STATE RESEARCH CENTER OF THE RUSSIAN FEDERATION
CONCERN CSRI ELEKTROPRIBOR, JSC

32nd SAINT PETERSBURG INTERNATIONAL CONFERENCE ON INTEGRATED NAVIGATION SYSTEMS



3 – 4 June 2025

Saint Petersburg, Russia

CO-SPONSORED BY:

- **INTERNATIONAL PUBLIC ASSOCIATION – ACADEMY OF NAVIGATION AND MOTION CONTROL (ANMC)**
- **NATIONAL RESEARCH UNIVERSITY ITMO**
- **CHINESE SOCIETY OF INERTIAL TECHNOLOGY (CSIT)**
- **JOURNAL GYROSCOPY AND NAVIGATION**

This is a collection of the plenary and poster papers of the 32nd Saint Petersburg International Conference on Integrated Navigation Systems (3–4 June 2025).

Editor-in-Chief

Academician of RAS
Vladimir G. Peshekhonov

ISBN 978-5-91995-110-0

© State Research Center of the Russian Federation
Concern CSRI Elektropribor, JSC, 2025

SESSION I – INTEGRATED SYSTEMS

PLENARY PAPERS

1.O.A. Stepanov, A.V. Motorin, V.P.Zolotarevich
Bayesian Estimation Algorithms for Solution of SLAM Problem 9

2.Zhuolin Wang, Yingjun Zhang, Haoze Zhang
Cooperative Image Enhancement for Vision-Based MASS Navigation: A MADRL Approach 14

PLENARY PAPER

3.P.A. Khmarski, A.O. Naumov
Adaptive Estimator of Coordinates and Motion Parameters of a Maneuvering Object Based on Data from a Fixed Bearing Sensor 22

POSTER PAPERS

4.N.N. Vasilyuk
Astronomical Calibration of a Strapdown Astroinertial Navigation System using a Simple Testbed 24

5.V.A. Pogorelov, S.V. Sokolov
The Principle of Building a Tightly Coupled Inertial Optical Navigation System 30

6.M.S. Selezneva, K.A. Neusypin
Evaluating of the Reliability of the Results for a Flight Experiment with an Inertial Navigation System 35

7.M.S. Selezneva, K.A. Neusypin
Increasing the Degree of Parametric Identifiability of an Error Model for an UAV Navigation System 38

8.Haoze Zhang, Yingjun Zhang, Zhuolin Wang
Performance Evaluation of Autonomous Berthing System for Unmanned Boat 41

9.A.A. Frolov, V.B. Pudlovskiy, I.S. Silvestrov
Test Method for INS/GNSS Receivers 47

10.Ye.G. Kharin, A.V. Yasenok, V.A. Kopelovich, E.B. Gorskiy
General Approach to Methods of Flight Tests for Airborne Flight and Navigation Equipment 51

11.O.A. Amosov, S.G. Amosova
Adaptive Method for Estimation and Navigation of Unmanned Vehicles Using Intelligent Technologies 56

12.A.F. Scherbatyuk
About a Method of AUV Group Positioning Using the Hydro Acoustical Navigation System with a Short Base Line 60

13.L.A. Martynova
Method of AUV Coordinates Correction during Observation in the Northern Latitudes 63

14.I.M. Kuznetsov, K.K. Veremeenko, M.V. Zharkov, A.N. Pronkin, R.Y. Zimin
Navigation System Neural Network Architecture Parameters Preformance Influence Investigation 67

15.N.I. Tsiopliakis
Fast Block LD Kalman Filter for Visual-Inertial Navigation Tasks 73

16.M.V. Kurdeneva, D.Kh. Mukhametzyanov
Numerical Stability of Various Discrete Kalman Filter Implementations when Solving Ill-Conditioned Problem 75

17.E.A. Zhirkov, E.R. Muratov
Application of Visual-Inertial Odometer for Autonomous Navigation of Unmanned Aerial Vehicle 79

