographic information, achieving visual three-dimensional positioning in the same floor, but it can't perceive the floor changes. Banerjee et al. [15] analyse the principle of the three-dimensional positioning based on the WiFi of the smartphone, and point out that the existing method has higher level precision and lower vertical precision and the three-dimensional positioning of fusing barometric altimeter and WiFi can improve the positioning accuracy of 25%. The performance is relatively limited.

In this paper, we use pitch angle velocity detection method to distinguish the level of walking, walking up and down stairs 2 types of movement patterns, and extract the different types of effective step. Based on Each effectively step of the different walking motion patterns, solving PDR three-dimensional positioning. A WiFi/PDR localization algorithm based on particle filter is proposed, which using the absolute position information of the WiFi positioning to eliminate error of the PDR. The indoor three-dimensional positioning can meet the needs of building floors to distinguish on intelligent terminals. Finally, the experiment is carried out in the WiFi network environment.

2 3D PDR Based on Pitch Rate Detection

The traditional PDR used the peak detection method, zero velocity update, Stand Phase detection method and other methods of acceleration signal detection, identify the step frequency and step size, combined with the heading outputted by sensor to calculate pedestrian location, and achieved a good two-dimensional plane positioning effect. However, only a single two-dimensional horizontal motion pattern is usually studied, and it is not possible to estimate the height change effectively. This paper we present a three-dimensional localization method of PDR based on pitch angle velocity detection, use the acceleration peak value to detect the effective step, detect the change of the pitch angular velocity to distinguish the pedestrian behavior pattern during the duration of the effective pitch, determine the relative elevation via the geometric information of building stairs.

2.1 Pitch Rate Detection to Distinguish Two Behavior Pattern

First, an effective step is detected by the peak detection algorithm [11], and then judged the continuous effective steps whether or not occur during the floor shift. It's the up and down stairs mode when satisfying the above-mentioned, otherwise it belongs to the horizontal movement mode. It is proposed to distinguish the beginning and the end of pedestrian stairs based on pitch angle velocity detection.

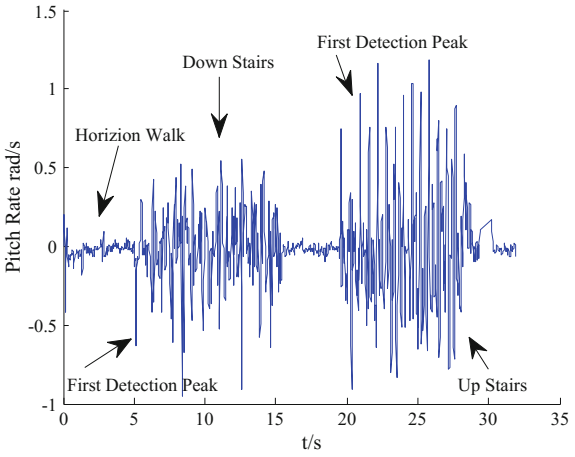


Fig. 1 Pitch rate

As the pedestrian walking gait may have an impact on the pitch angular velocity, three successive valid steps are selected to identify the motion pattern when taking into account that the number of steps between the floors. If a sudden change in the pitch angle of the three successively effective steps is detected, the mode is regarded as a staircase. The pitch angle speeds of pedestrian in different patterns show as Fig. 1.

The pitch angular velocity detection method can be expressed as: If the pitch angular velocity satisfies:

$$|\varphi_k| \geq \phi, \quad k = 1, 2, 3 \tag{1}$$

where φ_k is pitch rate, k is effective step, ϕ is pitch rate threshold according to the geometric information of the stairs. It means that three consecutive effective steps occur in the up and down stairs. There are two essential conditions, first, the detection of a single effective step pitch angle velocity is greater than the threshold, second, the effective step must be a one of continuous three, otherwise it will immediately switch to horizontal walking mode. Finally, determine the upstairs or downstairs according to the detection of the mutation of the pitch angular velocity of the first detection peak, it is upstairs if the first detection peak is positive, otherwise it's downstairs.

2.2 Solution of Pedestrian 3D Position

In this paper, a three-dimensional PDR localization based on pitch-angle velocity detection is proposed. After the detection of pitch angle velocity is completed, different types of behavior patterns can be distinguished. According to the characteristics of each mode, the pedestrian three-dimensional position estimation is

carried out for each detected effective step. When the behavior pattern is determined to be horizontal, the height remains unchanged and only the horizontal position is estimated. When it is determined to up and down stairs, the horizontal step length takes the width of staircase, and the vertical step length takes the height of staircase. The gyro and magnetic compass are used to estimate the current heading [11] and finally the three-dimensional indoor positioning is completed. Figure 2 shows the updated trajectory.

Unfirming the behavior of horizontal walking and up and down stairs to a projected equation, the three-dimensional coordinates are calculated as follows:

$$\begin{cases} x(k) = x(k-1) + mH_w \cos\theta(k) + (1-m)L\cos\theta(k) \\ y(k) = y(k-1) + mH_w \sin\theta(k) + (1-m)L\sin\theta(k) \\ z(k) = z(k-1) + mnH_h \end{cases} \tag{2}$$

where $\theta(k)$ is current heading angle, L is step length, H_w, H_h are width and height of staircase, m is sign of behavior pattern, horizontal walking is denoted as "0", and the up and down stairs is denoted as "1", n is sign of up and down floor, it is "1" when the pitch angular velocity detection peak is greater than 0, and "−1" when the peak is less than 0.

In this paper, when pedestrian horizontal walking, the use of PDR calculate the location coordinates, the vertical coordinates remain unchanged. When up and down mode, we don't directly use the pitch angle and pitch rate to calculation height changes. We use it as a basis for judging the behavior pattern and the location information is calculated by staircase information and effective steps. This kind of reckoning method makes use of the law of pedestrian movement. It reduces the influence of the accumulated errors of the sensor and improves the positioning accuracy effectively.

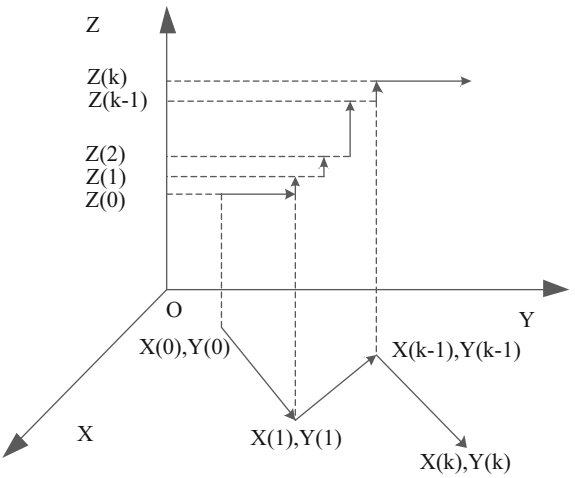


Fig. 2 Dead reckoning

3 Particle Filter Fusion WiFi Localization and 3D PDR

PDR positioning algorithm will produce cumulative error due to inertial device error drift and step frequency error detection, and it can only provide the relative position. For the PDR error accumulation problem, we provide absolute location information through WiFi positioning, and then combining with building information and pedestrian movement mode to fix the positioning results. Considering that the combinatorial system belongs to the nonlinear non-Gaussian filter system, the conventional linear filtering algorithm can't get a good filtering effect, and the particle filter algorithm is adopted to fusion [16]. Particle filter is a recursive Bayesian filter technique based on Monte Carlo sampling method. Through the evolution and propagation of a set of weighted particles to recursively approximate the posterior probability density function of the approximated states, it obtain other statistic on the state.

In this paper, the refinement algorithm that literature [4] proposed is used in WiFi positioning, and the result of coordinate estimation is $\left(\hat{x}_{wifi}, \hat{y}_{wifi}, \hat{z}_{wifi}\right)$. In the traditional PDR algorithm of building information fusion based on particle filter, the weight of the particle is set to zero when it is detected that the sample particle passes through the recording obstacle of the map information during the kth step, and it will be replaced by high weight particles in next time sampling. We further integrate the WiFi positioning results and pedestrian movement mode information in the calculation of the weight of the particles in the basis of particle filter algorithm based on the integration of building information. If a horizontal walking is detected in the kth step, the height state of the particle filter is reset to the floor height. The influence of the WiFi positioning error on the height estimation is reduced and the filtering performance is improved. The particle filter algorithm designed in this paper is described as follows:

1. Initialization. For the initial gait cycle $k = 0, N$ particles are generated according to the initial probability density $p(\mathbf{X}_0)$, the weight of each particle is taken as $\lambda_0^{(i)} = 1/N$, Where the Gauss distribution is used to$p(\mathbf{X}_0)$, The mean is the initial position of the target, which is noted by the WiFi location signed as $\left(x_{wifi,o}, y_{wifi,o}, z_{wifi,o}\right)$.
2. Prediction. According to the prediction Eq. (2.2), predicting the target location of next step $\widehat{\mathbf{X}}_k^{(i)}$.
3. Weight calculation. In [16], the reciprocal of the RSSI of WiFi or GSM network is used as the weight of the particle. The algorithm is simple but the precision is not high. We take the building map information and the deviation of sampling particle from WiFi location results as weight. Calculate as:

$$\lambda_k^{(i)} = \begin{cases} 0 & \text{collision occurred} \\ p\left(Z_k \middle| \mathbf{X}_k^{(i)}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left|\mathbf{X}_k^Z - \mathbf{X}_k^{(i)}\right|}{2\sigma^2}\right) & \text{No collision occurred} \end{cases} \tag{3}$$

where $\mathbf{X}_k^Z$ denotes a position information amount measurement, it's coordinates is obtained from WiFi 3D refinement algorithm $\mathbf{X}_k^Z = \left(\hat{x}_{wifi}, \hat{y}_{wifi}, \hat{z}_{wifi}\right), \mathbf{X}_k^{(i)}$ represents the position information of the *i*th particle at the time, σ represents the standard deviation of measurement noise, according to different environments RSSI time-domain floating selection. The weighting method in this paper means that the weight of the particle is closer to the coordinate of the WiFi centroid location.

4. Weighted normalization. The weights of sampled particles are normalized:

$$\lambda_k^{(i)\prime} = \lambda_k^{(i)} / \sum_{j=1}^{N} \lambda_k^{(j)} \tag{4}$$

5. State estimation. The filtered state estimates and the state probability distributions can be expressed as:

$$\hat{\mathbf{X}}_k^{(i)} = \sum_{i=1}^{N} \lambda_k^{(i)\prime} \widehat{\mathbf{X}}_k^{(i)} \tag{5}$$

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \sum_{i=1}^{N} \lambda_k^{(i)\prime} \delta\left(\mathbf{x}_k - \mathbf{x}_k^{(i)}\right) \tag{6}$$

6. Reset the height estimation. If the horizontal walking is detected, we can get that the current particle is not predicted by the previous particle from Eq. (2.2). The sampled particle remains unchanged. Resetting the height state estimation by resetting the height state to the altitude of current floor:

$$\hat{z}_k = round\left[\frac{\hat{z}_k}{F}\right] \cdot F \tag{7}$$

where, $\hat{z}_k$ is the height of the results, F is the floor height, $round[\cdot]$ represents rounding to the nearest floor.

7. Resampling. Particle filter will produce degenerate linear after a few iterative calculations, so resampling is an effective way to limit the degradation of the importance weight. When the number of valid samples is less than the threshold, resampling N new particles according to the probability density and eliminating low weight particles. The average distribution of particle weights is $1/N$, thereby limiting the degradation phenomenon.

4 Experiment Analysis

4.1 Experimental Preparation

In this paper, Huawei P9 mobile phone is regarded as a smart terminal location of the test equipment. This device incorporates the LSM-330 3-axis acceleration/gyro sensor module and the BCM-4355-XKUB WiFi controller module. Experimental site is in two and third floors of the experimental building. Ardon WiFi access points are deployed in each of the two floors (Fig. 3), for a total of 64 and using high precision GPS receivers calibrate the exact location to construct WiFi location networks. The gateway is connected to the PC via Ethernet, and the PC serves as the server for positioning. The wireless sensor networks use the time division multiplexing mechanism in the communication process. The gateway, mobile phone and WiFi nodes are operating in the 2.4 GHz bands. The mobile phone transmits the MEMS sensor measurement data together with the RSSI information and the beacon's ID to the gateway, and the gateway transmits the data to the PC The positioning function of the pedestrian is realized by the PC.

The experimental design of the routes shows in Fig. 3. The solid line indicates the third floor route, and the dotted line indicates the second floor. In the process of moving, the total number of steps is 483. Two experimental programs are designed for comparative study:

1. 3D PDR positioning
2. 3D localization of WiFi/PDR based on particle filter

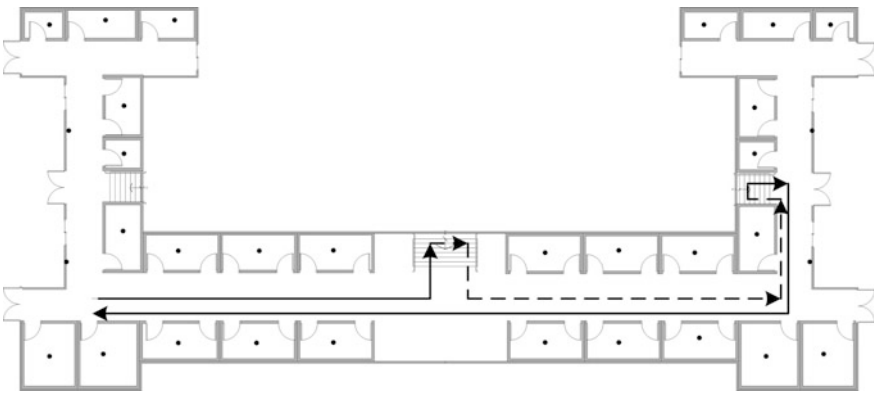


Fig. 3 Experiment design and environment

4.2 Analysis of Results

Respectively, positioning experiments are implied according to the design programs. In order to compare the accuracy of the two positioning schemes, the horizontal coordinate and the height are separately represented.

Figure 4 shows the horizontal positioning results of the two schemes. In the indoor environment, 3D PDR achieves the trajectory of the pedestrian tracking, and the average error of positioning is 4.7 m. In the absence of absolute position information correction, the positioning will be more and more deviated from the true track, due to the PDR inherent susceptibility to the cumulative error of the characteristics. After the turn, the locating trajectory deviation is more significant, which is manifested as the parallel spacing of different distances. From the experimental results of the 3D localization of WiFi/PDR based on particle filter, it can be seen that the accuracy of the multi-source fusion scheme is further improved that compared to PDR with an average error of 2.1 m. This is because the WiFi/PDR location algorithm based on particle filter can correct the accumulated error of PDR with the result of WiFi centroid location, and use the building information and behavior patterns to guide the sample particles to move to the real position.

Figure 5 shows the comparison of the height estimation results of the two schemes. Solid line indicates the height of the second and third floors relative to the

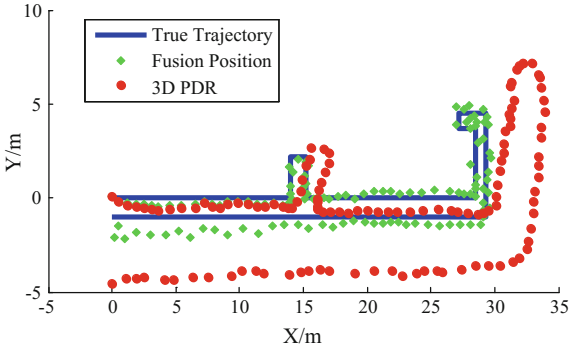


Fig. 4 Horizontal coordinate estimation

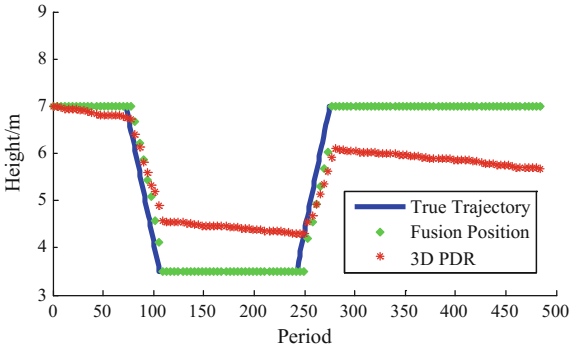


Fig. 5 Height estimation

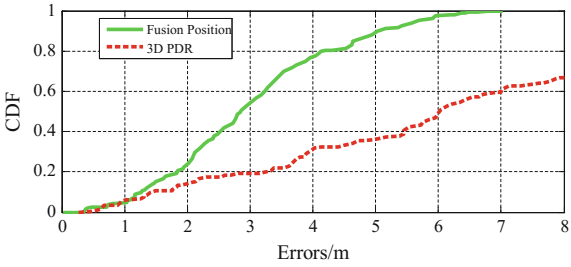


Fig. 6 Comparison of CDF

first floor, and dots indicate the height of the anchor point during walking. "Triangle" symbol is the height estimation result of 3D PDR location, and the height errors accumulated over times. Long-term use will not be able to make accurate tracking. "Asterisk" is the height estimation result of WiFi/PDR combination based on particle filter. The height changes only in up and down stairs and it is ladder-like. It is because of the using of pitch angle detection and floor information to estimate height, and the height is modified only when the behavior pattern of up and down stairs is detected. After the end of the upstairs or downstairs, the height will be set to floor height, and until up and down behaviors are detected again. Positioning error only occurred in the period of up and down stairs, and the average positioning error is 0.8 m which means it can effectively distinguish the floor information.

Figure 6 shows the comparison of the Cumulative Distribution Function (CDF) between the proposed scheme and the EKF-based three-axis MARG sensor PDR localization of the smartphone in the literature [11]. It can be found that the median error of the proposed technique is 2.9 m and the maximum error is 6.7 m. The median error of 3D PDR is 6.0 m. The positioning accuracy of the proposed technique is superior to that scheme, and the performance of positioning can be improved significantly by merging two positioning techniques which are mutually different but applied complementary.

5 Conclusion

Aiming at the problem that the existing indoor 3D positioning accuracy is low, this paper puts forward the multi-source fusion indoor layered positioning technology based on WiFi/PDR. Horizontal walking and up and down stairs movement patterns is distinguished using pitch rate detection method. The three-dimensional localization of the PDR is carried out for each effective walking step of the different walking motion patterns. A WiFi/PDR location algorithm based on particle filter is proposed to eliminate the accumulated error of PDR by using the absolute location information of WiFi location. The experimental results show that the three-dimensional positioning can be realized continuously. The average error of

the horizontal positioning is 2.2 m, the average estimation error is 0.8 m, and it can distinguish the floor and satisfy the needs of indoor high-precision positioning.