18.M.A. Klyushin
On Motion Control in the Problem of Spacecraft Inspection 82

19.P.S. Yakovleva, A.A. Kumarin
Estimating Torques for a Cubesat 6U Satellite with Deployable Solar Panels 84

20.M.D. Repin, V.S. Vyazmin
Adaptive Algorithms for Local Gravity Modeling from Airborne Gravimetry Data Using Spherical Point Masses and L_2-Regularization 88

PLENARY PAPERS

21.Luo Mao
Research on the Current Status and Future Development Trends of Intelligent Navigation Technology 92

22.O.A. Stepanov, Yu.A. Litvinenko
Federated Filtering Methods for Non-Recursive Navigation Data Processing 100

POSTER PAPERS

23.O.N. Maltseva
Gait Assessment of a Patient with a Modular Foot Prosthesis Using Marker-Based Video Analysis and Electromyography 105

24.T.A. Brovko, A.P. Malyshev, V.V. Pudlovsky
Comprehensive Assessment of the GLONASS System Efficiency Using a Simulation Model 108

25.V.I. Baburov, N.V. Vasilyeva, N.V. Ivantsevich
Comparison of Criteria for Choosing the Optimal Constellation of Navigation Satellites for Location Determinations in Conditions of Abnormal Use of the Navigation Field of the GNSS 111

26.Guoliang Yang, Yangkang Zhang, Zhaozi Zu
Research on Integrated Navigation Algorithm Based on Radio Navigation Station 114

27.E.S. Zemlyanyy, M.V. Tektov, K.A. Kravtsov
Calculation of Secondary Navigation Parameters Based on the Data of the Short-Range Radionavigation System 119

28.S.S. Golubev, A.B. Nemov, K.Yu. Cheredeev
On the Implementation of the Direction Finding Function in the On-Board DAAR of the UGV 121

29.Zhuo Liu, Gong Min Yan, Guang Bao Shan, Wei Huang, Yu Luo, Xiao-Kang Yang
Research on Time Slicing Square Root Kalman Filter under the Limitation of Hardware Resources 127

30.G.V. Smykov, K.K. Veremeenko
Research of Properties of Characteristics of the Correlation-Extreme Navigation System, Based on Radio Altimeter and Digital Terrain Map 132

31.O.I. Sauta, E.P. Vinogradova, A.V. Zhmurin, I.V. Savelieva
Autonomous Navigation Method When Flying over Mountainous and Flat Terrain 135

32.A.V. Chernodarov, N.P. Starostin
Adaptive Compensation of Error Estimates in Updatable Strapdown Inertial Navigation Systems between Observations 138

33.R.R. Bikmaev
Search Algorithm for Localization of a Ground Vehicle Using Geometric Primitives for Analyzing Trajectory Fields Obtained by Satellite Navigation 142

34.P.K. Kuznetsov, E.A. Schelokov
Technique and System of Contactless Optoelectronic Dual-Beam Time-of-Flight Determination of the Velocity Vector of Micrometeoroids 146

35.M.Yu. Belyaev, P.A. Borovikhin, A.M. Esakov
Objectives and Certain Results of Pointing ISS Scientific Instruments towards their Targets in Earth Science Experiments 150

36.E.V. Barinova, I.A. Timbai, E.D. Mironov, N.A. Elisov, I.A. Lomaka
Determination of Stable Equilibrium Positions of the SamSat-Ionosphere Nanosatellite under the Influence of Gravitational and Aerodynamic Moments 154

37.D.P. Avariaskin, I.V. Belokonov, M.S. Shcherbakov
Ballistic Design of Spacecraft Inspection Missions in Circular Orbits in a Normal Gravitational Field 158

SESSION II – INERTIAL SYSTEMS AND SENSORS

PLENARY PAPER

38.Yu.Yu. Broslavets, E.A. Polukeev, Yu.N. Travina, A.A. Fomichev, V.G. Semenov
Solid-State Laser Gyroscope Based on YAG:Cr4+: Features of Mode-Locking with KTP Phase Modulator Frequency Biasing 161

POSTER PAPERS

39.D.M. Kalikhman, E.A. Deputatova, N.V. Tarakanov, D.S. Gnusarev, L.Ya. Kalikhman
Vibration-Resistant Quartz Pendulum Accelerometer with Digital Feedback 171

40.V.M. Nikiforov, A.A. Gusev, K.A. Andreev, A.D. Sheiko, D.M. Kalikhman, E.A. Deputatova, N.V. Tarakanov
Prediction of Experimental Data of Pendulum Compensation Accelerometer by Moving Average Method by Means of Neural Network 175

41.S.A. Volobuev, E.V. Dranitsyna
Test Results of Compensation Pendulum Accelerometers 178

42.V.V. Sumarokov, E.A. Makhaev, E.A. Yakushina
Providing the Stability of the Pendulum Compensation Accelerometer and Development of its Adjustment Methodology 181

43.E.A. Yakushina, S.G. Romanenko
The Influence of Design Features of the Auto-Compensation Device on the Errors of the Precise Small Tilt Angle Meter 184

44.N.L. Yavorovskaya, T.N. Siraya
Determination of Characteristics of Motion Simulators Used in Navigation Equipment Testing 187

45.D.M. Kalikhman, E.A. Deputatova, N.V. Tarakanov, Yu.V. Chebotarevskii, V.M. Nikiforov
Three-Axis Simulation and Calibration Test Bench with Inertial Sensing Elements in Three Rotational Axes 190

46.N.A. Kiz
QD-Type Sun Sensor Calibration Methods for Small Spacecraft Orientation 194

47.L.S. Chelyshev, S.P. Timoshenkov
Development of Methodology and Study of Operational and Functional Characteristics of Microelectromechanical Sensors 199

48.A.N. Shevchenko, V.V. Chalkov
Studying the Modulation Factor Stability in the Magnetometric System of the Quantum Rotation Sensor 202

49.Ya.S. Savel'eva, L.V. Vodicheva, Yu.V. Parysheva
Calibrating Accelerometers During Operation of a Control System for Space Launch Vehicles 205

PLENARY PAPER

50.Pei Zhang, Yichen Wu, Zerun Zang, Bo Xie, Wei Hong, Yihua Wang, Honggang Chen, Yu Liu, Tianlin Xu
Research on the Technology for Improving Vibration Performance of High Precision Fiber Optic Gyroscope 209

POSTER PAPERS

51.N.V. Moiseev, A.A. Untilov S.V. Pavlova, R.G. Lyukshonkov
Method for FOG Dead Band Detection on a Stationary Base 218

52.Zerun Zang, Pei Zhang, Shitao Ji, Fengyuan Ma, Bin Ren, Wei Hong, Yichen Wu, Xudong Hu, Yu Liu
Research on the Optimization of Double-Layer Magnetic Shielding for Fiber Optic Gyroscope 220

53.Yichen Wu, Pei Zhang, Bo Xie, Yu Liu, Wei Hong, Bin Ren, Hui Cao, Zerun Zang, Yu Bai
Research on the Technology for Improving Long Term Performance of High Precision Fiber Optic Gyroscope by Hermetic Packing 223

54.Xiao Jie Ma, Jizhou Lai, Ji Xun Huan, Wen Long Feng, Xiang Tao Meng, Wei Hou
Analysis of Influence Factors on the Navigation Accuracy of Fiber Optic Gyro Inertial Measurement Unit in Vibration Environments 227

55.Fan Liu, Fanyu Song, Yu Ding, Bingze Lin, Yuelan Li, Zihan Wang
Study of the Characteristics of a Miniature Three-Axis Integrated Fiber-Optic Gyroscope Based on an Application-Specific Integrated Circuit (ASIC) 232

56.M.A. Basarab, B.S. Lunin, I.S. Stroganov, E.A. Chumankin
Neural Network Models for Thermal Drift Compensation of the Wave Solid-State Gyroscope 240

57.Yu Xuan Xie, Wei Wu, Jin Yu Li, Pan Yao, Yu Feng Tao, Hui Luo
Effect of Assembly Errors on the Static Capacitance of the Hemispherical Resonator Gyroscope 244