References

1. Wu C, Yang Z, Liu Y (2015) Smartphones based crowdsourcing for indoor localization. IEEE Trans Mob Comput 14(2):444–457
2. Liu XC, Wu ZF, Lin XK (2013) WLAN/MARG/GPS integrated positioning system based on a self-adaptive weighted algorithm. J Tsinghua Univ (Sci Technol) 53(7):955–960
3. Devanshi S, Agrawal S, Singh S (2014) 3D indoor localization based on Bluetooth technology: a brief review. Int J Comput Appl 97(8):31–33
4. Shu J, Liu LG, Chen YB et al (2012) 3D-RABLC: a novel refinement algorithm based on LQI confidence for three-dimensional localization. J Commun 33(7):125–134
5. Taekook K, Eui-Jik K (2014) A novel 3D indoor localization scheme using virtual access point. Int J Distrib Sens Netw 1–6
6. Monica S, Ferrari G (2016) A swarm-based approach to real-time 3D indoor localization: experimental performance analysis. Appl Soft Comput 43:489–497
7. Nakamura M, Akiyama T, Sugimoto M et al (2014) 3D FDM-PAM: rapid and precise indoor 3D localization using acoustic signal for smartphone. In: 2014 ACM international joint conference on pervasive and ubiquitous computing. ACM, pp 123–126
8. Yang SH, Kim HS, Son YH et al (2014) Three dimensional visible light indoor localization using AOA and RSS with multiple optical receivers. J Lightwave Technol 32(14):2480–2485
9. Zheng XL, Fu JQ (2015) Study on PDR and RSSI based indoor localization algorithm. Chin J Sci Instr 36(5):1177–1185
10. Vanini S, Giordano S (2013) Adaptive context-agnostic for transition detection on smart mobile devices. In: 2013 IEEE international conference on pervasive computing and communications workshops (PERCOM workshops). IEEE, pp 2–7
11. Tian Z, Zhang Y, Zhou M et al (2014) Pedestrian dead reckoning for MARG navigation using a smartphone. EURASIP J Adv Sig Process 65:1–9
12. Zheng L, Zhou W, Tang W et al (2015) A foot-mounted sensor based 3D indoor positioning approach. In: 2015 IEEE twelfth international symposium on autonomous decentralized systems (ISADS). IEEE, pp 145–150
13. Zhou W, Tang W, Zheng X et al (2016) A 3D indoor positioning system based on low-cost MEMS sensors. Simul Model Pract Theory 65:45–56
14. Chen GL, Meng XL, Wang Y (2015) Integrated WiFi/PDR/Smartphone using an unscented kalman filter algorithm for 3D indoor localization. Sensors 15(9):24595–24614
15. Banerjee D, Agarwal S K, Sharma P (2015) Improving floor localization accuracy in 3D spaces using barometer. In: Proceedings of the 2015 ACM international symposium on wearable computers. ACM, pp 171–178
16. Bojja J, Kirkko-Jaakkola M, Collin J et al (2014) Indoor localization methods using dead reckoning and 3D map matching. J Sig Process Syst 76(3):301–312

Efficient and Robust Convex Relaxation Methods for Hybrid TOA/AOA Indoor Localization

Enwen Hu, Zhong-liang Deng, Lu Yin, Di Zhu, Jun Lu and Yanping Zhao

Abstract In this chapter, we proposed a hybrid positioning method based on convex relaxation with time of arrival (TOA) and angle of arrival (AOA) measurements. Traditional maximum likelihood (ML) formulation for indoor localization is a nonconvex optimization problem. We exploit the relaxation methods to provide efficient convex solution. Besides, we apply this method to localization with hybrid TOA/AOA measurements firstly and the linear Cramer-Rao Bounds in the scenarios of error-free and erroneous location of sensor nodes are deduced, respectively. Simulations based on TC-OFDM signal system show that the proposed method is efficient and more robust as compared to the existing ML estimation and TOA or AOA based convex relaxation with or without error of sensor nodes location.

Keywords Convex relaxation · Hybrid · TOA · AOA · Indoor location · TC-OFDM

1 Introduction

In recent years, Location based services (LBSs) are a significant permissive technology and becoming a vital part of life. Today, more than 80–90% of people's time is spent indoors (Strategy Analytics). The global indoor location market is expected to grow from $935.05 million in 2014 to $4.42 billion by 2019, at an estimated Compound Annual Growth Rate (CAGR) of 36.5% from 2014 to 2019 (Research and Markets) with strong demand in healthcare, retail, hospitality, travel and other sectors. For localization in an outdoor environment, global navigation satellite systems (GNSS) such as GPS and BeiDou have been widely used in variety of applications. However, the main disadvantage of the GNSS approach is that

E. Hu (✉) · Z. Deng · L. Yin · D. Zhu · J. Lu · Y. Zhao
School of Eletronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
e-mail: owen.hu@bupt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_48

precise location is not possible, when there is no Line-Of-Sight (LOS) connection to at least four satellites, thus most applications function poorly in dense urban areas, indoors, underwater or underground with GNSS. Recently, various techniques have been proposed for indoor positioning system such as infrared (IR) [1], ultrasound [2], radio-frequency identification (RFID) [3], WiFi [4], Bluetooth [5], sensor networks [6], ultra-wideband (UWB) [7], geomagnetic [8], vision analysis [9] and Pseudolite [10]. Based on these fundamental technologies, many IPSs have been developed by companies, research institutes and universities. Each system takes advantage of a particular positioning technology or integrating some of these technologies. They make tradeoff between the performance and complexity of the IPSs.

With the increasing demand of high accuracy positioning, hybrid signal based localization is a good choice. Some nonlinear estimators including ML and nonlinear least squares (NLS) are investigated in the literature. Generally speaking, the corresponding cost functions are multi-model, so global optimization cannot be guaranteed. Convex optimization, particularly the semi-definite program (SDP) relaxation method, is a more recent positioning approach which strikes a balance between nonlinear and linear methods, namely, high accuracy and global convergence. Nonconvex optimization problem such as NLS and ML can be transformed into either a convex second-order cone program (SOCP), or a semidefinite program (SDP) by relaxation [11].

The rest of this paper is organized as follows. Section 2 states the problem formulation and the conventional ML solution of the problem. The SDP relaxation for hybrid TOA/AOA Measurements is discussed in Sect. 3 and localization with sensor location errors is presented in Sect. 4. Section 5 analyze the Linear Cramer-Rao Bound (LCRB) with and without error of sensor nodes locations. The proposed localization method is compared to an existing approach as well as the LCRB in Sect. 6. Conclusions are drawn in Sect. 7.

2 Problem Formulation

We consider a target localization problem using hybrid TOA/AOA measurements in 2-dimensional space. There are L non-collinear sensor nodes arbitrarily deployed at surveillance region, whose positions are known as $\mathbf{x}_i = [x_i, y_i]^T, (\mathbf{x}_i \in R^2)$, $i = 1, \ldots, N$, and a target node which has unknown coordinates i.e.$\mathbf{u} = [x, y]^T$, $(\mathbf{u} \in R^2)$. We assume that all sensor nodes' clocks are ideally synchronized and all sensor nodes receive signal with line of sight in this paper. Each sensor node can measure the TOA and AOA of the signal transmitted from the target. Let t_i and θ_i be the signal propagation delay from the target to the ith sensor node and angle measurement, respectively, which are modeled as:

$$\begin{cases} t_i = \frac{1}{c}\|\mathbf{u} - \mathbf{x}_i\| + t_0 + e_{ti} \\ \theta_i = \arctan\frac{y - y_i}{x - x_i} + e_{\theta i} \end{cases} \tag{1}$$

where c is the signal propagation speed, $\|\bullet\|$ denotes the Euclidean norm, t_0 is the local time at the sensors when the signal leaves the target, and $e_{ti}, e_{\theta i}$ is the unknown measurement error. Then the x and y coordinates of target node are given by

$$\begin{cases} \hat{x} = x_i + d_i \cos\theta_i \\ \hat{y} = y_i + d_i \sin\theta_i \end{cases} \tag{2}$$

where $d_i = c * t_i$ denotes the measured distance between the ith sensor node and target node, e_{xi}, e_{yi} are the errors of x and y coordinates corresponding to the measurements e_{ti} and $e_{\theta i}$, given by $e_{xi} = f_{xi}(e_{ti}, e_{\theta i})$ and $e_{yi} = f_{yi}(e_{ti}, e_{\theta i})$. Thus, Eq. (2) can be written in matrix form as [11]:

$$\mathbf{b} = \mathbf{A}\mathbf{u} + \mathbf{e}_u \tag{3}$$

where $\mathbf{A} = \begin{bmatrix} \mathbf{1}_{N\times 1} & \mathbf{0}_{N\times 1} \\ \mathbf{0}_{N\times 1} & \mathbf{1}_{N\times 1} \end{bmatrix}$, $\mathbf{u} = [x, y]^T$, $\mathbf{e}_u \in R^{2N\times 1}$, and $\mathbf{e}_u$ is the noise vector with zero mean vector and known covariance $E(\mathbf{e}_u \mathbf{e}_u^{\mathrm{T}}) = Q$ with $^{-1}$ and E and ^{T} being the inverse, expectation and transpose operators, respectively. $\mathbf{b}$ can be written as:

$$\mathbf{b} = \begin{bmatrix} x_1 + d_1\cos\theta \\ \vdots \\ x_N + d_N\cos\theta_N \\ y_1 + d_1\sin\theta_1 \\ \vdots \\ y_N + d_N\sin\theta_N \end{bmatrix} \tag{4}$$

Consequently, the maximum likelihood formulation of the target localization problem can be directly formulated as:

$$\underset{\hat{u}}{\text{minimize}}\ (\mathbf{b} - \mathbf{A}\hat{\mathbf{u}})^T Q(\mathbf{b} - \mathbf{A}\hat{\mathbf{u}}) \tag{5}$$

where $\mathbf{b}$, $\mathbf{A}$ and $\mathbf{Q}$ are given as before, respectively.

3 SDP Relaxation for Joint TOA/AOA Measurements

For ML target localization (5), the function is a nonconvex function of the unknown **u**. The ML solution of the localization problem, which is a non-convex optimization problem is presented above and it cannot achieve global optimization. Convex optimization, particularly the semi-definite program (SDP) relaxation method, is a more recent positioning approach which strikes a balance between nonlinear and linear methods, namely, high accuracy and global optimal convergence. One of the main contributions of this paper is the relaxation of the ML solution to a convex optimal problem, which can be efficiently solved by existing toolboxes in the literature. The main novelty in devising the new SDP relaxation formulation is to include the noise error as variables in the ML cost function. Equation (2) can be rewritten as:

$$\begin{cases} \hat{x} = x_i + d_i \cos\theta_i = x + e_{xi} \\ \hat{y} = y_i + d_i \sin\theta_i = y + e_{yi} \end{cases} \tag{6}$$

where e_{xi} and e_{yi} represent the error of x and y of ith sensor, respectively, which are nonlinear function of e_{ti} and $e_{\theta i}$ in Eq. (1).

Then the positioning error of $\mathbf{u} = [x, y]^T$ can be written as:

$$e_{xi}^2 + e_{yi}^2 = \begin{bmatrix} (x_i + d_i \cos\theta_i - x)^2 \\ +(y_i + d_i \sin\theta_i - y)^2 \end{bmatrix} \tag{7}$$

Then we obtain

$$e_i^2 = e_{xi}^2 + e_{yi}^2 = d_i^2 + \|\mathbf{u} - \mathbf{x}_i\|_2^2 - 2d_i[(x - x_i)\cos\theta_i + (y - y_i)\sin\theta_i] \tag{8}$$

where the auxiliary variable $e_i^2 = e_{xi}^2 + e_{yi}^2$ is introduced. The target localization can be written as

$$\begin{aligned} &\underset{\mathbf{u}}{\text{minimize}} \sum_{i=1}^{N} e_i^2 \\ &\textit{subject to } e_i^2 = d_i^2 + \|\mathbf{u} - \mathbf{x}_i\|^2 - 2d_i[(x - x_i)\cos\theta_i + (y - y_i)\sin\theta_i] \\ &\qquad i = 1, 2, \ldots, N \end{aligned} \tag{9}$$

Each nonlinear constraint in (4) can be converted to a linear one by spreading $\|\mathbf{u} - \mathbf{x}_i\|^2$, we can obtain

$$\|\mathbf{u} - \mathbf{x}_i\|^2 = \mathbf{u}^{\mathrm{T}}\mathbf{u} - 2\mathbf{x}_i^{\mathrm{T}}\mathbf{u} + \mathbf{x}_i^{\mathrm{T}}\mathbf{x}_i \tag{10}$$

and

$$2d_i[(x-x_i)\cos\theta_i+(y-y_i)\sin\theta_i]=2d_i(\mathbf{u}-\mathbf{x}_i)^T\vartheta_i,\ \vartheta_i=[\cos\theta_i,\sin\theta_i]^T \tag{11}$$

The constraints in (9) can be given by

$$d_i^2+\mathbf{x}_i^T\mathbf{x}_i+2d_i\mathbf{x}_i=2(\mathbf{x}_i+d_i\vartheta_i)^T\mathbf{u}-\tilde{R}+e_i^2 \tag{12}$$

where $\tilde{R}=\mathbf{u}^T\mathbf{u}$. As a result, we can rewrite (4) as

$$\begin{aligned}&\underset{\zeta}{\text{minimize}}\ \sum_{i=1}^{N}q_i\\ &\textit{subject to}\ y=A\zeta\\ &\qquad q_i=e_i^2,\ i=1,2,\ldots,N\\ &\qquad \tilde{R}=\mathbf{u}^{\mathrm{T}}\mathbf{u}\end{aligned} \tag{13}$$

where

$$y=A\zeta$$

$$y=\begin{bmatrix}d_1^2+\mathbf{x}_1^T\mathbf{x}_1+2d_1\mathbf{x}_1\\ d_2^2+\mathbf{x}_2^T\mathbf{x}_2+2d_2\mathbf{x}_2\\ \vdots\\ d_N^2+\mathbf{x}_N^T\mathbf{x}_N+2d_N\mathbf{x}_N\end{bmatrix} \tag{14}$$

$$A=\begin{bmatrix}2(\mathbf{x}_1+d_1\vartheta_1)^T & -1 & 1 & 0 & \cdots & 0\\ 2(\mathbf{x}_1+d_1\vartheta_1)^T & -1 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ 2(\mathbf{x}_N+d_N\vartheta_N)^T & -1 & 0 & 0 & \cdots & 1\end{bmatrix} \tag{15}$$

$$\zeta=\left[\mathbf{u},\tilde{R},\tilde{\mathbf{q}}\right]^T \tag{16}$$

with $\tilde{\mathbf{q}}=\begin{bmatrix}e_1^2 & e_2^2 & \cdots & e_N^2\end{bmatrix}$. Note that the optimization problem described in (13) is still nonconvex as the constraints is nonconvex. Hence, it is still difficult to obtain the solution. The semidefinite relaxation is employed such that these two equalities can be relaxed into convex inequalities, $q_i\geq e_i^2$ and $\tilde{R}\geq\mathbf{u}^T\mathbf{u}$. Adopting the relaxation approach, we now have transformed the ML problem into a convex optimization problem:

$$\begin{aligned} &\underset{\zeta}{\text{minimize}} \sum_{i=1}^{N} q_i \\ &\textit{subject to } y = A\zeta \\ &\qquad q_i \geq e_i^2, \quad i = 1, 2, \ldots, N \\ &\qquad \tilde{R} \geq \mathbf{u}^{\mathrm{T}}\mathbf{u} \end{aligned} \tag{17}$$

The global optimal solution to the unknown source position can be derived by interior point methods such as SeDuMi [12].

4 Localization with Sensor Location Errors

In the previous sections, we assume that the sensors locations are known exactly; that is, $\mathbf{x}_i$ is accurately known. However, the sensors locations may not be known perfectly because of imperfections in sensor nodes in sensor deployment, positioning, and delay of position updating. We are interested in making source localization robust with the sensor location errors. In this section, we take the sensor location uncertainty into account and present a robust target localization method. The erroneous sensor locations can be modeled as:

$$\begin{cases} z_{x,i} = x_i + x_{e,i} \\ z_{x,i} = x_i + x_{e,i} \end{cases} \tag{18}$$

where $x_{e,l}$ and $y_{e,l}$ are the zero-mean disturbances in the lth sensor location and $z_{e,i} = \left[x_{e,i}, y_{e,i}\right]^T \rightarrow \mathrm{N}(0_{2\times 1}, \boldsymbol{\sigma}_i)$ and $\boldsymbol{\sigma}_i$ is the covariance matrix for $z_{e,i}$ which is assumed known. The maximum likelihood estimation of the target location in the presence of sensor location errors is conducted with the function in the similar form to (9) given by

$$\begin{aligned} &\underset{e_i, \mathbf{u}, \mathbf{z}_{e,i}}{\text{minimize}} \sum_{i=1}^{N} e_i^2 + \sum_{i=1}^{N} \mathbf{z}_{e,i}^T \boldsymbol{\sigma}_i^{-1} \mathbf{z}_{e,i} \\ &\textit{subject to} \\ &e_i^2 = d_i^2 + \left\| \mathbf{u} - \left(\mathbf{z}_i - \mathbf{z}_{e,i}\right) \right\|^2 \\ &\quad - 2d_i \left[\left(x - \left(z_{x,i} - x_{e,i}\right)\right) \cos\theta_i + \left(y - \left(z_{y,i} - y_{e,i}\right)\right) \sin\theta_i \right] \\ &i = 1, 2, \ldots, N \end{aligned} \tag{19}$$

And the constrains can be expressed by

$$d_i^2 + \mathbf{u}^T\mathbf{u} - 2\mathbf{u}^T\mathbf{z}_i - \mathbf{z}_i^T\mathbf{z}_i - 2d_i(\mathbf{u} - \mathbf{z}_i)^T\vartheta_i = e_i^2 + 2(d_i\vartheta_i + \mathbf{u} - \mathbf{z}_i)^T\mathbf{z}_{e,i} + \mathbf{z}_{e,i}^T\mathbf{z}_{e,i} \tag{20}$$

Let

$$\begin{aligned} X_1 &= \left[\mathbf{z}_{e,1}, \mathbf{z}_{e,2}, \ldots, \mathbf{z}_{e,N}\right] \\ X_2 &= \left[d_1\vartheta_1 + \mathbf{u} - \mathbf{z}_1, d_2\vartheta_2 + \mathbf{u} - \mathbf{z}_2, \ldots, d_N\vartheta_N + \mathbf{u} - \mathbf{z}_N\right] \\ Y &= X_2^T X_1 \\ R_2 &= \mathbf{z}_{e,i}^T\mathbf{z}_{e,i} \end{aligned} \tag{21}$$

Thus

$$\begin{aligned} \underset{\zeta}{\text{minimize}} \ & \sum_{i=1}^{N} q_i + \sum_{i=1}^{N} \mathbf{z}_{e,i}^T\boldsymbol{\sigma}_i^{-1}\mathbf{z}_{e,i} \\ \textit{subject to } & y = A\zeta \\ & q_i \geq e_i^2, \ i = 1, 2, \cdots, N \\ & \tilde{R}_1 \geq \mathbf{u}^T\mathbf{u} \\ & \tilde{R}_2 \geq \mathbf{z}_e^T\mathbf{z}_e \end{aligned} \tag{22}$$

where

$$\begin{aligned} y &= \left[\, d_1^2 - z_1^T z_1 + 2d_1 z_1^T\vartheta_1 \quad \cdots \quad d_N^2 - z_N^T z_N + 2d_N z_N^T\vartheta_N \,\right]^T \\ A &= \left[I_N, 1_N, -1_N, 2diag\left(z_1^T - d_1\vartheta_1, \ldots, z_N^T - d_N\vartheta_N\right)\right] \\ \zeta &= \left[\mathbf{q}^T, diag(Y), \tilde{R}_2, \mathbf{u}^T\right] \\ \mathbf{z}_e &= \left[\mathbf{z}_{e,1}, \ldots, \mathbf{z}_{e,N}\right] \end{aligned} \tag{23}$$

5 Linear Cramer-Rao Lower Bound (LCRB) for Joint TOA/AOA Measurements

5.1 Linear Cramer-Rao Bound with Known Sensor Locations

CRB as the optimal performance indicator for unbiased estimator is widely applied in the localization and positioning system. In this section, the error bounds of the localization will be analyzed by using the joint TOA/AOA measurements based on Fisher Information Matrix (FIM). Let $\hat{\mathbf{x}}_i = [\hat{x}_i, \hat{y}_i]^T$ be the estimate of the ith location