58.Wei Cheng, Shun Qing Ren, Bo Qi Xi, Zhen Tian, You Huan Ning, Chang Hong Wang
Dynamic Modeling and Mass Balancing of Hemispherical Resonators with Variable Thickness 249

59.S.E. Perelyaev, A.V. Alekhin
The Implementation of the Hilbert Transform of the Primary Signal of a Wave Solid-State Gyro 253

60.S.E. Perelyaev, S.T. Lesnoy
Features of Digital WSG Algorithms Implementation on Modern System-on-Chip Platforms 256

61.V.M. Kotlov, S.E. Perelyaev
Excitation Mode with Continuous Restoration of the Velocities of Movement of Elastic Deformations of a Single-Axis Solid-State Gyroscope 259

62.A.A. Maslov, D.A. Maslov, I.V. Merkuryev, I.G. Ninalalov
Increasing the Accuracy of a Micromechanical Gyroscope with a Ring Resonator in a Multimode Operation Mode 262

PLENARY PAPER

63.D.B. Pazychev, A.S. Fadeev, V.N. Gerdi
Ultra-Compact Magnetic Heading Sensor for UAV 266

POSTER PAPERS

64.D.G. Gryazin, T.V. Paderina
Magnetic Compass with a Correction System Built into the Autopilot Circuit 273

65.Xiaoge Ning, Jixun Huang, Jianxun Li
An Adaptive Initial Alignment Filtering Algorithm for Strapdown Inertial Navigation under Large Misalignment Angles 277

66.L.V. Vodicheva, Yu.V. Parysheva, Ya.S. Savel'eva, D.I. Kabanova, N.A. Shmakov
Calibration of a Strapdown Inertial Measurement Unit with Two Angular Rate Sensors 281

67.A.V. Chernodarov, P.S. Gorshkov, I.V. Mitrofanov, A.P. Patrikeev
Adaptive Robust Filtering and Calibration of Inertial MEMS Sensor Signals 285

68.S.V. Topilskaya, A.O. Kolesnikov, Yu.V. Bakin
Accuracy Improvement of SINS on Laser Gyroscopes under Vibration 289

69.W.F. Luo, J. Xu, W.M. Geng, J.M. Tang, T.J. Wang
Construction and Verification Analysis of Digital Model for Dynamic Characteristics of Inertial Measurement Unit System 292

70.D.V. Akulin, M.V. Mentyukov, K.O. Ivanov
Infrared Sensor-Based Attitude Determination System for CubeSat Satellites 300

71.E.S. Minchenkova, D.D. Nikeev, P.V. Gaenko, K.A. Karimulin
Autonomous Navigation Module for Aircraft 304

72.V.V. Matveev, A.V Kalikanov, A.N. Khomyachkova
TRIAD Method for Solving Orientation Problem 307

73.N.A. Efimov, A.A. Golovan, V.L. Hirnaya, V.V. Sidelnikova, I.A. Papusha
The Adaptation and Application of SINS Fine Alignment in the Case of a Vessel's Rolling 311

SESSION III – RELEVANT ISSUES OF THEORY

PLENARY PAPER

74.S.E. Perelyaev, V.F. Zhuravlev
On the Control of Spatial Elastic Standing Waves in a Complete Spherical Shell 314

POSTER PAPERS

75.Yu.N. Chelnokov
New Quaternion Algorithms of SINS in Euler Parameters for Determining the Inertial Orientation of Moving Objects: Using the Structures of Runge-Kutta Methods 323

76.A.V. Molodenkov, Ya.G. Sapunkov, T.V. Molodenkova
Time-Quasi-Optimal Spacecraft Angular Acceleration Constructed according to the Poinsot Kinematic Interpretation 327

77.I.A. Pankratov, Yu.N. Chelnokov
The Problem of Spacecraft Optimal Flight to the Orbit of a Moon Satellite Using Limited Thrust and New Regular Quaternion Equations of the Perturbed Spatial Restricted Three-Body Problem in KS-variables 332

78.Ke Fang, Ti Jing Cai
The Kinematic Models of SINS and Its Errors on the Se2(3) Group: Analysis and Comparison 336

79.A.D. Sheyko, V.M. Nikiforov, A.A. Gusev, K.A. Andreev, M.L. Simakov, A.V. Shevchenko
Quadcopter Motion Control System under Disturbances Based on Fuzzy Logic 340

SESSION I – INTEGRATED SYSTEMS

Bayesian Estimation Algorithms for Solution of SLAM Problem

O.A. Stepanov
Concern CSRI Elektropribor, JSC,
ITMO University
St. Petersburg, Russia
soalax@mail.ru

A.V. Motorin
Concern CSRI Elektropribor, JSC,
ITMO University
St. Petersburg, Russia
motorin.a@mail.ru

V.P. Zolotarevich
Concern CSRI Elektropribor, JSC,
ITMO University
St. Petersburg, Russia
zolotarevich@yandex.ru

***Abstract* — The problem of simultaneous localization and mapping and algorithms of its solution are considered within the framework of the Bayesian approach. Their relationship with solution algorithms based on graph optimization methods is discussed. It is shown that such algorithms are actually based on the combined use of the Bayesian approach and graph optimization methods.**

Keywords— nonlinear estimation, Bayesian approach, factor-graph, optimization, simultaneous localization and mapping problem.

I. INTRODUCTION

The paper [1] discusses the relationship and differences between Bayesian estimation algorithms [2–7] and algorithms based on the use of graph optimization methods, also called factor-graph optimization algorithms [8–12]. This paper is a continuation of these studies, carried out using the example of solving the Simultaneous Localization and Mapping problem (SLAM).

The paper is structured as follows. Section 2 presents the traditional Bayesian statement and general solution of the estimation problem considered and discusses its relationship with the solution of the system of linear equations (SLE or linear system). Section 3 gives the mathematical problem statement of the SLAM within the Bayesian approach, discusses the differences from the traditional problem statement and notes its computational complexity conditioned by the high dimension of the estimated state vector. This, in turn, necessitates the use of new algorithms for obtaining an estimate. The stages of their development are presented, including those based on the transition to the information form of representing the a posteriori density and the feasibility of applying the SLE solving methods that are closely related to graph optimization methods. Conclusions are presented in Section 4.

II. TRADITIONAL ESTIMATION PROBLEM STATEMENT AND ITS SOLUTION WITHIN THE BAYESIAN APPROACH

The traditional problem of joint processing of measurements from different sensors can be formulated as a problem of nonlinear estimation of an *n*-dimensional random vector described by a shaping filter

$$x_k = f_k\left(x_{k-1}, u_k\right) + \Gamma_k w_k \qquad (1)$$

by *m*-dimensional measurements of the following form:

$$y_k = h_k(x_k) + v_k. \qquad (2)$$

In these relations x_k is an *n*-dimensional random Gaussian vector of estimated parameters with a given a priori probability density function (PDF) $p(x_0) = N(x_0; \overline{x}_0, P_0)$; u_k is a vector of known input control signals and/or navigation sensor signals; w_k is a centered discrete Gaussian system white noise with known covariance matrix Q_k; y_k is a *p*-dimensional vector of measurements; v_k is a centered discrete Gaussian white noise of measurements, with known covariance matrix R_k, Γ_k is a known matrix; $f_k(\bullet)$, $h_k(\bullet)$ are known nonlinear vector functions describing the dynamics of the state vector and the measurement model; k is a discrete time index.

The concept of the Bayesian estimation problem solution consists in obtaining an optimal mean-square estimate of the state vector $\hat{x}_k^{opt}(Y)$ based on measurements $Y = (y_1^T, ..., y_k^T)^T$ and the corresponding conditional covariance matrix of estimation errors $P_k^{opt}(Y)$.