$\mathbf{x}_i = [x_i, y_i]^T$ obtained from the joint TOA t and AOA θ measurements. Covariance matrix is calculated by

$$Var(\hat{\mathbf{x}}) = E\left[(\hat{\mathbf{x}} - E(\hat{\mathbf{x}}))(\hat{\mathbf{x}} - E(\hat{\mathbf{x}}))^T\right] = \begin{bmatrix} \sigma_{\hat{x}}^2 & \sigma_{\hat{x}\hat{y}} \\ \sigma_{\hat{y}\hat{x}} & \sigma_{\hat{y}}^2 \end{bmatrix} \tag{24}$$

where $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_N]$, $\sigma_{\hat{x}}^2$ and $\sigma_{\hat{y}}^2$ are the Mean Square Errors (MSE) of $[\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_N]$ and $[\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_N]$, $\sigma_{\hat{x}\hat{y}}$ and $\sigma_{\hat{y}\hat{x}}$ are the covariance between x and y, and between y and x, respectively. For simplicity, we assume that e_{ti} and $e_{\theta i}$ in Eq. (1) follow the Gaussian distribution with the mean zero and variance σ_1 and σ_2, respectively, and the joint TOA/AOA measurements follow the bivariate normal distribution. Thus the PDF of the TOA/AOA measurements with respect to $\mathbf{u}$ is

$$f_{\mathbf{u}}(TOA/AOA) = \prod_{i=1}^{N} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}} \exp\left(-\frac{\varpi_{12}}{2(1-\rho_{12}^2)}\right) \tag{25}$$

where ρ_{12} is the correlation coefficient of t_i and θ_i, $\varpi_{12} = \frac{\xi_1^2}{\sigma_1^2} + \frac{\xi_2^2}{\sigma_2^2} - 2\rho_{12}\frac{\xi_1}{\sigma_1}\frac{\xi_2}{\sigma_2}$, and based on (1), we have

$$\begin{cases} \xi_1 = t_i - \frac{1}{c}\|\mathbf{u} - \mathbf{x}_i\| + t_0 \\ \xi_2 = \theta_i - \arctan\frac{y - y_i}{x - x_i} \end{cases} \tag{26}$$

If $\hat{\mathbf{u}}_i$ is calculated from the unbiased estimate of the joint TOA/AOA measurements at the ith location $\mathbf{u}$. Then, the CRB tells us that

$$Var(\hat{\mathbf{u}}) \geq \left[I^{-1}(\mathbf{u})\right] \tag{27}$$

where $I(\mathbf{u})$ is the Fisher Information Matrix:

$$I(\mathbf{u}) = -\frac{\partial^2 \mathrm{In}\, f_{\mathbf{u}}(TOA/AOA)}{\partial \mathbf{u}^2} \tag{28}$$

where $f_{\mathbf{u}}(TOA/AOA)$ is the Probability Distribution Function (PDF) of $\mathbf{u}$ with respect to t_i and θ_i. Let

$$I(\mathbf{u}) = \begin{bmatrix} I_{xx}(\mathbf{u}) & I_{xy}(\mathbf{u}) \\ I_{yx}(\mathbf{u}) & I_{yy}(\mathbf{u}) \end{bmatrix} \tag{29}$$

where

$$\begin{aligned} I_{xx}(\mathbf{u}) &= -E\left(\frac{\partial^2 \mathrm{In}\, f(\mathbf{u})}{\partial x^2}\right), \quad I_{xy}(\mathbf{u}) = -E\left(\frac{\partial^2 \mathrm{In}\, f(\mathbf{u})}{\partial xy}\right) \\ I_{yx}(\mathbf{u}) &= -E\left(\frac{\partial^2 \mathrm{In}\, f(\mathbf{u})}{\partial xy}\right), \quad I_{yy}(\mathbf{u}) = -E\left(\frac{\partial^2 \mathrm{In}\, f(\mathbf{u})}{\partial y^2}\right) \end{aligned} \tag{30}$$

Thus the maximum accuracy of the two dimensional localization can be given by the LCRB [12]

$$Var(\hat{\mathbf{u}}) \geq \frac{I_{xx}(\mathbf{u}) + I_{yy}(\mathbf{u})}{\det[I(\mathbf{u})]} \tag{31}$$

where $\det[I(\mathbf{u})] = I_{xx}(\mathbf{u})I_{yy}(\mathbf{u}) - I_{xy}(\mathbf{u})^2$, and

$$\begin{aligned} I_{xx}(\mathbf{u}) &= -\frac{1}{2\left(1-\rho_{12}^2\right)}\left[A_1\frac{\partial \xi_1}{\partial x} + A_2\frac{\partial^2 \xi_1}{\partial x^2} + A_3\frac{\partial \xi_2}{\partial x} + A_4\frac{\partial^2 \xi_2}{\partial x^2}\right] \\ I_{xy}(\mathbf{u}) &= -\frac{1}{2\left(1-\rho_{12}^2\right)}\left[B_1\frac{\partial \xi_1}{\partial x} + B_2\frac{\partial^2 \xi_1}{\partial x\partial y} + B_3\frac{\partial \xi_2}{\partial x} + B_4\frac{\partial^2 \xi_2}{\partial x\partial y}\right] \\ I_{yy}(\mathbf{u}) &= -\frac{1}{2\left(1-\rho_{12}^2\right)}\left[C_1\frac{\partial \xi_1}{\partial y} + C_2\frac{\partial^2 \xi_1}{\partial y^2} + C_3\frac{\partial \xi_2}{\partial y} + C_4\frac{\partial^2 \xi_2}{\partial y^2}\right] \\ I_{yx}(\mathbf{u}) &= I_{xy}(\mathbf{u}) \end{aligned} \tag{32}$$

and

$$\begin{aligned} A_1 &= \frac{2}{\sigma_1^2}\frac{\partial \xi_1}{\partial x} - \frac{2\rho_{12}}{\sigma_1\sigma_2}\frac{\partial \xi_2}{\partial x}, \quad A_2 = \frac{2\xi_1}{\sigma_1^2} - \frac{2\rho_{12}\xi_2}{\sigma_1\sigma_2} \\ A_3 &= \frac{2}{\sigma_2^2}\frac{\partial \xi_2}{\partial x} - \frac{2\rho_{12}}{\sigma_1\sigma_2}\frac{\partial \xi_1}{\partial x}, \quad A_4 = \frac{2\xi_2}{\sigma_2^2} - \frac{2\rho_{12}\xi_1}{\sigma_1\sigma_2} \end{aligned} \tag{33}$$

$$\begin{aligned} B_1 &= \frac{2}{\sigma_1^2}\frac{\partial \xi_1}{\partial y} - \frac{2\rho_{12}}{\sigma_1\sigma_2}\frac{\partial \xi_2}{\partial y}, \quad B_2 = \frac{2\xi_1}{\sigma_1^2} - \frac{2\rho_{12}\xi_2}{\sigma_1\sigma_2} \\ B_3 &= \frac{2}{\sigma_2^2}\frac{\partial \xi_2}{\partial y} - \frac{2\rho_{12}}{\sigma_1\sigma_2}\frac{\partial \xi_1}{\partial y}, \quad B_4 = \frac{2\xi_2}{\sigma_2^2} - \frac{2\rho_{12}\xi_1}{\sigma_1\sigma_2} \end{aligned} \tag{34}$$

$$C_i = B_i, \quad i = 1, 2, 3, 4 \tag{35}$$

5.2 Linear Cramer-Rao Bound with Sensor Location Errors

The location information of sensor nodes are usually not exactly known. We would like to analyze the effect of node location error on the performance limit of source localization. In the preceding development of the LCRB with known sensor locations, we have obtained the closed-form of LCRB when the sensor locations are exactly known. Similar to the previous analyzation, the ξ_1 and ξ_2 in (26) are changed to

$$\begin{cases} \xi_1' = t_i - \frac{1}{c}\|\mathbf{u} - \mathbf{z}_i\| + t_0 \\ \xi_2' = \theta_i - \arctan\frac{y - z_{y,i}}{x - z_{x,i}} \end{cases} \tag{36}$$

As is known, the LCRB with the sensor location errors are developed with functions in the similar form to (31–35). But the errors of sensor locations will have the effect on the performance limit of source localization.

6 Simulation Results

In this section, simulations are conducted to evaluate the performance of the proposed methods. We exploit the wireless insite software based on ray tracing to generate the TOA and AOA measurements including 25 paths in the scenario shown in Fig. 1 where 7 sensor nodes in red circle are configured. The MATLAB toolbox SeDuMi [13] is utilized for the algorithm realization where the solver SDPT3 [14] is employed. We compare the performance of the proposed methods with the existing state-of-the-art algorithms. We mainly compare the root mean square errors (RMSEs), average positioning errors and the effect of the number of sensor nodes in the scenarios of error-free and erroneous location of sensor nodes, respectively.

In Fig. 2, we selected 15 test points (TPs) in black rectangle to evaluated the performance of proposed method by Monte Carlo simulations according to (17) and (22) in the scenarios with and without error of sensor nodes location. (Ex_1, Ey_1), (Ex_2, Ey_2), Er_1 and Er_2 represent the coordinate estimation errors and location errors in the error-free and erroneous sensor nodes location scenarios, respectively. The simulation results listed in Table 1 show that the positioning accuracy with unknown sensor nodes is larger than other scenario which the locations of all sensors are exactly known. And the TPs (e.g. #11, #12, #13, #14, #15) lie outside the convex hull of the sensor nodes, the positioning error is larger than others, which is the manifestation of the convex hull problem [15].

We investigate the performance of the proposed method for hybrid TOA/AOA-based localization with SDP relaxation in the scenario where the locations of all the sensor nodes are exactly known. Comparisons with LCRB, traditional ML [12], and TOA based location [15] with SDP are also made. We

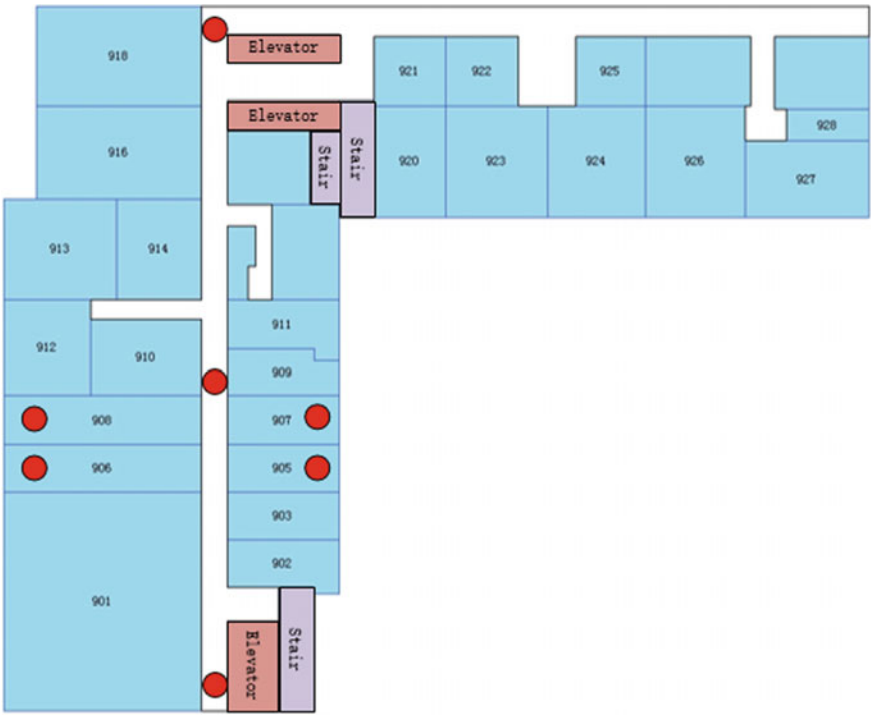


Fig. 1 Simulation scenario and sensor nodes distribution (color figure online)

consider the root mean square error (RMSE) versus the variances of TOA and AOA measurements represented by σ_θ and σ_t respectively. From Fig. 3, we can know that the proposed method outperforms the traditional ML and TOA based location and match with the optimal performance predicted by LCRB in Sect. 5.1. Figure 4 shows the RMSE of TOA based, traditional ML and proposed method when the locations of sensor nodes is uncertainty and we know that the proposed method is more robust than other methods, but closely follows the LCRB and is worse than the scenario without error of sensor nodes.

The simulations in Figs. 5 and 6 are conducted in the scenario where all the TPs lie outside the convex hull and the variances of TOA and AOA are the same as in Figs. 3 and 4. We note that the RMSE performance in Figs. 5 and 6 are now worse compared to Figs. 3 and 4, respectively. As shown in Table 1, when the TPs outside the convex hull of sensor nodes, Eqs. (17) and (22) will suffer from a convex hull problem and the performance will be degraded and mismatch with the LCRB.

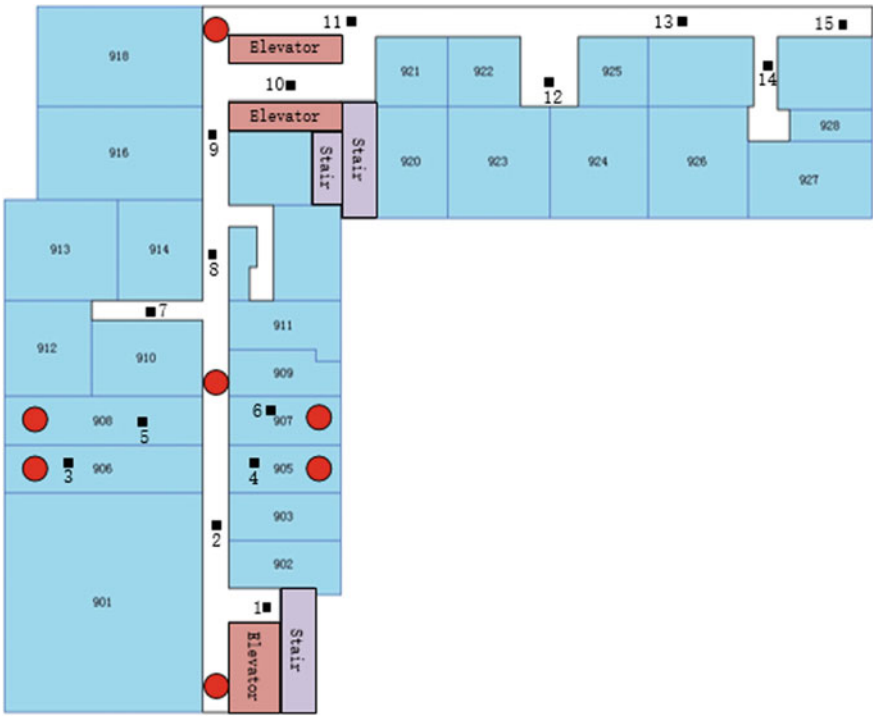


Fig. 2 Test points distribution

Table 1 RMSE of location

TP	Ex_1	Ey_1	Er_1	Ex_2	Ey_2	Er_2
1	0.57	0.87	1.04	1.36	1.97	2.39
2	0.76	0.49	0.90	1.28	1.04	1.65
3	0.64	1.24	1.40	1.25	1.36	1.85
4	1.06	1.52	1.85	1.45	1.21	1.89
5	1.2	0.78	1.43	1.67	1.43	2.20
6	1.47	0.48	1.55	1.75	2.01	2.67
7	0.94	1.25	1.56	0.81	1.05	1.33
8	0.87	1.23	1.51	1.14	1.27	1.71
9	0.91	1.31	1.60	1.89	1.21	2.24
10	1.04	1.25	1.63	1.76	0.98	2.01
11	1.56	1.87	2.44	2.54	2.06	3.27
12	1.02	2.46	2.66	2.01	1.92	2.78
13	2.79	1.87	3.36	2.69	2.04	3.38
14	2.11	1.98	2.89	2.93	1.98	3.54
15	1.82	1.41	2.30	2.19	1.91	2.91

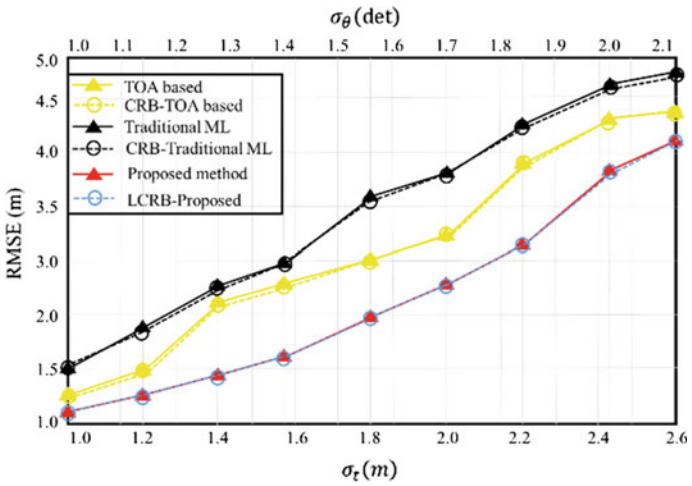


Fig. 3 RMSE of target location estimate versus σ_θ and σ_t with sensor location exactly known

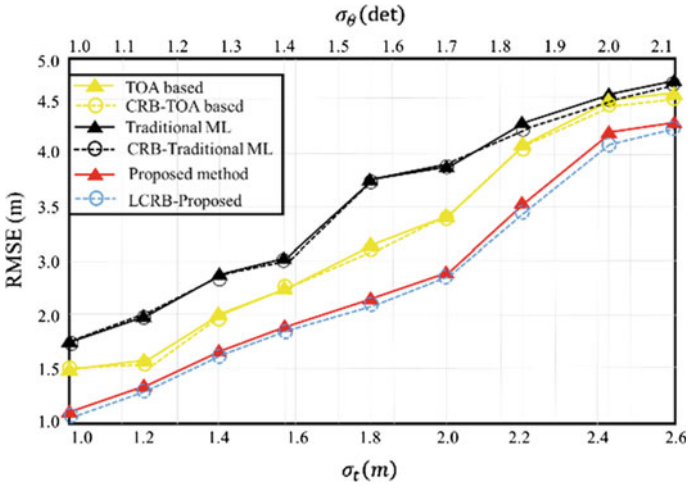


Fig. 4 RMSE of target location estimate versus σ_θ and σ_t with sensor position uncertainty

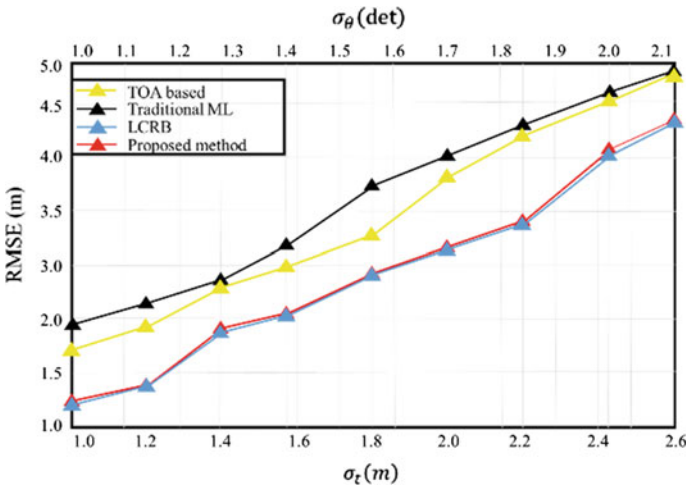


Fig. 5 RMSE of target location estimate with TPs outside the convex hull of error-free sensor nodes locations

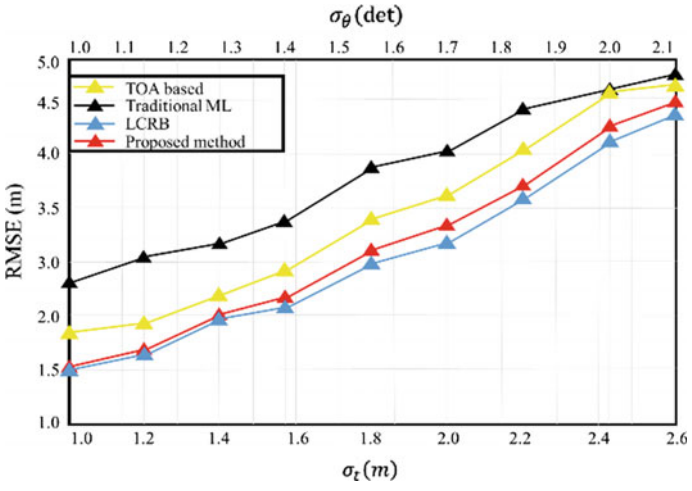


Fig. 6 RMSE of target location estimate with TPs outside the convex hull of erroneous sensor nodes locations

7 Conclusions

In this paper, SDP relaxation-based hybrid TOA/AOA positioning method is firstly proposed and conformed efficient and more robust than TOA based location method. The SDP relaxations for both the error-free and erroneous locations of sensor nodes are given and LCRB are induced in the different scenarios. Simulation results show that the relaxation methods can provide superior performance and matching with the performance predicted by LCRB. The extension of this work to sensor node selection deserves further investigation for higher accuracy and more robustness.