This estimate $\hat{x}_k^{opt}(Y)$ is defined as the mathematical expectation of the a posteriori PDF $p\left(x_k / Y\right)$ [2,3,6]:

$$\hat{x}_k^{opt}\left(Y\right) = \int x_k p\left(x_k / Y\right) dx_k \,. \qquad (3)$$

Calculating the optimal estimate (3) in nonlinear problems (1), (2) is often non-trivial, but many recursive estimation methods have been developed to solve such problems, including the extended Kalman filter (EKF) and other filters based on the Gaussian approximation of the PDF [3–5] (cubature, unscented Kalman filter), solutions based on the numerical representation of the PDF $p\left(x_k / Y\right)$ and calculation of the integral (3) using sequential Monte Carlo methods (particle filter), as well as combined methods such as Rao-Blackwellization [3–5, 13,14].

In this paper we will focus on the simplest linearization method, which involves the expansion of nonlinear functions from (1) in a series in the form:

The study was supported by the grant of the Russian Science Foundation No. 23-19-00626, https://rscf.ru/project/23-19-00626/

$$f_k(x_{k-1},u_k) \approx f_k(x_k^{\pi_1},u_k) + \Phi_k(x_{k-1} - x_k^{\pi_1}),$$
$$h_k(x_k) \approx h_k(x_k^{\pi_2}) + H_k(x_k - x_k^{\pi_2}), \quad (4)$$

where Φ_k, H_k, are the Jacobian matrices of functions $f_k(x_{k-1},u_k)$, $h_k(x_k)$, calculated at the linearization points $x_k^{\pi_1}, x_k^{\pi_2}$. As a result, the problem (1), (2) is reduced to a linear Gaussian estimation problem.

In [1] it is shown that when using a set of non-recursively processed measurements, prerequisites are created for increasing the accuracy of the algorithms for solving the problem based on linearization and ensuring their consistency. Taking this into account, we will further focus on using a set of measurements in the algorithm. In this case, the estimate (3) will agree with the estimate that minimizes a criterion of the following form:

$$J(X) = \left(\begin{bmatrix} H \\ \Phi \end{bmatrix} X - \begin{bmatrix} Y \\ \tilde{u} \end{bmatrix}\right)^T \begin{bmatrix} R^{-1} & 0 \\ 0 & G^{-1} \end{bmatrix} \left(\begin{bmatrix} H \\ \Phi \end{bmatrix} X - \begin{bmatrix} Y \\ \tilde{u} \end{bmatrix}\right) \quad (5)$$

where $X = \begin{bmatrix} x_0^T & \dots & x_k^T \end{bmatrix}^T$ is a composite estimated vector that includes the state vectors x_k at all moments of time, $G^{-1} = blkdiag\left(P_0^{-1} \quad \left(\Gamma_1 Q_1 \Gamma_1^T\right)^{-1} \quad \dots \quad \left(\Gamma_k Q_k \Gamma_k^T\right)^{-1}\right)$, $blkdiag(\bullet)$ is a block-diagonal matrix, $\tilde{u} = \begin{bmatrix} \bar{x}_0^T & -\tilde{u}_1^T & \dots & -\tilde{u}_k^T \end{bmatrix}^T$, $\tilde{u}_k = f_k(x_k^{\pi_1},u_k) - \Phi_k x_k^{\pi_1}$, $Y = \begin{bmatrix} \tilde{y}_1^T & \dots & \tilde{y}_k^T \end{bmatrix}^T$, $\tilde{y}_k = y_k - h_k(x_k^{\pi_2}) + H_k x_k^{\pi_2}$, $H = \begin{bmatrix} 0 & blkdiag(-H_1 & \dots & -H_k) \end{bmatrix}$, $R = blkdiag(-R_1 \quad \dots \quad -R_k)$,

$$\Phi = \begin{bmatrix} E & 0 & \dots & 0 \\ -\Phi_1 & E & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & -\Phi_k & E \end{bmatrix}.$$

Following [1], we will explain the relation to the problem of solving SLE. In [1] it is noted that minimization of the criterion (5) is equivalent to solving an overdetermined SLE of the form:

$$AX = b, \quad (6)$$

in which

$$A = S\begin{bmatrix} H \\ \Phi \end{bmatrix}, \; b = S\begin{bmatrix} Y \\ \tilde{u} \end{bmatrix}, \quad (7)$$

where the matrix S satisfies the equation.

$$\begin{bmatrix} R^{-1} & 0 \\ 0 & G^{-1} \end{bmatrix} = S^T S. \quad (8)$$

Solving the SLE (6) by the least squares method (LSM) is reduced to SLE:

$$\Omega X = \xi, \quad (9)$$

where $\Omega = A^{\mathrm{T}}A$, $\xi = A^{\mathrm{T}}b$. In this case, Ω is the information matrix of the posterior PDF $p(X/Y)$, and ξ is its information vector [2,3]. Direct obtaining of the inverse matrix Ω^{-1} is computationally difficult due to its large dimension. The solution of the SLE (9) can be found using the Cholesky decomposition of the matrix Ω (such a decomposition exists because the matrix Ω is symmetric and positive definite):

$$\Omega = LL^{\mathrm{T}}, \quad (10)$$

where L is a lower triangular matrix (Cholesky multiplier). The solution of SLE is then reduced to a two-stage procedure. At the first stage (forward) the solution to the problem is found:

$$L\tilde{X} = \xi, \quad (11)$$

for $\tilde{X} = L^T X$, and on the second stage (reverse) the estimate $\hat{X}$ is calculated from the given ratio:

$$L^T X = \tilde{X}. \quad (12)$$

The problems (11), (12) are solved simply by forward and back substitution, since L is a triangular matrix, in the Russian literature such an algorithm is also called the sweep method for block matrices [15]. Moreover, in the considered linear problem L will be block-bidiagonal [16], which allows obtaining a recursive estimation algorithm. Recursive forward substitution for solving (11) will match with the Kalman filter (KF), and the back substitution for (12) will match with the Rauch-Tung-Stribel smoothing algorithm [16]. In a nonlinear problem, when choosing the propagation as linearization points $x_k^{\pi_1}, x_k^{\pi_2}$, the expression (11) corresponds to the solution of the problem using the extended Kalman filter (EKF). As a result, in general, the solution to the problem of finding an estimate $\hat{X}$ will coincide with the solution to the smoothing problem for all components of the vector X . A comparison of recursive and non-recursive algorithms for solving the problem (1), (2) is carried out in [1].

Here we also note that the optimal estimate (3) in the linear (or linearized) Gaussian case matches with the estimate corresponding to the maximum of the a posteriori PDF.

$$\hat{x}_k(Y) = \arg\max_{x_k} p(x_k / Y). \quad (13)$$

III. SLAM Problem Statement and its Solution in the Framework of the Bayesian Approach and The Relationship with Factor Graph Optimization Algorithms

SLAM problem can be generally represented as a Bayesian problem of estimating a joint state vector $\mathbf{x}_k = \begin{bmatrix} x_k^T & M^T \end{bmatrix}^T$ described by a system of equations:

$$x_k = f_k(x_{k-1},u_k) + \Gamma_k w_k,$$
$$M_k = M_{k-1} = M, \quad (14)$$

for p-dimensional measurements of the form:

$$y_k = h_k(x_k, M) + v_k, \quad (15)$$

where $M = \begin{bmatrix} m_1^T & \dots & m_j^T \end{bmatrix}^T$ is a constant vector describing the map, and x_k is a state vector of the robot.

The first feature of the SLAM problem is the presence of a map vector M consisting of a large number of subvectors m_j, independent of each other. These subvectors may differ in their physical nature and dimensionality depending on the description of the environment map, the sensors used to observe it, and the front-end pre-processing algorithms. They may represent either the coordinates of individual map landmarks [17,18] or the occupancy of the map grid cells if it is represented as an occupancy grid map [19], or parameters that determine the rendering (drawing) of the image when using the NErF [20] or 3D Gaussian splatting technique [21]. However, common to all SLAM problems is the fact that the vector M consists of a large number of independent constant subvectors m_j, the number j of which depends on the size of the map being studied. In the process of solving the problem, i.e. studying the map, the dimension of the vector M also grows. New map objects are added, and the number of subvectors m_j grows accordingly, describing them. In this paper, attention is not focused on this, since the introduction of new elements into the problem in the process of its solution does not pose a problem.