Acknowledgements This work is supported by the National Key Research and Development Program (NO. 2016yfb0502001), National Natural Science Foundation of China (No. 61372110) and National High Technology Research and Development Program of China (No. 2014AA123103).

References

1. Moreno V, Zamora MA, Skarmeta AF (2016) A low-cost indoor localization system for energy sustainability in smart buildings. IEEE Sens J 16(9):3246–3262
2. Sun Z, Purohit A, Chen K et al (2011) PANDAA: physical arrangement detection of networked devices through ambient-sound awareness. In: Proceedings of the 13th international conference on ubiquitous computing. ACM, pp 425–434
3. Yang L, Chen Y, Li XY et al (2014) Tagoram: real-time tracking of mobile RFID tags to high precision using COTS devices. In: Proceedings of the 20th annual international conference on mobile computing and networking. ACM, pp 237–248
4. Zhuang Y, Syed Z, Li Y et al (2015) Evaluation of two WiFi positioning systems based on autonomous crowd sourcing on handheld devices for indoor navigation
5. Faragher R, Harle R (2015) Location fingerprinting with bluetooth low energy beacons. IEEE J Sel Areas Commun 33(11):2418–2428
6. Pagano S, Peirani S, Valle M (2015) Indoor ranging and localisation algorithm based on received signal strength indicator using statistic parameters for wireless sensor networks. IET Wirel Sensor Syst 5(5):243–249
7. Maalek R, Sadeghpour F (2016) Accuracy assessment of ultra-wide band technology in locating dynamic resources in indoor scenarios. Autom Constr 63:12–26
8. Liu K, Motta G, Ma T et al (2016) Multi-floor indoor navigation with geomagnetic field positioning and ant colony optimization algorithm. In: 2016 IEEE symposium on service-oriented system engineering (SOSE). IEEE, New York, pp 314–323
9. Suh J, You S, Choi S et al (2016) Vision-based coordinated localization for mobile sensor networks. IEEE Trans Autom Sci Eng 13(2):611–620
10. Xu R, Chen W, Xu Y et al (2015) A new indoor positioning system architecture using GPS signals. Sensors 15(5):10074–10087
11. Biswas P, Lian TC, Wang TC et al (2006) Semidefinite programming based algorithms for sensor network localization. ACM Trans Sens Netw (TOSN) 2(2):188–220
12. Khan MW, Salman N, Kemp AH et al (2016) Localisation of sensor nodes with hybrid measurements in wireless sensor networks. Sensors 16(7):1143
13. Sturm JF (1999) Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim Methods Softw 11(1–4):625–653

14. Toh KC, Todd MJ, Tutuncu RH (1999) SDPT3—a MATLAB software package for semidefinite programming, Optim Methods Softw 11:545–581
15. Tseng P (2007) Second-order cone programming relaxation of sensor network localization. SIAM J Optim 18(1):156–185
16. Chan FKW, So HC, Ma WK et al (2013) A flexible semi-definite programming approach for source localization problems. Digit Signal Proc 23(2):601–609

Hyperbola Positioning Scheme Based on Continuous Entangled Light and Bell State Direct Measurement

Guan Fang, Chunyan Yang, Yang Bai, Xianglin Wang, Xiang Li and Kun Chen

Abstract Limited by signal power or bandwidth, traditional navigation measures have ceiling hard to reach so that there are some problems of lacking navigation precision, susceptible to interference, insufficient measure distance and so on. With the rapid development of quantum mechanics as well as quantum information theory and the advance of relevant theory and technology, quantum's characteristics of squeezing effect and entanglement are applied to modern navigation and positioning system, which can greatly improve the precision and confidentiality to break the limitation of positioning precision due to restricted radio power and bandwidth. In this way, quantum positioning can be realized with advantages of ultra-precision, long range and strong anti-interference as well as noise resistance. According to that, a hyperbola positioning scheme based on continuous entangled light and Bell state direct measurement is proposed. Based on theory of hyperbola positioning, continuous entangled light is used as transmitting signals and in target receiving terminal Bell state direct measurement system is used to check two ways of continuous entangled light and delay one of them. Then, coherent processing is conducted, and the delay time of the maximum correlation value is the time difference between the two ways of continuous entangled light. According to the time difference, the distance difference can be calculated to obtain a group of hyperbola. In this way, another two base points are chosen to repeat the process above to get the second hyperbola so as to conducting the positioning. By theoretical analysis and parameter estimation, the theoretical positioning precision and error range in given conditions can be achieved.

Keywords Continuous entangled light · Quantum positioning · Hyperbola positioning · Bell state

G. Fang (✉) · C. Yang · Y. Bai · X. Wang · X. Li · K. Chen
Collage of Information and Navigation, Air Force Engineering University, Xi'an 710077, China
e-mail: 984225821@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_49

1 Introduction

As for traditional navigation systems, restricted by the elements like the transmitting power and band width of classical radio signals, their positioning precision faces the technological limitation that cannot be broken. Hyperbolic positioning is a simple and actionable one of traditional navigation methods, but its precision is not high so that it increasingly cannot meet people's needs for high-precision positioning. With the development of quantum theory and technology, quantum entanglement and squeezing effect can be used for precision measurement to improve navigation precision by breaking through shot noise limit (SNL) and even approaching Heisenberg limit (HL). In 2001, the team of Dr. Vittorio Giovannetti from MIT [1–3], firstly came up with the Quantum Positioning System (QPS) based on quantum mechanics and quantum information theory. In 2004, Dr. Thomas B. Bahder explained in detail the method of adopting entangled bi-photon and baseline interferometry to realize precision quantum positioning [4]. Then, Professor Yang Chunyan and others analyzed error factors on the basis and made deep analysis of the measuring method of multi-structural grouping entanglement [5, 6]. But the entangled photons used in this method belonged to separating variable entanglement of low generation efficiency, and the detection by non-ideal detector of relatively low efficiency was difficult to be applied. Comparably, continuous-variable (CV) entangled state was certain entanglement of high generation efficiency, and it was quite easy to realize detection of high quantum efficiency by using the photoelectric converter to detect continuous entangled light. In 2012, Zhu Jun and others from Shanghai Jiao Tong University (SJTU) [7], realized the cm-level measuring accuracy by measuring the second-order correlation characteristic of light field, and then realized hundred-micrometer measurement by utilizing HBT interferometer. In 2016, Wang Xi and others proposed the ranging scheme [8] based on continuous entangled light and balanced homodyne detection [9], which theoretically solved the low efficiency of separating variable entanglement and limited measuring range. However, the experiment of balanced homodyne detection was in need of respectively measuring the local light and the reference light, and the measurement of quadrature amplitude and quadrature phase was not conducted simultaneously, therefore, it was in need of changing relative phases of light and there were several phases to be locked, which made the experiment difficult to implement and had higher requirements for system stability. Zhang Jing and others from the Institute of Optoelectronics of Shanxi University [10] proposed the Bell-state direct detection method that could extract directly and simultaneously two groups of quadrature components without local light, and was merely in need of locking phase once, which made the experiment easier to conduct.

In view of the problems and shortcomings existing in classical radio navigation and quantum ranging method, a kind of hyperbolic ranging difference positioning method based on two-mode squeezed light [11] and bell-state direct state is proposed in this paper: it adopts hyperbolic positioning theory and uses two-mode squeezed lights as measuring signals which are transmitted form two base stations

precisely located to overcome the problem of the low generation and detection efficiency of separating variable entanglement. The two beams of entangled lights reach the target waiting for measuring via different routes. When reaching the receiving end of target, the two light beams can be processed with adjustable delay and their relative phases can be adjusted, and then Bell-state direct detection system is used to simultaneously detect two groups of quadrature components of measuring signal lights and extract their quadrature components to make sum and difference. The delay time of the least amplitude and phase difference is the time difference of arrival so as to measure the optical path difference between the two beams of light and hyperbolic positioning can further be realized.

2 Theoretical Model

2.1 Two-Mode Squeezed Light

Quantum entanglement can be divided into continuous variables entanglement and separating variable entanglement [12, 13], according to Hilbert space dimension of quantum state. The one whose eigenvalue is discrete spectrum is separating variable entanglement, while the one of continuous-spectrum eigenvalue is continuous variable entanglement. And the quadrature component of light field is a typical kind of continuous variable entanglement.

$\hat{X}$ and $\hat{Y}$ respectively represent the quadrature amplitude component and quadrature phase component of single-mode light field. As for any light field states, the product of $\hat{X}$ and $\hat{Y}$ quantum mean square fluctuation satisfies the Heisenberg uncertainty relations:

$$\left\langle \left(\Delta\hat{X}\right)^2 \right\rangle \left\langle \left(\Delta\hat{Y}\right)^2 \right\rangle \geq 1 \tag{1}$$

Here $\left\langle \left(\Delta\hat{X}\right)^2 \right\rangle = \left\langle \left(\hat{X} - \langle\hat{X}\rangle\right)^2 \right\rangle$, $\left\langle \left(\Delta\hat{Y}\right)^2 \right\rangle = \left\langle \left(\hat{Y} - \langle\hat{Y}\rangle\right)^2 \right\rangle$ when equality holds $\left\langle \left(\Delta\hat{X}\right)^2 \right\rangle = \left\langle \left(\Delta\hat{Y}\right)^2 \right\rangle = 1$ and the state of light field is minimum uncertainty state, namely, the coherent state. The quantum fluctuation with the value of 1 is called as shot noise limit (SNL). The fluctuation of classical light quadrature component is always equal or greater than SNL, while the optical squeezed state [10] is non-classical state and squeezed light means one of its quadrature component fluctuations is squeezed to the one less than SNL. When $\left\langle \left(\Delta\hat{X}\right)^2 \right\rangle < 1$, it is the squeezed state of amplitude component. When $\left\langle \left(\Delta\hat{Y}\right)^2 \right\rangle < 1$, it is the squeezed state of phase component. The coherent state and the squeezed state in phase space are shown as Figs. 1 and 2.

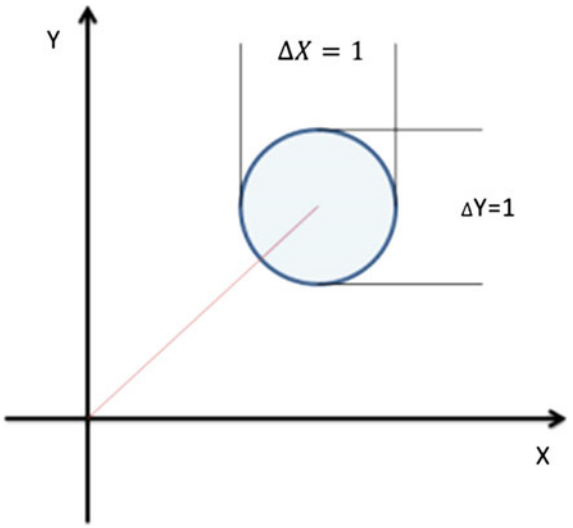


Fig. 1 The expression of the coherent state in phase space

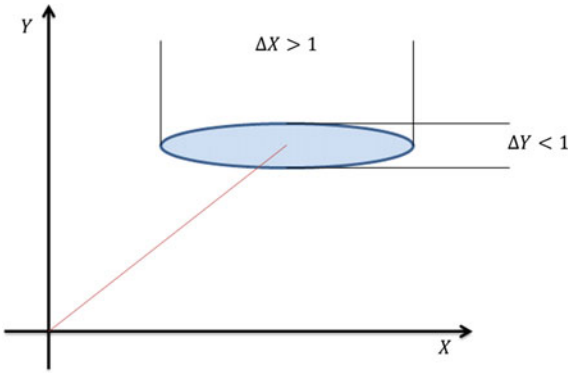


Fig. 2 The expression of the squeezed state in phase space

As for two-mode light field, $\hat{X}_1$, $\hat{Y}_1$ and $\hat{X}_2$, $\hat{Y}_2$ respectively represents its quadrature components. Its amplitude sum and phase difference or amplitude difference and phase sum could be commuting. In such a state, the measurement of quadrature component of one light field can determine the corresponding quadrature component of the other, which introduces the entanglement into two-mode light field. The entanglement state where quadrature amplitudes are positively associated, while quadrature phases are reflexively associated, can be expressed as:

$$\left\langle \left(\Delta\left(\hat{X}_1 - \hat{X}_2\right)\right)^2 \right\rangle = 0, \quad \left\langle \left(\Delta\left(\hat{Y}_1 + \hat{Y}_2\right)\right)^2 \right\rangle = 0 \tag{2}$$

The entanglement state where quadrature amplitudes are reflexively associated, while quadrature phases are positively associated, can be expressed as:

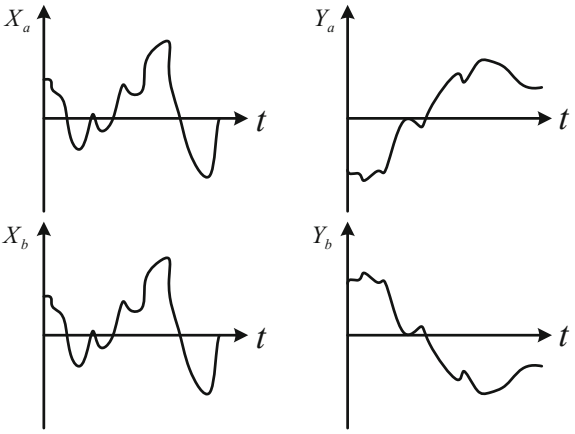


Fig. 3 Diagram of positively associated quadrature phases and reflexively associated quadrature amplitudes

$$\left\langle \left(\Delta\left(\hat{X}_1 + \hat{X}_2\right)\right)^2 \right\rangle = 0, \quad \left\langle \left(\Delta\left(\hat{Y}_1 - \hat{Y}_2\right)\right)^2 \right\rangle = 0 \tag{3}$$

As Fig. 3 shows.

In such a two-mode light field, joint variable noise is squeezed. Thus, the entanglement of two groups of quadrature components is two-mode squeezed state.

3 Production of Two-Mode Squeezed Light

The conversion in nonlinear optical parameters is a typical method to produce the squeezed state [14, 15]. A continuous beam of light with higher frequency is injected as injection light into Nondegenerate Optical Parametric Amplifier (NOPA), and two beams of continuously entangled light with lower frequency are output. Such two beams of light are highly associated so as to show the quantum squeezing characteristic. As Fig. 4 shows.

The a and b are used as annihilation operators of two light field modes:

$$X = \frac{1}{\sqrt{2}}\left(a + a^+ + b + b^+\right) \tag{4}$$

$$P = \frac{-i}{\sqrt{2}}\left(a - a^+ + b - b^+\right) \tag{5}$$

Fig. 4 Diagram of parametric down-conversion

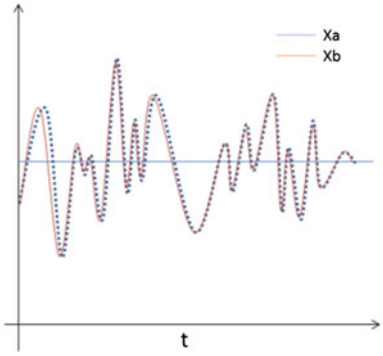


Fig. 5 Quadrature component fluctuations of signals of two beams

They satisfy $[X, P] = 2i$, so $\Delta X \Delta P \geq 1$ is achieved and relative phases can be locked as $\varphi = 0$ by phase locking technology. In this way, two beams of squeezed light where quadrature amplitudes are reflexively associated, while quadrature phases are positively associated are produced.

The quadrature amplitude components, $\hat{X}_a$ and $\hat{X}_b$, of two beams of entangled light, a and b, is taken as an example. In ideal conditions, once squeezing is strong enough to determine the value of $\hat{X}_a$, $\hat{X}_b$ can be calculated, and $\hat{X}_b$'s uncertain region, $\Delta\hat{X}_b$, is quite small. The quadrature amplitude fluctuations of two beams of entangled light are associated to far higher degree than classical association. They can be seen as two truly random signals that are almost the same with each other. As Fig. 5 shows.

Noticeably, the two beams of entangled light are produced simultaneously in the same space. Therefore, when entangled light is applied into navigation as transmitting signal, it can be ensured that its time is totally synchronized in generation end, which is decided by quantum entangle properties. It eliminates the error caused by desynchronized time in traditional radio transmission and positioning precision is further improved.

4 Plan Design

4.1 Positioning Theory

The plan uses two-mode squeezed light as the measurement signal based on the hyperbola locating theory and sets a entanglement source in the geometric midpoint of two base stations. The two generating quadrature phases are positive correlation. The Quadrature amplitude is inverse correlation and the constant entanglement light $\hat{a}$, $\hat{b}$ are transmitted to the terminal A and B respectively and reach the target M to be measured through different path. We delay transaction the first reached beam of

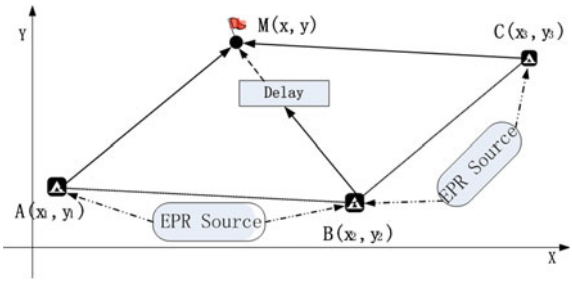


Fig. 6 Schematic diagram of hyperbolic positioning

light (assume beam of light $\hat{b}$ reach first). Then enter the Bell detection system to extract the quadrature component, differ the quadrature phases component and collaborate quadrature amplitude modulation. In an ideal compression strength, when the two beam of lights don not have the light equation, the two corresponding quadrature component is full quantum correlation and the results of the differ and collaborate the quadrature component is infinitely tend to 0. When the optical path difference is greater than some value of number, the two corresponding quadrature components don not have connections, adjusting the delay of beam of light $\hat{b}$ until the results of the differ and collaborate of two quadrature components have the minimum value. The delayed time of $\hat{b}$ is the cost time of beam of light $\hat{b}$ than beam of light $\hat{a}$ for the optical path difference, which can be obtained by $\Delta L = c \times \Delta t$, then we can ensure the double hyperbolic positioning, and choose another two base stations, we can ensure another double hyperbolic positioning. The cross of the hyperbolic positioning can ensure the location of the target. The positioning principle is as Fig. 6.

If the base stations A, B, C are known, which are (x_1, y_1), (x_2, y_2), (x_3, y_3), the coordination of the target to be measured is (x, y), and we can get the distance difference Δd_1, Δd_2 by the corresponding positioning equation:

$$\sqrt{(x_1 - x)^2 + (y_1 - y)^2} - \sqrt{(x_2 - x)^2 + (y_2 - y)^2} = \Delta d_1 \tag{6}$$

$$\sqrt{(x_3 - x)^2 + (y_3 - y)^2} - \sqrt{(x_2 - x)^2 + (y_2 - y)^2} = \Delta d_2 \tag{7}$$

4.2 The Detection Process

When the transmit signal two-mode squeezed light reaches the target, we measure it by the Bell detection system and achieve the extraction of quadrature component by transferring the beam of light to the light current. The Bell is essentially makes use of the coherent measurement of interferometer Mach-Zehnder. The experimental scheme of making use of unbalanced interferometer Mach-Zehnder to detect the

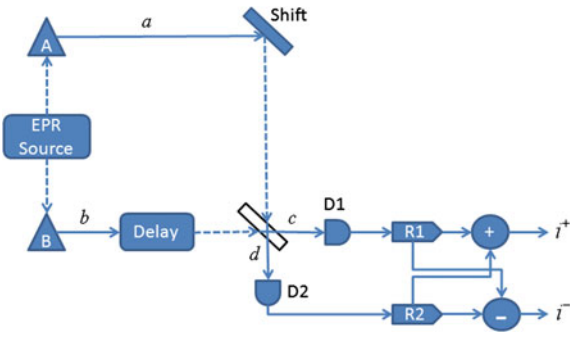


Fig. 7 Schematic diagram of the detection process

pulse squeezed light phase light is put forward by O. Glockl, U. Andersen and so on in 2004. In 2000, Zhang Jing of ShanXi Photoelectric Institute put forward the direct detection method of Bell. Compared to the balanced homophone detection, this method does not need local light and can get the noise of the quadrature amplitude modulation and quadrature phase. The detection process is as Fig. 7.

In the plan, by using the annihilation operator the two-model squeezed light $\hat{a}$ and $\hat{b}$ can be expressed as:

$$\hat{a}(t) = \alpha + \delta\hat{a} \tag{8}$$

$$\hat{b}(t) = \alpha + \delta\hat{b} \tag{9}$$

The two beam of lights are transmitted from different base stations, and by different path to reach the target to be measured, the Bell detection system is responsible for the detection, firstly before the beam of lights $\hat{a}$ and $\hat{b}$ entering the 50:50 beam splitters, they must undergo the phase adjustment, control the phase difference of the light beam is $\pi/2$ (the adjustment is periodical and the adjustment scope is known, the error of optical path difference can be compensated at last), because the entanglement light generated by OPA is orthogonal polarization, we need to use the half glass sheet to make the polarized direction the same, assume light beam $\hat{b}$ reaches first, and make an adjustable delay process which is Δt, the optical path difference of the two beam light when transmitting is ΔL, and input the light field of the beam splitter can be expressed as:

$$\hat{l}_a(t) = \alpha + \delta\hat{a}(t) \tag{10}$$

$$\hat{l}_b(t) = \hat{b}(t-\tau)e^{i\pi/2} = i\alpha + i\delta\hat{b}(t-\tau+\Delta t) \tag{11}$$

The two optical output fields of the light splitter are:

$$\begin{aligned}\hat{c}(t) &= \frac{1}{\sqrt{2}}[\hat{l}_a(t) + \hat{l}_b(t)] \\ &= \frac{1}{\sqrt{2}}[(\alpha + i\alpha) + (\delta\hat{a}(t) \\ &\quad + i\delta\hat{b}(t - \tau + \Delta t))]\end{aligned} \tag{12}$$

$$\begin{aligned}\hat{d}(t) &= \frac{1}{\sqrt{2}}[\hat{l}_a(t) - \hat{l}_b(t)] \\ &= \frac{1}{\sqrt{2}}[(\alpha - i\alpha) + (\delta\hat{a}(t) \\ &\quad - i\delta\hat{b}(t - \tau + \Delta t))]\end{aligned} \tag{13}$$

After two detectors, the two light fields are transmitted to two light current, which can be expressed as:

$$\begin{aligned}\hat{i}_c = c^+(t)c(t) &= \alpha^2 + \frac{1}{2}\alpha[\delta X_a(t) \\ &\quad - \delta Y_b(t - \tau + \Delta t) + \delta Y_a(t) + \delta X_b(t - \tau + \Delta t)]\end{aligned} \tag{14}$$

$$\begin{aligned}\hat{i}_d = d^+(t)d(t) &= \alpha^2 + \frac{1}{2}\alpha[\delta X_a(t) \\ &\quad + \delta Y_b(t - \tau + \Delta t) - \delta Y_a(t) \\ &\quad + \delta X_b(t - \tau + \Delta t)]\end{aligned} \tag{15}$$

$\delta X = \delta a^+ + \delta a$ and $\delta Y = i(\delta a^+ - \delta a)$ are the noise component of the cross polarization and quadrature phase. Then the two light currents are sent to the radio frequency beam splitter and are divided into two parts. Different from the beam splitting, the light current splitters do not have additional noise and by the minus of the corresponding light current can get two electric signal:

$$\begin{aligned}i_+ &= \frac{1}{\sqrt{2}}(i_c + i_d) \\ &= \sqrt{2}\alpha^2 + \frac{1}{\sqrt{2}}\alpha[\delta X_a(t) + \delta X_b(t - \tau + \Delta t)]\end{aligned} \tag{16}$$

$$\begin{aligned}i_- &= \frac{1}{\sqrt{2}}(i_c - i_d) \\ &= \frac{1}{\sqrt{2}}\alpha[\delta Y_a(t) - \delta Y_b(t - \tau + \Delta t)]\end{aligned} \tag{17}$$

The connecting parts of the electric signal are the photoelectric current fluctuation of two quadrature components light current:

$$\delta i_+ = \frac{1}{\sqrt{2}}\alpha[\delta X_a(t) + \delta X_b(t - \tau + \Delta t)] \tag{18}$$

$$\delta i_- = \frac{1}{\sqrt{2}}\alpha[\delta Y_a(t) - \delta Y_b(t - \tau + \Delta t)] \tag{19}$$

So far, we can extract the amplitude and the phase difference of two quadrature component at the same time. Owing to the quadrature component of the constant entanglement lights a, b are highly related, they can be seen as the almost same signal. We can see that when the delayed time Δt is the same as the time of optical path difference of two light path τ, the quadrature phase of two light fields are full quantum positive correlation and the quadrature amplitude is negative correlation. The noise of quadrature phase difference and the quadrature amplitude sum are both less than SNL. In an ideal squeezed intensity, the two values infinitely tend to 0:

$$\delta i_+ = \frac{1}{\sqrt{2}}\alpha[\delta X_a(t) + \delta X_b(t - \tau + \Delta t)] = 0, \quad \Delta t = \tau \tag{20}$$

$$\delta i_- = \frac{1}{\sqrt{2}}\alpha[\delta Y_a(t) - \delta Y_b(t - \tau + \Delta t)] = 0, \quad \Delta t = \tau \tag{21}$$

We can make use of one of the two quadrature components to get the time difference of arrival. The Bell state direct detection used in the plan can simultaneously get the measuring results of two quadrature components. By the comparison of two results, the accountability and accuracy of measuring results can be further enhanced.

5 Performance Analysis

5.1 Estimation of Detection Performance Parameters

Compared to the problem existing in balanced homodyne detection that local light is stronger than signal light so that local light can lead to detector saturation, the two compressed beams in this scheme are used as signal light and their light intensity is almost same, so the result of difference calculation is effective. And the result of Bell state direct detection is the quantum noise of the two sets of quadrature components. So using associated noise to do parameters analysis [16, 17] in this scheme is more straight and more effective. The correlation noise variance of the two-mode compressed light can be expressed by the compressive strength r:

$$\left\langle \delta^2\left(\hat{X}_a + \hat{X}_b\right)\right\rangle = \left\langle \delta^2\left(\hat{Y}_a - \hat{Y}_b\right)\right\rangle = 2e^{-2r} \tag{22}$$

$$\left\langle \delta^2\left(\hat{X}_a - \hat{X}_b\right)\right\rangle = \left\langle \delta^2\left(\hat{Y}_a + \hat{Y}_b\right)\right\rangle = 2e^{-2r} \tag{23}$$

The higher the compressive strength, the smaller the associated noise variance is. When $r \to \infty$, $\left\langle \delta^2\left(\hat{X}_a + \hat{X}_b\right)\right\rangle = \left\langle \delta^2\left(\hat{Y}_a - \hat{Y}_b\right)\right\rangle \to 0$, When the two compressed light is near synchronization in time, it is significantly smaller than the SNL trough, as shown in Fig. 8.

When the correlation noise is less than SNL, the detection precision meets the requirements. And then use the range in which the correlation noise is less than the SNL as the detection error range Δt. The experiment results show that the correlation of the two entanglement signals will disappear completely when the delay time exceeds 27 ns [18], so the delay time error range Δt has to be within 27 ns to meet the effective measurement requirement. As for measuring the least error, analog-digital converter (ADC) should be adopted in practical experiment or simulation experiment to collect correlated noises and they are input into computer for processing to within 27 ns, choose the time interval τ_{min} of two sampling points where minimum values are the closest, and it is the minimum error value of measuring delay time. Therefore, when estimating the minimum error, its precision is limited by AD converter's sampling frequency. The higher its frequency is, the less the minimum error range of measuring delay time will be and the higher precision will be. Currently, the highest sampling frequency of AD converter can be over 2GSPS, so in ideal condition, the minimum error of measuring delay time τ_{min} can be up to picosecond and the measuring accuracy of corresponding optical path difference is up to millimeter.

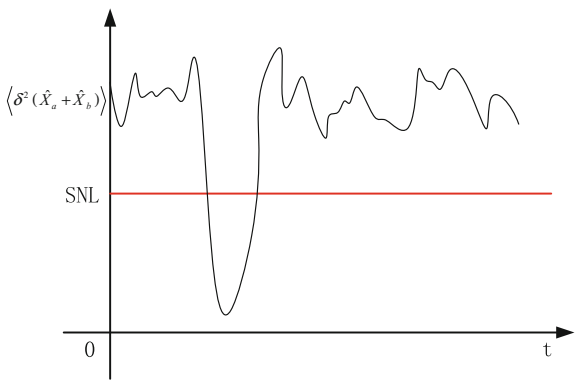


Fig. 8 Schematic diagram of correlation noise measurement results change with the delay time

5.2 Positioning Accuracy Estimation

By estimating the detection error range of the delay time, the optical path difference of the two beams can be estimated, according to the formula:

$$\Delta L = c \times \Delta t \tag{24}$$

Due to the limited application of AD converter of ultrahigh sampling frequency and it cannot be popularly used in most fields, we estimate positioning precision when assuming that the sampling frequency of the detector is 500 MHz. In such conditions, the error range of the time difference of the arrival time is $\Delta t = 2$ ns. In this case, the error range of the optical path difference of the two measuring light is 0.6 m. According to the hyperbola positioning principle, the error range of the positioning is estimated.

As shown in Fig. 9, it is assumed that the straight-line distance between the two base points A, B is known and the optical path difference has been measured and the y-coordinate y of the target point M (x, y) has been determined by another pair of positioning hyperbolas The abscissa x of the target point can be estimated according to the geometric properties of the hyperbola. The standard hyperbolic equation is:

$$\frac{x^2}{a^2} - \frac{y^2}{c^2 - a^2} = 1 \tag{25}$$

Optical path difference $\Delta l = l_1 - l_2 = 2a$ after introducing the optical path difference measurement error, it is rewritten as follows:

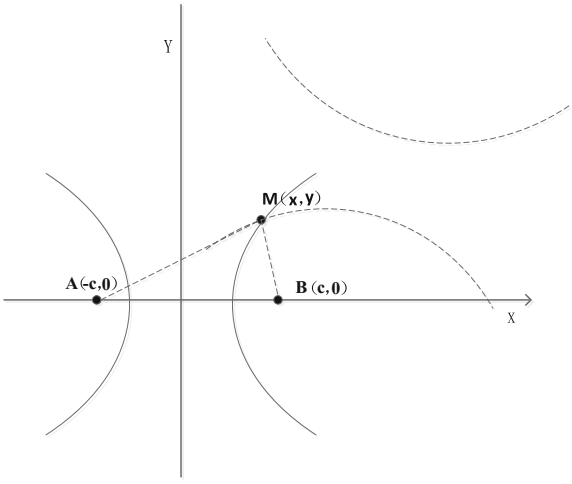


Fig. 9 Positioning principle

$$\frac{x^2}{(a+\alpha)^2} - \frac{y^2}{c^2-(a+\alpha)^2} = 1 \tag{26}$$

At this point the abscissa can be expressed as:

$$x' = \sqrt{\left[1 + \frac{y^2}{c^2-(a^2+\alpha^2)}\right]}(a+\alpha) \tag{27}$$

Then the error range of the abscissa is:

$$\begin{aligned}\Delta x &= x' - x \\ &= (a+\alpha)\sqrt{\left[1 + \frac{y^2}{c^2-(a^2+\alpha^2)}\right]} - a\sqrt{\left[1 + \frac{y^2}{c^2-a^2}\right]}\end{aligned} \tag{28}$$

In practical navigation positioning, the precision is affected by many factors like geometrical relationships among base stations, the length of baseline, the range of detection and so on. So, we need to analyze positioning precision under some given conditions. Assuming that the angle between two baselines is about 90° and the straight-line distance between A and B is 1000 m, which means 2c = 1000, the optical path difference Δl = 600 = 2a. Assuming y = 400 m, then Δx can be set by formula (28) to about 0.849 m. Under the assumption of the same conditions, it is estimated that the error range of ordinates y is Δy = 0.849 m. Therefore, the positioning range of this scheme can be roughly determined within the rectangular range where the horizontal and vertical coordinate errors are 0.849 × 0.849 m^2. As shown in Fig. 10.

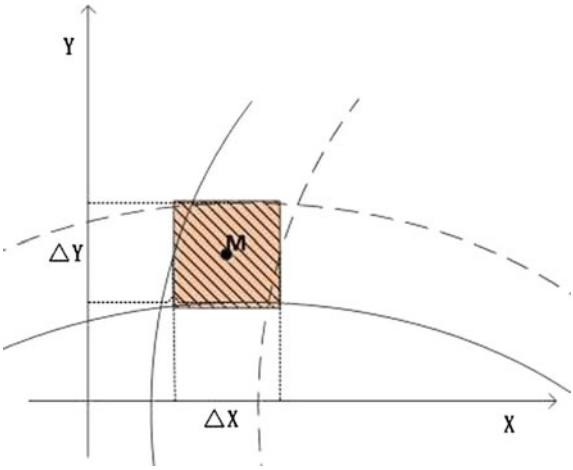


Fig. 10 Estimated positioning range

6 Conclusion

A quantum location scheme based on continuous variable entanglement and Bell state direct detection is proposed. According to theoretical research, under certain conditions, the precision of measurement is directly proportional to the compressive strength of the entangled light. Based on the hyperbolic positioning principle and the existing experimental data, this program gives the theoretically estimated value of positioning accuracy and positioning error range under some given conditions. By optimizing the parameter estimation algorithm and other improvement measures, there is still much room for improvement in this area. Compared with the positioning method of the discrete-variable entangled state, this scheme has the advantages of high efficiency and high detection efficiency. At the same time, compared with the balanced zero-beat detection, the direct detection method of Bell states is easier to implement and the detection process is simpler and more direct. Furthermore, the proposed method is suitable for two beams with the same intensity, which is beneficial to improve the effectiveness. The next step will be simulation experiments, further investigate the impact of various parameters on the positioning accuracy and then analyze. The preliminary theoretical analysis shows that the scheme has some feasibility and can provide a new idea and reference for the application of continuous variable entangled states and quantum navigation and positioning methods.

References

1. Giovannetti V, Lloyd S, Maoccone L (2001) Quantum-enhanced positioning and clock synchronization. Nature 412(6845):417–419
2. Giovannetti V, Lloyd S, Maoccone L (2004) Quantum-enhanced beating the standard quantum limit. Science 306(5700):1330–1336
3. Giovannetti V, Lloyd S, Maoccone L (2002) Positioning and clock synchronization through entanglement. Phys Rev A 65(2):022309
4. Bahder TB (2004) Quantum positioning system [EB/OL] Cornell, Tex. Cornell University Library. arXiv:quant-ph/0406126
5. Yang C, Wu D, Yu Y et al (2011) Method of enhancing the measurement of the time of arrival by utilizing multi-structured grouped-cetin-glad quantum pulse. J Beijing Univ Posts Telecommun (Natural Science Edition) 34(6):34–38 (in Chinese)
6. Yang C, Wu D, Yu Y, Zhao X (2009) Research on optinal constellation distribution of interferometric quantum positioning system, Acta Geod Cartogr Sin 393(12):1–6
7. JunZhu YuanLu, Huang Peng, Liu Jingneng, Zeng Guihua (2011) Clock synchronization through second-order coherence of thermal light. AIP Conf Proc 13(63):117
8. Wang X, Chen S, Wu D, Yang C, Chen K, Li X (2016) Quantum ranging scheme based on two-mode squeezing light. Acta Opt Sin 36(7):0727001
9. Liu K (2012) Generation of continuous-variable entanglements and conditional cloning. Shanxi University, Taiyuan, pp 12–14
10. Zhang J, Peng K (2000) Quantum teleportation and dense coding by means of bright amplitude-squeezed light and direct measurement of a Bell state. Phys Rev A 62:064302

11. Sun H, Liu K, Zhang J et al (2015) Quantum precision measurement based On squeezed light. Acta Phys Sin 64(23):210–234
12. HorodcckiR. HorodcckiP. HorodcckiM.etal. (2007) Quantum entanglement, Rcv Mod Phys 81(2): 865–942
13. Braunstein SL, van Loock P (2005) Quantum information with continuous variables. Rev Mod Phys 77(2):513–577
14. Ou ZY, Pereira SF, Kimble HJ et al (1992) Realization of the Einstein-Podolsky-Rosen paradox for continuous variables. Phys Rev Lett 68(25):3663–3666
15. Villar AS, Cruz LS, Cassemiro KN et al (2005) Generation of bright two-color continuous variable entanglement. Phys Rev Lett 95(24):243–263
16. Jia X, Duan Z, Yan Z (2013) Dependence of continuous variable entanglement enhancement on experimental parameters. Acta Opt Sin 33:030006
17. Shen Y, Deng X, Jia X (2011) Improvement of continuous variable entanglement enhancement experiment. Acta Sin Quantum Opt 17(2):113–118
18. Marino AM, Pooser RC, Boyer V et al (2009) Tunable delay of Einstein-Podolsky-Rosen entanglement. Nature 457(7231)
19. Zhang Y, Yu X-D, Di K, Li Wi, Zhang J (2013) Looking the phase of balanced homodyne detection system for squeezed light. Acta Phys Sin 62:084204
20. Durhuus T, Joergensen C, Mikkelsen B et al (1994) All optical wavelength conversion by SOA's in a Mach-Zehnder configuration. IEEE Photonics Technol Lett 6(1):53–55
21. Glöckl O, Andersen UL, Lorenz S et al (2004) Sub-shot-noise phase quadrature measurement of intense light beams. Opt Lett 29(16):1936–1938
22. Hétet G, Glöckl O, Pilypas KA et al (2006) Squeezed light for bandwidth limited atom optics experiments at the Rubidium D1 line. J Phys B: At Mol Opt Phys 40(1):221–226

Application of the Multipath Hemispheric Model in Monitoring Receiver

Tianxiang Su, Yuanlan Wen, Wei Feng, Guangming Liu, Hao Lan, Jiangwei Ma, Yun Zhang and Xu Zhang

Abstract Multipath effect is the primary factor affecting the quality of GNSS observation data, and thus affects the measurement accuracy of monitoring receivers. Giving the characteristics of the monitoring receiver, the application of the multipath hemispherical model is used to the three types of satellites of the BeiDou system and the model precision is analyzed in this paper. The experimental indicates that the multipath hemispherical model has a good effect on the multipath improvement of monitoring receiver.

Keywords Multipath · Hemispherical model · Monitoring receiver

1 Introduction

With the rapid development of GNSS, errors such as satellite ephemeris error, clock error, ionospheric error, and tropospheric error can be further reduced by modeling. But for the multipath error, it is mitigated difficultly owing its own particularity, has become a major error affecting GNSS precise positioning. There are three main methods to reduce multipath: receiver antenna design, receiver baseband signal processing and data post-processing methods. It has an obvious effect that reducing the selection of receiver antenna type and the proper installation environment. The receiver baseband signal processing can also suppress the multipath effect by using narrow correlation and other processing techniques. However, after adopting these two techniques, multipath effects are not completely eliminated, and multipath errors still exist in the observation data [1, 2]. The method of the data post-processing can be applied to estimate the multipath error by modeling and filtering online or offline, so as to further weaken the multipath effect.

In recent years, more and more attention has been compensated to the method of further reducing the multipath effect at the data processing level, and scholars at

T. Su (✉) · Y. Wen · W. Feng · G. Liu · H. Lan · J. Ma · Y. Zhang · X. Zhang
No. 143, Leiguang Road, Haojiang District, Shantou, Guangdong, China
e-mail: tianxiang_su@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_50

home and abroad have conducted extensive research on the analysis of multipath error. Feng Xiaochao et al. made a relative experiment on the multipath effect of the pseudorange measurement of GNSS receiver, and carried on the experimental analysis of the related factors which influence the multipath effect [3]. Zhang Liyun, Li Guowei studied the modeling of the pseudorange multipath error of GPS system by using spherical function and trigonometric series respectively, and comparing and analysing the error accuracy of multipath model fitting [4, 5]. Wu Xiaoli et al. proposed CNMC algorithm to mitigate multipath of GEO satellite of BeiDou system, improving the pseudo-range positioning accuracy [6]. Axelrad et al. proposed a SNR-based weighting algorithm to reduce the multipath effect [7]. Azarbad et al. modeled multipath signals by using wavelet analysis and Vondrak filters to reduce multipath errors [8, 9]. Larson et al. established the ARTA adaptive filter by analyzing the time series of the positioning error for GPS system satellites, and the multipath errors were removed [10]. Zhongping et al. proposed a Sidereal filter based on the observed data to make a good estimate of the multipath error [11]. Ye Shirong et al. studied the Carrier Phase Multipath of the BeiDou GEO, IGSO, MEO and other satellites, and achieved very good results [12]. Guo Jinyun, Song Le, Shi Qiang et al. had established a multipath hemispherical model for non-difference, single-difference and double-difference observations respectively, and achieved good results [13–15].

In this paper, the mechanism of multipath generation is described firstly. And then the characteristics of multipath effects is analysed. Finally, the hemispherical multipath model is used to mitigate multipath error on the monitor receiver, and further experiments verify the hemispheric multipath model reliability and prediction accuracy.

2 Multipath Effect Analysis

2.1 Multipath Effect

Multipath is the measurement error caused by the receiver receiving the direct signal transmitted by the satellite while also receiving other indirect reflected signals, interference of the indirect transmitted signal to the direct signal, resulting in distortion of the received signal. The principle of multipath is shown in Fig. 1.

2.2 Repeatability Analysis of Satellite

BeiDou satellite navigation system is composed of GEO satellite, IGSO satellite and MEO satellite, the orbital period also has a significant difference. The calculation method of the orbit overlap period is given [10]:

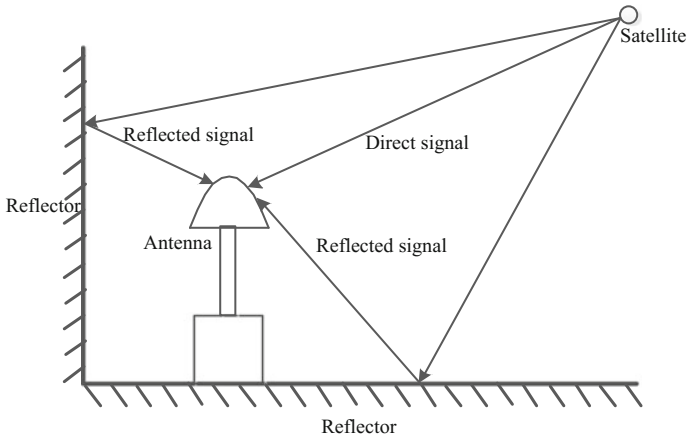


Fig. 1 The principle of multipath signal

$$n = \sqrt{\frac{GM}{a^3}} + \Delta n \tag{2.1}$$

$$T = 4\pi/n \tag{2.2}$$

The orbital overlap period of GEO and IGSO satellites is about one day, while that of MEO is about 7 days [12, 15].

2.3 *Analysis of Monitoring Multipath Receiver*

Monitoring receiver is used for ground monitoring stations to collect the original data of the navigation satellite, is the major equipment of ground control system for data acquisition, which data is foundation for satellite orbit determination, ephemeris generation, differential correction and various model parameter calculation, has an irreplaceable position. The precision of the monitoring receiver is high, and its higher performance required for continuity, stability and reliability [16].

Due to the limitation of the location requirements of the monitor station [17], the monitoring receiver environment is relatively fixed, which will not change in the short term, and monitoring receivers are static receivers, thus the relative environment of the receiver is relatively unchanged, which benefits for the multipath modeling of receiver.

As the multipath effect leading the phase distortion and amplitude distortion only relates to the elevation and azimuth, but not relating the satellite, the multipath effect of the monitoring receiver is equal when the different satellites cross the same location. The multi-path effect has the same space-time invariance, the size of the

multipath effect of the same altitude and azimuth is the same for different satellites, and does not change with time and specific satellite [13].

In this paper, the pseudo-range carrier combination method is used to calculate the multipath error value [6]. Suppose that MP1, M2 are B1, B2 multipath, ρ_1, ρ_2 and φ_1, φ_2 are B1, B2 pseudorange measurements and carrier phase measurements respectively, λ_1, λ_2 are B1, B2 wavelengths. The multipath error calculation formula is as follows:

$$MP_1 = \rho_1 + \frac{1+\alpha}{1-\alpha}\varphi_1 \cdot \lambda_1 - \frac{2}{1-\alpha}\varphi_2 \cdot \lambda_2 \tag{2.3}$$

$$MP_2 = \rho_2 + \frac{2\alpha}{1-\alpha}\varphi_1 \cdot \lambda_1 - \frac{1+\alpha}{1-\alpha}\varphi_2 \cdot \lambda_2 \tag{2.4}$$

where, $\alpha = \frac{f_1^2}{f_2^2}$.

Taking GEO-5, IGSO-3 and MEO-1 as examples, the multipath errors are shown in Figs. 2, 3 and 4 respectively.

It can be seen from Figs. 2, 3 and 4 that the trend of multipath effect of GEO-5, IGSO-3 and MEO-1 is basically the same on 27th and 28th. Figure 4 illustrates the MEO-1 satellite's early entry in a week, mainly because its orbital overlap is not an entire week.

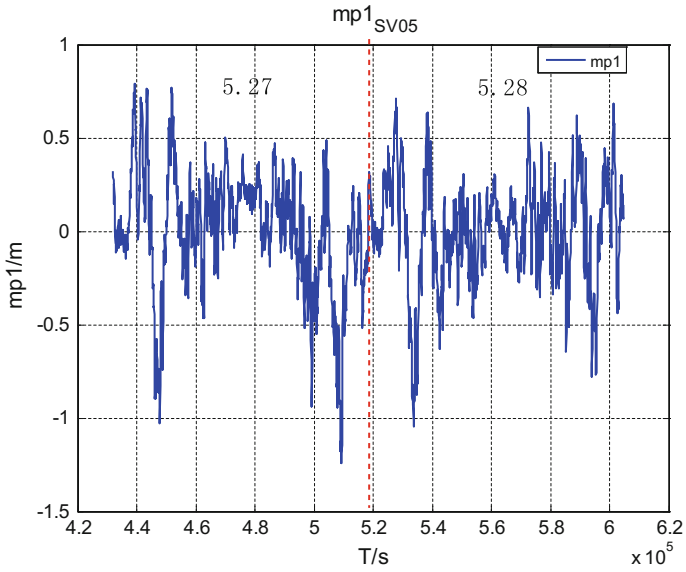


Fig. 2 The multipath error of GEO-5

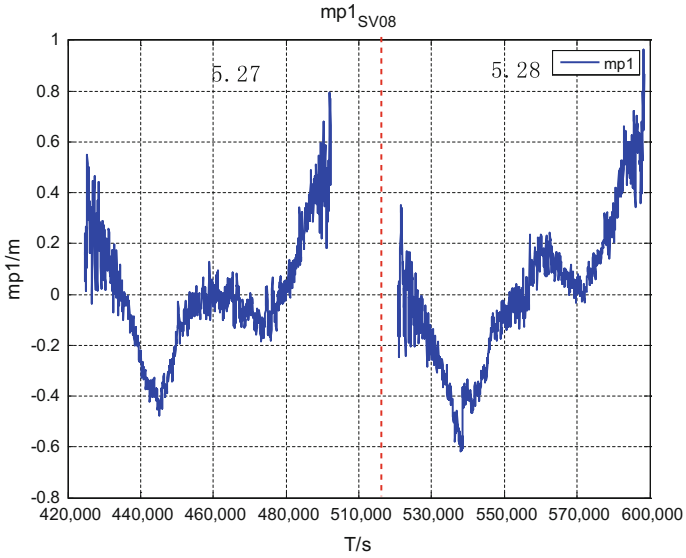


Fig. 3 The multipath error of IGSO-3

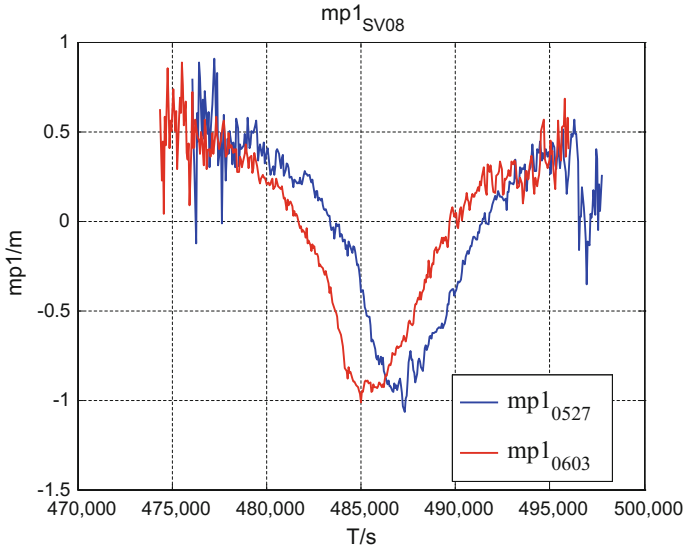


Fig. 4 The multipath error of MEO-1

3 Multipath Hemispherical Model

The receiver multipath effect is only related to the elevation and azimuth of the satellite. Therefore, the multipath model can be established by dividing the multipath into elevation and azimuth.

The multipath hemispherical model is the statistical average of the multipath, which is means of the multipath of all the satellites passing through the same position. Because of the different orbital heights of the three types of satellites, the changes of the elevation and azimuth are different, so the meshing can be treated differently according to different types of satellites.

3.1 Multipath Triangular Series Modeling

There are many methods of modeling Multipath, such as the polynomials, the spherical harmonics and trigonometric series at el. The trigonometric model is applied in this paper. θ is the elevation and ϕ is the azimuth. The trigonometric series of the multipath can be expressed as follows:

$$MP(\theta,\phi) = a_0 + \sum_{i=1}^{N} [a_i \cos(i\theta) + b_i \sin(i\theta) + c_i \cos(i\phi) + d_i \sin(i\phi)] \tag{3.1}$$

where, N is the order, a_i, b_i, c_i, d_i is the undetermined parameter.

3.2 Fitting Algorithm

In this paper, the least squares algorithm is used to solve the unknown parameters,

$$\begin{bmatrix} MP_1 \\ MP_2 \\ \vdots \\ MP_m \end{bmatrix} = AX + \varepsilon \tag{3.2}$$

where A is the matrix of coefficients, m is the number of observations, determined by the sampling frequency, X is the parameter to be solved, the dimension of X is determined by the order of triangular series, namely (4N + 1) × 1 dimension, $N(0,\sigma_\rho)$ is Gaussian white noise, and σ_ρ is the range error.

$$X = \left(A^T A\right)^{-1} A^T \cdot MP \tag{3.3}$$

Since the elevation and angle of elevation change slowly between neighboring observations, the correlation matrix A is strong, which leads to a serious ill-condition in the process of solving, and the method of ridge estimation is used [18].

$$X = (A^T A + K \cdot I)^{-1} A^T \cdot MP \tag{3.4}$$

where, K ridge parameter is 1e−10, which is a unit matrix.

3.3 The Prediction of Hemispherical Map Model

After the parameters are calculated from Sect. 3.2, the receiver multipath is predicted according to the $0.1 \times 0.1^\circ$ grid.

$$\overline{MP} = B \cdot X \tag{3.5}$$

where $\overline{MP}$ is the predicted value of multipath, B is the coefficient matrix which contains the sine and cosine functions of elevation and azimuth, and X is the estimated parameter.

4 Multipath Effect Analysis

In this paper, BeiDou monitoring receiver is adopted. The data acquisition location is located in Guangdong Province from May 27th to June 3rd, 2016, sampling frequency is 1 s and elevation cut-off is 15°. The multipath error is analyzed for GEO-5, IGSO-3 and MEO-1 in this paper.

The celestial projection of multipath error for GE0-5, IGSO-3 and MEO-1 on monitoring receiver is respectively shown in Figs. 5, 6 and 7. In the figure, the value corresponding to the color is the value of the multipath error.

Since GEO satellites are stationary relative to the earth, the changes in elevation and azimuth are small, as shown in Fig. 5. While the IGSO satellite and MEO satellite are movable relative to the earth, the elevation and azimuth change faster than GEO satellite. Theoretically, the higher the elevation angle at which the receiver receives the satellite signal, the smaller the multipath error should be. The IGSO-3 satellites have different multipath effects on the east and west sides. When the IGSO-3 satellite is on the east side of the receiver, the IGSO-3 satellite can basically meet the trend of higher elevation angle and smaller multipath. However, when IGSO-3 satellites are located on the west side, the multipath effect decreases and then increases with the rising elevation angle, and then becomes smaller, which is not consistent with the conventional trend.

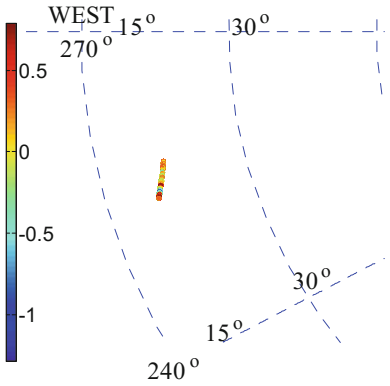


Fig. 5 The sky map of GEO-5 multipath error

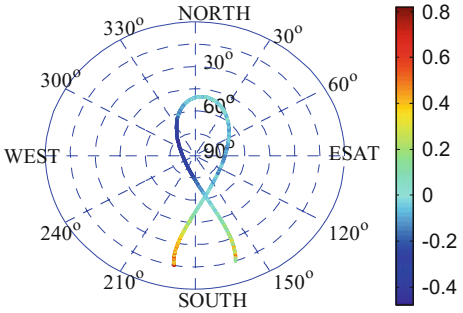


Fig. 6 The sky map of IGSO-3 multipath error

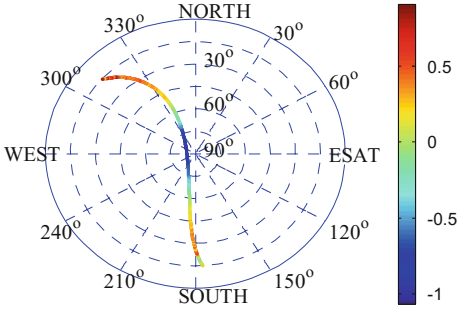


Fig. 7 The sky map of MEO-1 multipath error

In Fig. 7, this abnormal trend is more obvious for MEO satellites. When MEO-1 arrives, the elevation angle is very small, thus the multipath is very large. As the elevation angle becomes larger, the multipath decreases, but when the MEO satellites are at a high elevation angle in the west of the monitoring receiver, multipath does not decrease but becomes larger. As MEO satellite gradually exits, the elevation angle is low, multipath effect gradually becomes larger.

In general, it is believed that the higher the elevation angle, the smaller the multipath error. This situation is based on the concept which the receiver is in the ideal environment. That is, the environment of the receiver in the horizontal direction is substantially symmetrical. The multipath curves in Figs. 6 and 7 show that the multipath error is also related to the azimuth angle. Therefore, in the process of multipath modeling, elevation angle and azimuth angle as two important parameters, the number is equal, which is different from the trigonometric series modeling [5]. Several pools on the west side of the monitoring receiver are likely to cause larger multipath of the IGSO-3 and MEO-1 at high elevation.

It can be seen from Fig. 8 that the hemispheric multipath fitting curve can reflect the multi-path trend, while the forecast model only in a few obvious changes in the place of the forecast is better, other parts of the forecast did not achieve the desired effect. The trend is basically consistent with the multi-path trend. Figures 9 and 10

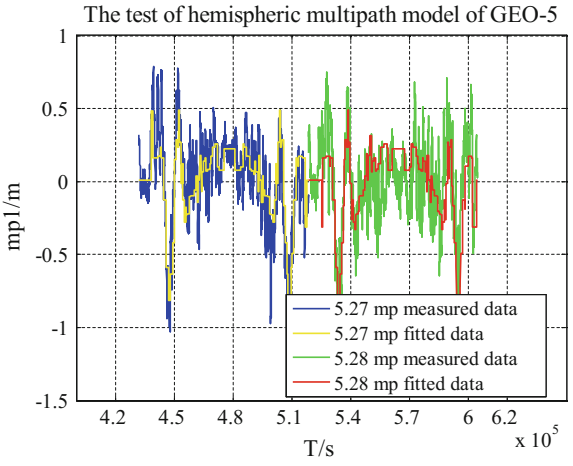


Fig. 8 The test of hemispheric multipath model of GEO-5

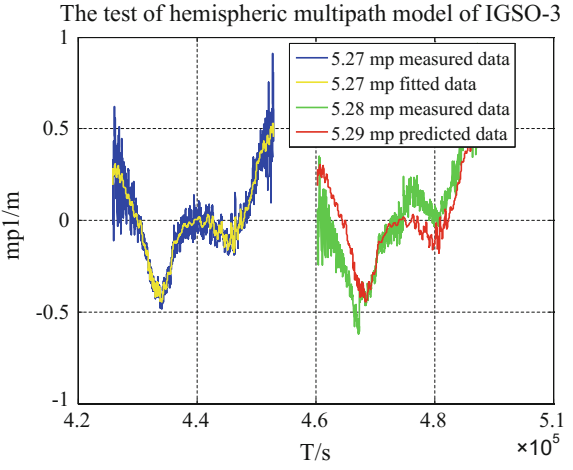


Fig. 9 The fitting error curve of GEO-5

also indicate that the prediction of the hemispheric multipath of GEO-5 is not very satisfactory. The main reason is that the elevation and azimuth change is too small, leading to the strong correlation between the observed data, thus resulted that the model parameter calculation does not reflect the trend of change, but at several important inflection point forecast model is better.

It can be seen from Figs. 11, 12 and 13 that the hemispheric multipath model has a good fitting effect on IGSO-3 satellite, and the model forecast can reflect the multi-path trend. It can be seen from Fig. 12 that the effect of hemispherical multipath model fitting is good. While in Fig. 13, although the trend of multipath effect is reflected in the whole model forecasting process, but the prediction error

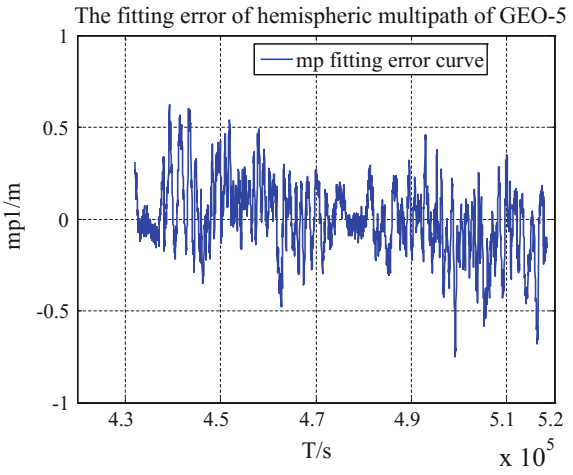


Fig. 10 The prediction error curve of GEO-5

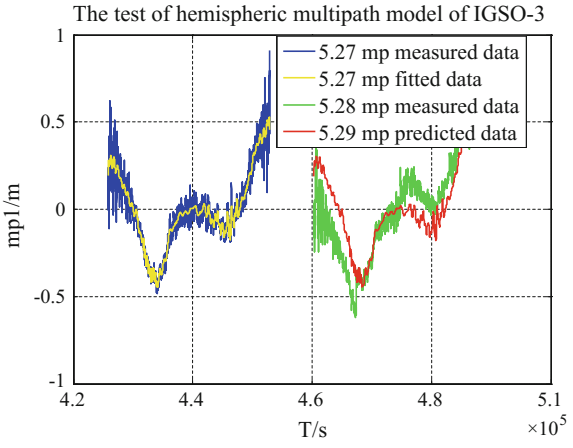


Fig. 11 The test of hemispheric multipath model of IGSO-3

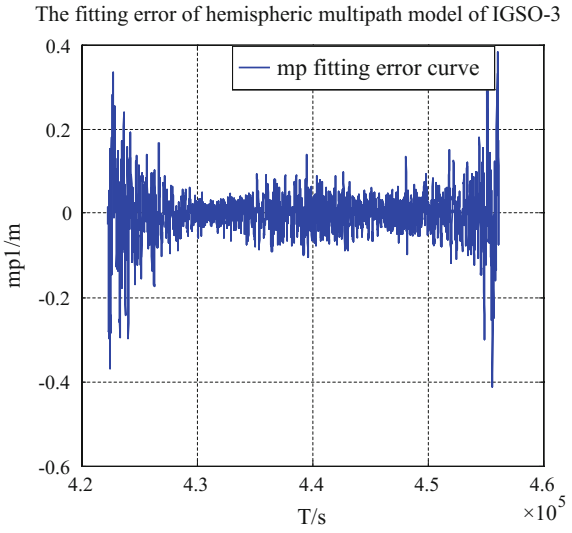


Fig. 12 The fitting error curve of IGSO-3

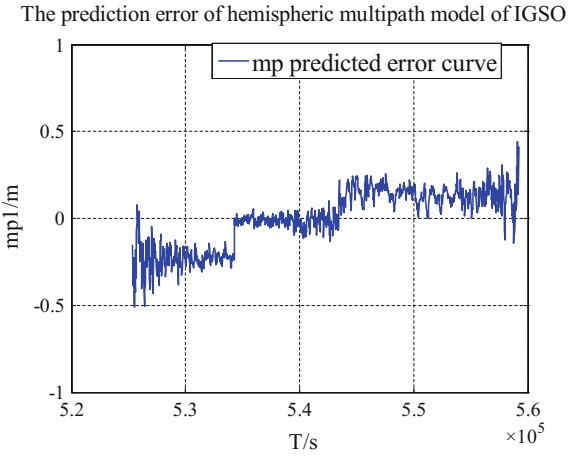


Fig. 13 The prediction error curve of IGSO-3

behaves that the first half is larger, the latter half is smaller, there is a systematic bias. This may be related to the adjustment of satellite position on that day (Fig. 14).

Figures 15, 16 and 17, it can be seen that the application of the hemispherical multipath model in MEO-1 is better; the trend of the multipath is reflected in both the fitting curve and the forecast curve.

As can be seen from Table 1, the effect of the hemispherical multipath model is not the same for the three types of satellites, the applicability for the IGSO and MEO satellite is better than GEO satellite in general. Due to the high orbit of GEO

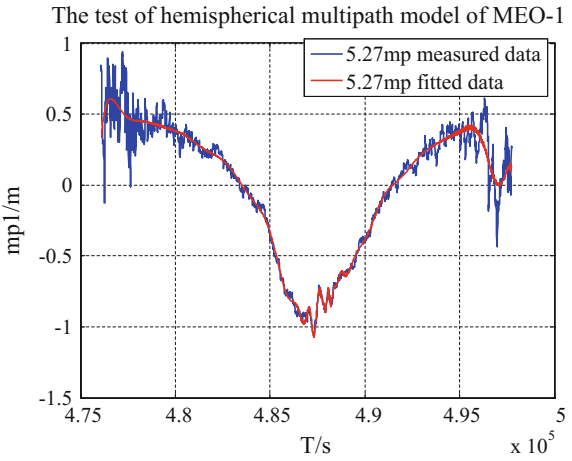


Fig. 14 The fitting curve of hemispheric model of MEO-1

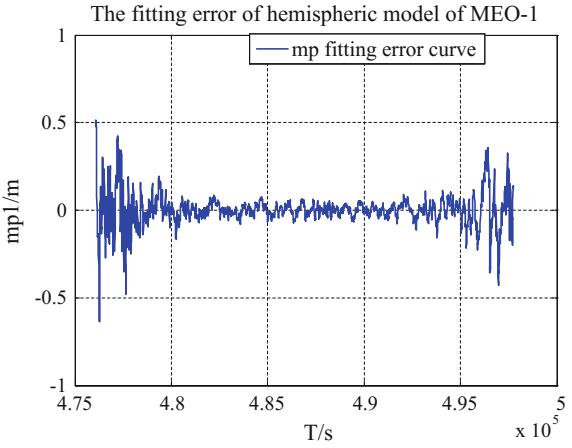


Fig. 15 The fitting error curve of hemispheric model of MEO-1

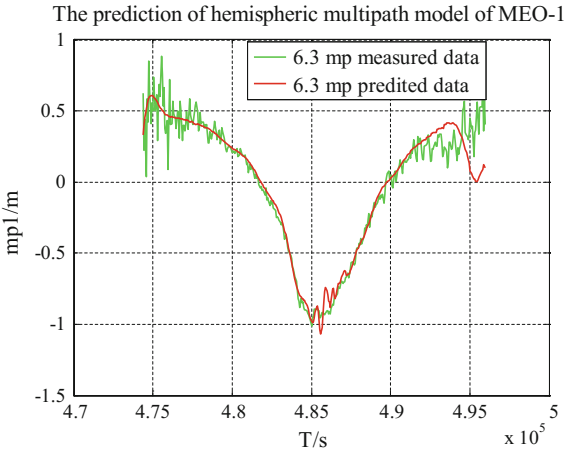


Fig. 16 The prediction of hemispheric multipath model of MEO-1

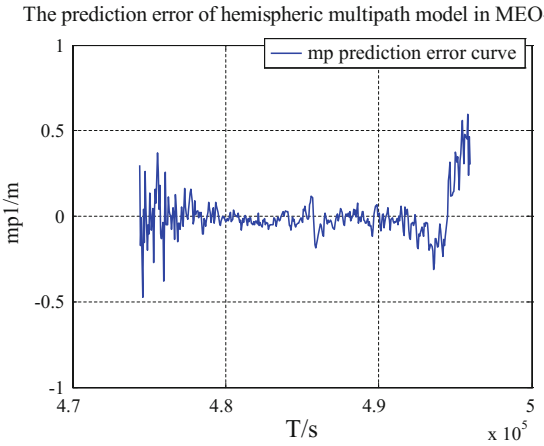


Fig. 17 The prediction error curve of hemispheric multipath model of MEO-1

Table 1 The applicability of the hemispherical multipath model to three types of satellites

Satellite	Fitting accuracy	Prediction accuracy
GEO-5	0.2076	0.2437
IGSO-3	0.0622	0.1654
MEO-1	0.0911	0.1248

satellites, the elevation and azimuth change slowly, leading to the strong correlation of the data. However, there is a weak correlation with the observation data of the IGSO and MEO satellites, because the changes of the elevation and azimuth are quicker than GEO.

In addition, the order of trigonometric series is not directly proportional to the accuracy of multipath prediction, but when the order is more than certain degree, the accuracy of model prediction will decrease gradually.

5 Conclusions

This paper preliminary discusses the application of the hemispherical multipath model in non-difference data of the BeiDou monitoring receiver. The multipath errors of code-reduced carrier are analyzed for GEO, IGSO and MEO, and the multipath errors are predicted according to the model. The experiment shows that the hemispherical multipath model has some feasibility in non-difference data of the monitoring receiver. However, the base data of the parameter estimation only takes 1 day, and the accuracy of the monitoring receiver angle can only reach 0.1°, and the change of elevation angle and azimuth angle are small, especially for the GEO satellite, resulting that the change of multipath can't be reflected though the calculated parameters. The next step is to continue to optimize the order of the

trigonometric series and to explore different multipath modeling methods for the three types of satellites to further refine the model parameters and to more accurately model the multipath errors.

References

1. Xie G (2011) Principle of GPS and receiver design. Publishing House of Electronics Industry, Beijing, pp 384–386
2. Kaplan ED, Hegarty CJ (2008) Understanding GPS: principles and applications. Publishing House of Electronics Industry, Beijing, pp 184–190
3. Feng X, Jin G, Fan J et al (2013) Experimentation and analysis of multipath effect in pseudo-range measurement of GNSS receiver. Mod Electron Tech 05:77–81
4. Li G, Guo J, Yuan Y et al (2013) Studies of modeling multipath effect at GPS station. Sci Surveying Mapp 38(3):7–9
5. Zhang L, Huang W, Ming D et al (2015) Simulation modeling of multipath effect for satellite navigation monitoring station receiver based on a segmentalization technique. J Geodesy Geodyn 35(1):106–110
6. Xiaoli W, Shanshi Z, Gang W et al (2012) Multipath error detection and correction for GEO/IGSO satellites. Science China (Physics, Mechanics & Astronomy) (7):1297–1306
7. Axelrad P, Larson K, Jones B (2005) Use of the correct satellite repeat period to characterize and reduce site-specific multipath errors
8. Azarbad MR, Mosavi MR (2014) A new method to mitigate multipath error in single-frequency GPS receiver with wavelet transform. GPS Solutions 18(2)
9. Satirapod C, Rizos C (2005) Multipath mitigation by wavelet analysis for GPS base station applications. Surv Rev 295(38):2–10
10. Duncan C, Larson K (2007) Finding the repeat times of the GPS constellation. GPS Solution 1(11):71–76
11. Zhong P, Ding X, Yuan L et al (2010) Sidereal filtering based on single differences for mitigating GPS multipath effects on short baselines. J Geodesy 2(84):145–158
12. Ye S, Chen D, Liu Y et al (2015) Carrier phase multipath mitigation for BeiDou navigation satellite system. GPS Solutions 4(19):545–557
13. Song L (2015) Application of multipath hemispherical model for short baseline multipath effect elimination. East China Normal University
14. Guo J, Li G, Qiao L et al (2014) Modeling GPS multipath effect based on spherical cap harmonic analysis. Trans Nonferrous Met Soc China (6):1874–1879
15. Shi Q, Dai W, Zeng F et al (2016) The BDS multipath hemispherical map based on double difference residuals and its application analysis. Changsha
16. Su T, Feng X, Ma J et al (2016) Overview of testing and evaluation technology for satellite navigation monitoring receiver. Changsha, Hunan Province, China
17. Gong L, Ma X, Feng X et al (2014) Specification for the address-choosing and construction of GNSS monitoring station. Digit Commun World 02:10–13
18. Zhang J, Cai H (1995) Aircraft test statistic. National University of Defence Technology, Changsha

The Test and Analysis on Pulse Signal Detection Abilities of the X-ray Detector MCP for Pulsar Navigation

Qingyong Zhou, Lizhi Sheng, Ziqing Wei, Siwei Liu, Kun Jiang, Chun Chen, Jianfeng Ji, Hongfei Ren and Gaofeng Ma

Abstract The Micro-Channel Plate detector (MCP) is an X-ray detecting system with a mature technology, which is suitable for space observation and navigation application of X-ray pulsar. For analyzing the X-ray pulse signal observation capability of MCP detector, a long-time experiment of MCP detector with different radiation flux and different background noise were carried out by the ground testing system. Meanwhile, a set of evaluation methods of the pulse observation capability of X-ray detector was established, those expressions of the signal-to-noise ratio (SNR), the correlation of pulse profile (R), the accuracy of pulse time of arrival (σ_{TOA}), the minimum detectable power $(P_{\min})$ are derived from the photon counting model of the ground testing system. In the whole testing experiment which took almost a week, 8 groups of 10000 s X-ray photon TOAs dataset were received by the 20 cm^2 MCP detectors, the best periods were searched and gained, the observable pulse profiles were replicated, those characteristic parameters of pulse profile were estimated. The result of this experiment shows that the MCP detector has a good capability of X-ray pulse signal observation and reconstruction, the SNR, R, σ_{TOA} of the observed pulse profile at the low pulse signal flux $\left(f_p = 0.05\,\mathrm{ph/cm^2/s}\right)$ are (35.73, 88.38%, 51.53 μs) and are (35.73, 88.38%,

Q. Zhou (✉) · G. Ma
Geographic Spatial Information Institute, Information Engineering University, Zhengzhou 450052, China
e-mail: zjlzqy1986@163.com

Q. Zhou · Z. Wei · S. Liu · J. Ji · H. Ren
Xi'an Research Institute of Surveying and Mapping, Xi'an 710054, China

Q. Zhou · Z. Wei · S. Liu · J. Ji · H. Ren
State Key Laboratory of Geo-information Engineering, Xi'an 710054, China

L. Sheng · C. Chen
Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China

K. Jiang
Beijing Tracking and Communication Technology Research Institute, Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2017 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 438,
DOI 10.1007/978-981-10-4591-2_51

51.53 μs) at strong background noise (B = 16), so, the MCP detector has a good observation ability of a certain faint X ray pulsar. The result also has revealed that the minimum detectable power $(P_{\min})$ of MCP detector is closely related to the power intensity of the X-ray pulsar simulator, is less correlation with the intensity of background noise. The $P_{\min}$ has the minimal value which is also called the minimum detectable cutoff power. When the X-ray pulse signal received by MCP detector is weak, $P_{\min}$ is close to its minimal value (1×10^{-14} W).

Keywords Micro-channel plate (MCP) detector · Ground testing system · X-ray pulsar simulator

1 Introduction

The pulsar is a type of neutron star with high-speed rotation and strong magnetic-filed, whose periodic pulse signal is extremely stable and time characteristics is clear, just like the "Lighthouse" in the universe. So, by using X-ray millisecond pulsar as a natural and permanent super stable signal source, the autonomous navigation of spacecraft in deep space can be achieved for long time. The X-ray pulsar navigation is such a process that the photons from pulsar are received by X-ray detectors and then the pulse times of arrive are estimated, the orbit and clock errors of spacecraft are determined by fusing with the orbit dynamic information. In general, X-ray pulsar navigation system always includes the X-ray detector, the pulsar parameter database, satellite clock, auto-navigation software and other components, especially, the X-ray detector is a key component of the pulsar navigation system, whose sensitivity of detection directly affects X-ray signal detection ability, thus affects the space observation ability and navigation application performance for X-ray pulsar.

The MCP detector is a kind of detector for X-ray pulsar space observation or navigation application. It has the advantages of high time measurement accuracy and large-scale integration. At present, a large number of performance tests of the MCP detector has been carried out. The relevant test results show that the signal of X-ray photon can be acquired well and read out accurately by the MCP detector. Because the radiation flux of the X-ray millisecond pulsar used in space navigation is weak, and the space particle background noise is strong, so testing and evaluating the pulse detection ability of MCP detector for X-ray pulsar in weak pulse signal or strong background noise environment is very important. For analyzing the X-ray pulse signal observation capability of MCP detector, a long-time experiment of MCP detectors under different radiation flux and different background noise were carried out by the ground testing system which is described in reference [1]. This paper just reports the result of this testing experiment of MCP detector.

2 The Ground Testing System and Evaluation Method

2.1 MCP Detector and Ground Test System

Micro-channel plate (MCP) is a two-dimensional array with a small channel composed of electronic multiplication devices, with millions of parallel electron multiplication micro-channel. Two MCP can achieve 10^5–10^6 times the electron gain. The working principle of MCP detector is shown in Fig. 1, When the X-ray photons pass through a collimator, and will produce photoelectrons with photocathode in photoelectric effect, and then repeat collisions with the inner wall of the passage, the final output of the large number of electrons are received by the anode. The electric pulse is formed by the fast current sensitive amplification, constant ratio timing CFD (Constant Fraction Discriminator) and time digital conversion circuit TDC (Time-to-Digital Converter) to obtain the time of arrive of photons.

The experiments on the pulse signal detection abilities of the X-ray detector MCP are carried out on X ray detector ground testing system. The schematic figure of X ray detector ground testing system is shown in Fig. 2, the X-ray ground testing system mainly includes two parts: X-ray pulse signal modulation and generation part and X-ray photon receiving and processing part. X-ray pulse signal modulation and generation part produces the discrete X ray photon beam which can recover the same profile by X-ray pulsar simulation source (grid controlled X ray tube here) with pulse signal. The X-ray photon receiving and processing part receives X-ray photons by MCP detector, and deals with those data.

The generation and reception of the whole X-ray photons are in the vacuum pipeline of the ground testing system, the length of vacuum pipeline is 1.2 m, and the test environment temperature is 20 °C. The time-frequency stability of system is about 10^{-10}. The more information of X ray pulsar simulation source and ground test system please refers to reference [2, 3].

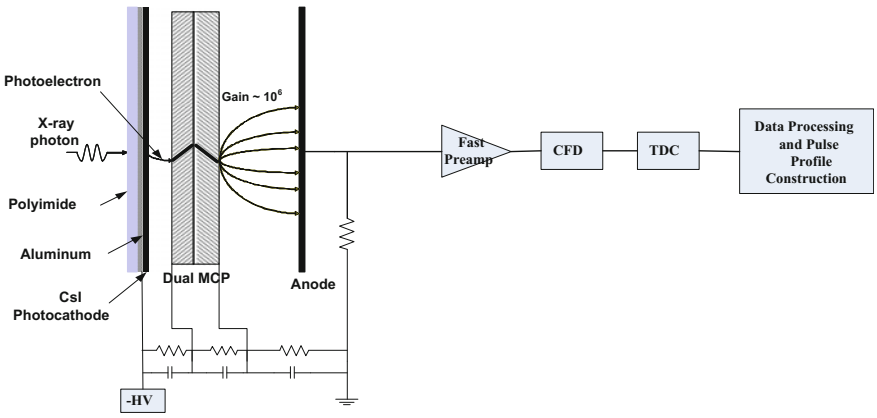


Fig. 1 Schematic of photon detecting and readout for MCP detector[3]

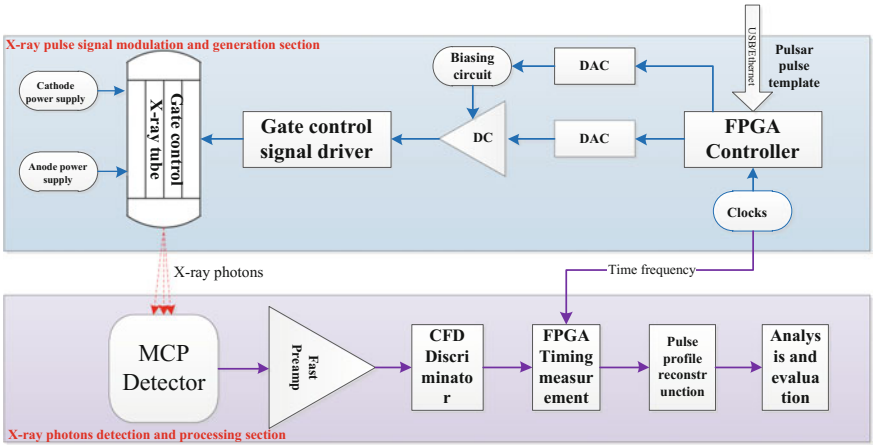


Fig. 2 The sketch map of X-ray detector ground testing system

2.2 *The Evaluation Method*

The X-ray pulsar simulation source can control the radiation power of the grid controlled X-ray tube and modulate the X-ray photons with controllable flow and period by loading the pulse contour template on the host computer. The simulated X-ray pulsar for testing is a typical navigation pulsar Crab, whose profile template is obtained by dealing with RXTE satellite spatial observation data and whose modulation pulse period is 33.4 ms. Figure 4 shows the normalized standard profile of Crab pulsar, the part above red line is considered as pulse photon counts, N_{Pulse} X-ray photons from the pulsar, the other part below red line is considered as non-pulse photon count $N_{non-Pulse}$, generally from the pulsar cloud or other space background radiation. Define the background noise intensity $B = N_{non-Pulse}/N_{Pulse}$, indicating the intensity of the background noise.

The MCP detector receives discrete the X-ray photons and records its arrival time. Ideally, we can recover the best observation pulse profile by using modulated pulse periods. The reconstruction of accurate pulse profile is very sensitive to the precision of period value. However, due to the effect of the self-noise of MCP detector, photon limited and other unknown factors, the pulsar rotation period will change a little, so the value of pulse period need to be researched for obtaining the best pulse period, then the high-SNR pulse profile is replicated. The periodic search algorithm adopts the quantity χ^2 evaluation method.

The pulse detection capability of MCP detector can be evaluated by the quality of observing pulse signal and the minimum detectable power. The former includes the similarity of observation pulse, signal to noise ratio (SNR) and TOA accuracy, the latter describes the minimum detection ability of MCP detector. Those equations can be derived from the photon counting model.

According to the working principle of the ground testing system, the photon counts of MCP detector output mainly includes three parts: (1) X-ray pulse photon counts N_{Pulse}, (2) X-ray non-pulse photon counts $N_{non-Pulse}$, always are considered as radiation from the nebula of pulsar and space particle background (3) background counts N_b from MCP detector itself. The counting statistics caused by the latter two kinds of photon are both random white noise.

$$N = N_{pulse} + N_{non-Pulse} + N_b \tag{1}$$

Assuming that the geometric effective area of the MCP detector is A and the effective observation time is t. Three kinds of photon flux density can be estimated below:

$$\begin{aligned} f_{Pulse} &= N_{Pulse}/t/A \\ f_{non-Pulse} &= N_{non-Pulse}/t/A \\ f_b &= N_b/t/A \end{aligned} \tag{2}$$

where $A = 20\,\mathrm{cm}^2$, $\mathrm{t} = 10000\,\mathrm{s}$. The SNR of the observation signal is usually expressed as the ratio of the pulse photon counts of source to the observation error σ_{noise} of the signal.

$$\begin{aligned} SNR &= \frac{N_{Pulse}}{\sigma_{noise}} = \frac{N_{Pulse}}{\sqrt{(N_b + N_{non-Pulse}) + N_{Pulse}}} \\ &= \frac{f_{Pulse}}{\sqrt{(f_b + f_{non-Pulse}) + f_{Pulse}}} \cdot \sqrt{A \cdot t}\big|_{A=20} \end{aligned} \tag{3}$$

The SNR of observational pulse signal received by MCP detector in the pulse mode also can be analyzed with time changes by Eq. (3).

The minimum detectable power $P_{\min}$ of MCP detector, also called the noise equivalent power (NEP), is defined as the incident X-ray power when SNR of pulse signal gained by per unit area (1 cm^2) X-ray detector in per unit time (1 s) is 1. The minimum detectable power is related to the incident X ray photon energy. In the experiments, the characteristic energy spectrum is at 4.5 keV.

$$\begin{aligned} P_{\min} &= P(t = 1\,\mathrm{s})\big|_{SNR=1, A=1\,\mathrm{cm}^2} = hv \cdot f_{Pulse}/\eta \\ &= 4.5 \times 10^3 \times 1.602 \times 10^{-19} f_{Pulse}/\eta \end{aligned} \tag{4}$$

where $h = 6.63 \times 10^{-34}\,\mathrm{J\,s}$ is the Planck constant, v is the frequency of X ray photon, $7.2 \times 10^{-16}\,\mathrm{J}$ is the corresponding energy of the Ti target spectrum 4.5 keV, η is the detection efficiency. The detection efficiency of the MCP detector at the 4.5 keV energy is about 9.15%. Inserting $SNR = 1$ into Eq. (3), and the reasonable term is maintained:

$$f_{Pulse} = \frac{1 + \sqrt{1 + 4((f_b + f_{non-Pulse}))}}{2} \tag{5}$$

Combination Eqs. (4) and (5), the new formula can be simplified below:

$$P_{\min} \approx 4 \times 10^{-15}(1 + \sqrt{1 + 4 \cdot (f_b + f_{non-Pulse})})\mathrm{W} \tag{6}$$

where W is the power unit watts. It is thus clear that MCP detector at 4.5 keV energy has the minimum detectable power. When there is no noise, the minimum detectable power $P_{\min} \approx 8 \times 10^{-15}$ W. The background noise of X ray detector is closely related to the operating environment temperature and the output voltage threshold. The background noise of MCP detector is strong, its background flow density is $f_b \approx 0.3$ ph/cm^2/s, so $P_{\min} \approx 1 \times 10^{-14}$ W. Equations (3) and (6) are the calculation formulas of the pulse signal SNR and the minimum detectable power of the MCP detector respectively.

The formulation for calculating correlation between the observation pulse profile and the loaded standard pulse profile:

$$R \equiv \max\{\mathrm{R}(\tau)|\tau\} \tag{7}$$

where $R(\tau) = \frac{\int_{-\infty}^{+\infty} p(\mathrm{t}) \cdot f(\mathrm{t}+\tau)\mathrm{dt}}{\sqrt{\int_{-\infty}^{+\infty} p^2(\mathrm{t})\mathrm{dt} \cdot \int_{-\infty}^{+\infty} f^2(\mathrm{t})\mathrm{dt}}}$, $p(\mathrm{t})$ is the standard pulse profile, see in Fig. 3, $f(t)$ is observational pulse profile.

The precision of Pulse time of arrive (TOA) can be estimated by pulse profile characteristic, which represents the resolution of pulse arrival time based on single observation:

$$\sigma_{TOA} = FWHM/SNR \tag{8}$$

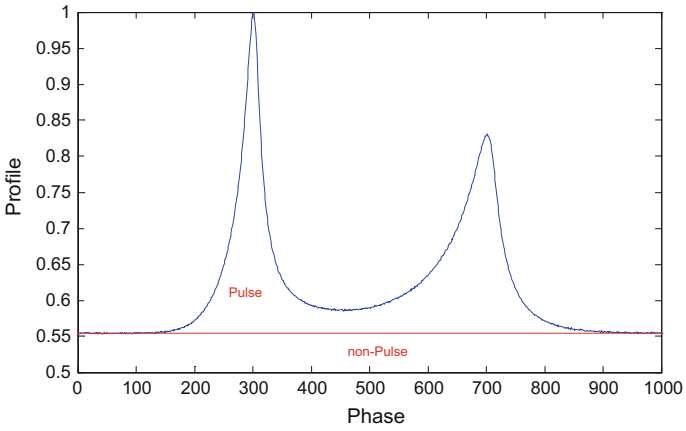


Fig. 3 The united standard profile of Crab pulsar

Table 1 The different parameters of MCP detector which reflect the pulse signal detection ability

I (μA)	B	N (ph)	f (ph/cm^2/s)	f_{Pulse} (ph/cm^2/s)	$f_b + f_{non-Pulse}$ (ph/cm^2/s)	P_B (Hz)	SNR	$P_{\min}$ ($\times 10^{-15}$ W)	R (%)	σ_{TOA} (μs)
65	1	4,209,261	21.05	3.4	17.65	29.9403720998	331.79	37.85	95.44	14.72
50	1	1,826,109	9.13	3.16	5.97	29.9404387727	467.20	23.95	98.72	6.70
25	1	903,345	4.52	1.96	2.55	29.9404178995	413.27	17.39	99.08	5.62
5	1	198,937	0.99	0.26	0.73	29.9404141866	117.17	11.92	95.12	13.12
1	1	75,481	0.38	0.05	0.33	29.9403931065	35.73	10.09	88.38	51.53
25	4	922,084	4.61	1.08	3.53	29.9403903122	225.80	19.55	96.84	7.43
25	8	947,729	4.74	0.62	4.12	29.9403917319	126.70	20.72	93.72	19.78
25	16	913,485	4.57	0.31	4.26	29.9403757422	64.89	20.99	89.72	29.24

where, FWHM is the half bandwidth of the observational pulse profile, and the FWHM of Crab pulsar accounts for about 9.1% of the whole period.

3 Experimental Results and Analysis

Two types of MCP detectors have been tested by ground testing system to evaluate the ability of X ray pulse signal acquisition and recovery of MCP detector with different radiation flow and background noise intensity. The first kind of tests fixed the background noise intensity B = 1, adjusted the X ray radiation intensity by changing the current of X-ray tube, five groups of different radiation flux experiments, in which the corresponding I were 65, 50, 25, 5, 1 A, have been carried out. The second kind of tests fixed radiation flux (I = 25 A), changed the background noise intensity B by loading different pulse profile template, three sets of experiments were conducted in which the background noise intensity B were 4, 8, 16. The observation time of each experiment is 10000 s, The normalized pulse profile loaded by the pulsar simulation source is Crab.

The MCP detector was placed at 80 cm from the X ray pulsar simulation source, the vacuum chamber vacuum degree remains at 10^{-4} Pa, the test environment temperature is 20 °C.

In the experiment, the MCP detector can receive the data about the time of arrival of each photon under different radiation flux and different background noise intensity. The method to evaluate pulse signal in Sect. 2.2 was adopt. For each datasets of observation, the pulse period was searched, and the observable pulse profile was estimated by the estimated value, seen in Fig. 3. The pulse flux and non-pulse flux were calculated, the SNR of observation pulse, the pulse correlation R, σ_{TOA} were estimated, the minimum detectable power of the MCP detector was analyzed, the related results are shown in Table 1.

From Table 1 and Fig. 4, The MCP detector has a good X-ray pulse signal ability of acquisition and procession, can recover observed pulse profile under weak pulse photon flux and high background noise intensity, the main conclusions are summarized as follows:

(1) The observable pulse profile at the weakest radiation flux 0.05 ph/cm^2/s can be recovered by MCP detector with area of 20 cm^2 in 10000 s, seen in Fig. 5e. the SNR of pulse profile is up to 35.73, R is 88.38%, σ_{TOA} is 51.53 μs. For high background noise, MCP detector also has good pulse signal detection ability. When the B = 16, in other words, the background noise intensity is 16 times than the pulse flux, the observable pulse profile also can be recovered, the SNR, R, σ_{TOA} of observable pulse profile are 64.89, 89.72%, 29.24 μs respectively. So, the MCP detector can realize the space observation of Crab, PSR B1509 and other typical pulsar for navigation. If the area of MCP detector can be enlarged and the observation time can be prolonged, the detection ability of pulse signal can be improved.

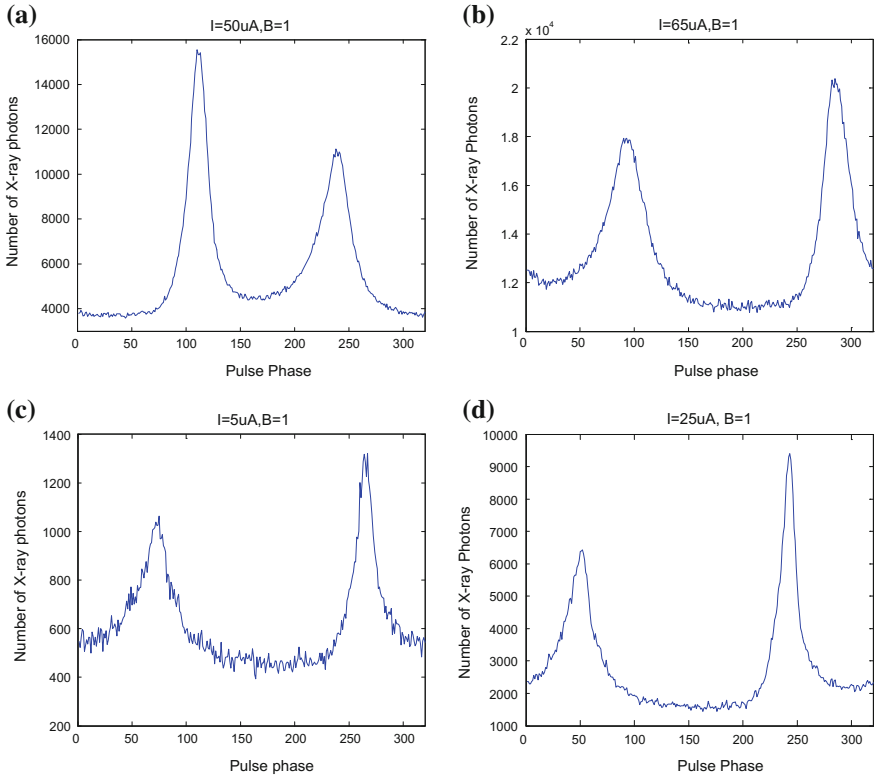


Fig. 4 The 8 observed pulse profiles from experiment. **a** I = 65 μA, B = 1; **b** I = 50 μA, B = 1; **c** I = 25 μA, B = 1; **d** I = 5 μA, B = 1; **e** I = 1 μA, B = 1; **f** I = 25 μA, B = 4; **g** I = 25 μA, B = 8; **h** I = 25 μA, B = 16

(2) The SNR of observable pulse profile raises with the increase of current intensity I of X-ray simulation source, decreases with the increase of the background noise intensity B, both relationship are nonlinear. The minimum detectable power $P_{\min}$ of MCP detector is closely related to the current intensity I, which is independent of the background noise intensity B. When the flux of pulse signal received by the MCP detector (e.g. I = 1 A) is low, the $P_{\min}$ of the MCP detector approaches its extreme value 1×10^{-14} W. the SNR of observable pulse profile basically raises with the increase of I, and decreases with the increase of B. although with strong pulse flux and weak background noise, the pulse correlation R of observable pulse profile is better than 90%. However, when with weak pulse flux (I = 1 μA) and strong background noise (B = 16), both R of two observable pulse profile is less than 90%. The precision of pulse time of arrive σ_{TOA}, which is used to describe the accuracy of observational pulse profile, is inversely proportional to the SNR of pulse profile. The σ_{TOA}

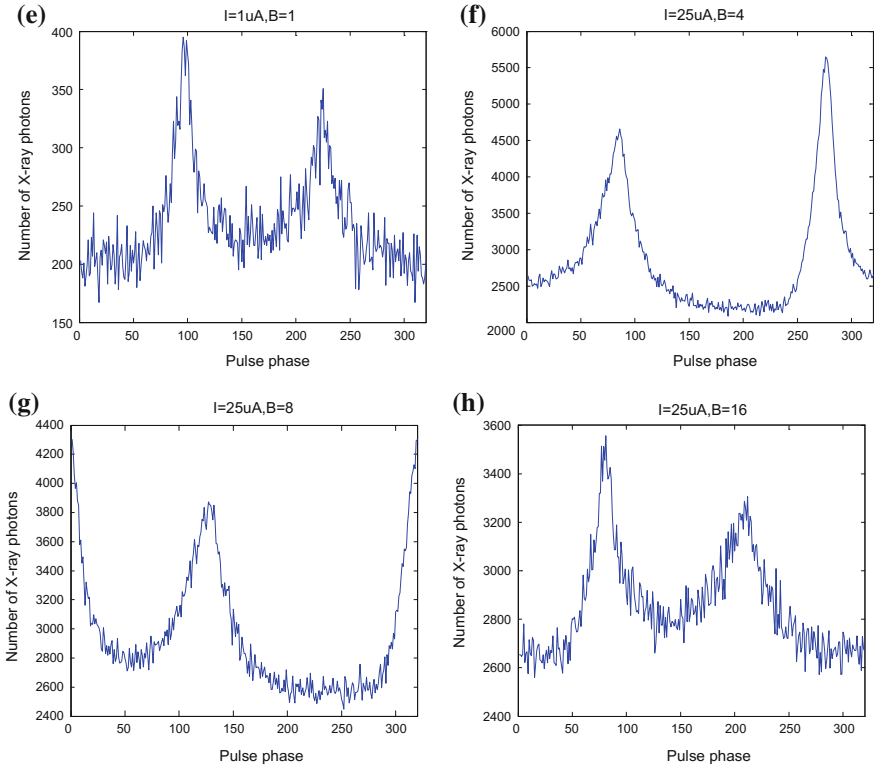


Fig. 4 (continued)

improves with the increase of current I of X-ray simulation source, and decreases with the increase of the background noise intensity B. The σ_{TOA} of the observable pulse profile which is received by 20 cm^2 MCP detector in 10000 s is better than 100 μs in low pulse flux (I = 1 A) or strong background noise (B = 16), this result also means that the MCP detector can be applied for pulsar navigation.

4 Summary and Discussion

Through a long-time test of MCP detector in different X-ray radiation flux and different background noise intensity, MCP detector has good X-ray pulse signal observation and recovery capability, can obtain the pulse profile with high SNR and R. Both of two observable pulse profile which received by 20 cm^2 MCP detector with low flux ($f_p = 0.05\,ph/cm^2/s$) or high noise (B = 16) are recovered in this test, whose characteristic parameters of pulse profile SNR, R, TOA are (35.73,

88.38%, 51.53 μs), (64.89, 89.72%, 29.24 μs) respectively, those results proved that the MCP detector has a faint X ray pulsar observation ability. In the real space environment detection, pulsar radiation flow is very weak, even the strongest Crab pulsar is only 1.54 ph/cm^2/s, and the flux of other pulsar is usually below 10^{-2} ph/cm^2/s. Due to the influence of celestial motion, solar activity and geomagnetic field disturbance, the space environment of spacecraft is very complicated, and there are many kinds of background radiation particles which changes with time. The volatility of the spatial dispersion of X-ray background may have a great influence on the pulsar navigation. The tests in this paper only consider a constant space background noise. In further, we can construct a radiation model based on the real space environment data, and assess the impact of MCP detector from space X-ray background with the mathematic simulation method.

References

1. Sheng LZ, Zhao BS, Zhou T et al (2013) Performance of the detection system for X-ray pulsar based navigation. Acta Photon Sin 42(9):1071–1076
2. Hu HJ (2012) Research of MCP-based X-ray photon counting detector and key technology for X-ray pulsar navigation. University of Chinese Academy of Sciences, Beijing
3. Sheng LZ (2013) Research of key technologies for X-ray pulsar simulation and X-ray detector. University of Chinese Academy of Sciences, Beijing
4. Hanson JE (1996) Principles of X-ray navigation. Department of Aeronautics and Astronautics, Stanford University, Stanford
5. Sheikh SI (2005) The use of variable celestial X-ray sources for spacecraft navigation. Department of Aerospace Engineering, University of Maryland, Maryland
6. Wang YD, Zheng W, Sun SM et al (2013) X-ray pulsar-based navigation system with the errors in the planetary ephemerides for Earth-orbiting satellite. Adv Space Res 51:2394–2404
7. Tremsin AS, Siegmund OHW (1999) Spatial distribution of electron cloud footprints from microchannel plates: measurements and modeling. Rev Sci Instrum 70(8):3282–3288
8. Siegmund OHW, Tremsin AS, Vallerga JV et al (2003) High resolution cross strip anodes for photon counting detectors. Nucl Instrum Meth A 504:177–181
9. Saito M, Saito Y, Asamura K (2007) Spatial charge cloud size of microchannel plates. Rev Sci Instrum 78:023302
10. Chen BM, Zhao BS, Hu HJ et al (2011) X-ray photon-counting detector based on a micro-channel plate for pulsar navigation. Chin Opt Lett 9(6):060404
11. Song J, Zhao BS, Sheng LZ et al (2014) Research on shared anode used for large area array MCP detector. Acta Photon Sin 43(8):0823002
12. Song J, Zhao BS, Sheng LZ et al (2015) Selection of MCP for array X-ray pulsar navigation detector. Opt Precis Eng 23(2):402–407
13. Song J (2014) Research on key technology of MCP X-ray detector used for XNAV. University of Chinese Academy of Sciences, Beijing
14. Sheng LZ, Zhao BS, Wu JJ et al (2013) Research of X-ray pulsar navigation simulation source. Acta Phys Sin 62(12):129702
15. Ge MY (2012) The X-ray emission of pulsars. Graduate University of Chinese Academy of Sciences, Beijing
16. Zhou QY, Ji JF, Ren HF (2013) Quick search algorithm of X-ray pulsar period based on unevenly spaced timing data. Acta Phys Sin 62(1):019701