The second feature of the SLAM problem is that the measurements (15) at individual moments of time k depend only on a part of the components of the vector M. Since usually only a small section of the map is subject to measuring at each moment of time, the measurement function $h_k(x_k, M)$ includes a relatively small number of subvectors m_j. In addition to this, the problem of data association arises, i.e. determining which map parameters are currently measured. To simplify the following conclusions, the solution to this problem is also left outside the scope of this work.

The first feature makes the application of the above-considered recursive algorithms, such as EKF-SLAM, computationally inefficient for solving the problem (14), (15) . The main problem of EKF-SLAM [17] is the need to calculate and store the full covariance matrix for the vector $\mathbf{x}_k$, which causes a quadratic dependence of the computational complexity of the algorithm on the map dimension, and makes such an algorithm unacceptable for large spaces. The application of other recursive methods for solving nonlinear problems, such as the use of Rao-Blackwellization over x_k, called Fast-SLAM [18], preserves the quadratic dependence of the amount of calculations on the number of map objects, due to their constant presence in the state vector (14). The main reason for abandoning EKF-SLAM and Fast-SLAM at present is their computational complexity.

At the same time, these same features of the SLAM problem allowed us to obtain computationally efficient algorithms for its solution by representing it in a non-recursive form, similar to (5) and the application of SLE solution methods (6), (9). Representing the criterion (5) for the problem (14), (15) in the form:

$$J(\mathbf{X}) = \left(\begin{bmatrix} H \\ \mathbf{\Phi} \end{bmatrix} \begin{bmatrix} X \\ M \end{bmatrix} - \begin{bmatrix} Y \\ \tilde{\mathbf{u}} \end{bmatrix} \right)^T \begin{bmatrix} R^{-1} & 0 \\ 0 & \mathbf{G}^{-1} \end{bmatrix} \left(\begin{bmatrix} H \\ \mathbf{\Phi} \end{bmatrix} \begin{bmatrix} X \\ M \end{bmatrix} - \begin{bmatrix} Y \\ \tilde{\mathbf{u}} \end{bmatrix} \right)^T \quad (16)$$

where $\mathbf{X} = [\, X^T M^T \,]^T$,

$$\mathbf{\Phi} = \begin{bmatrix} \Phi & 0 \\ 0 & E \end{bmatrix}, \ \mathbf{G}^{-1} = \begin{bmatrix} G^{-1} & 0 \\ 0 & P_M^{-1} \end{bmatrix}, \ \tilde{\mathbf{u}} = \begin{bmatrix} \tilde{u} \\ 0 \end{bmatrix},$$

Φ, G , $\tilde{u}$ are calculated similarly to (5) for the model (14), and P_M is the a priori matrix of covariances of the vector M. In this case, the information matrix corresponding to the problem (16) and information vector can be written as:

$$\Omega = \Phi^T \mathbf{G}^{-1} \Phi + H^T R^{-1} H \quad (17)$$

$$\xi = \Phi^T \mathbf{G}^{-1} \tilde{\mathbf{u}} - H^T R^{-1} Y \,. \quad (18)$$

Thus, the optimal Bayesian solution also reduces to solving a SLE of the form (9), complicated by the fact that the dimension of the matrix (17) and vector (18) increased by j, and the dimension of the recursive problem at each step also increased by j. To simplify the solution, the following procedure is proposed in [8], taking into account the first feature of the SLAM problem – the fact that the vector M consists of a set of independent subvectors. Since the matrix (17) is an information matrix of the joint PDF of vectors X and M $p(\mathbf{X}/Y) = p(X, M/Y)$ we can find the parameters of the particular PDF $p(X/Y) = \int p(\mathbf{X}/Y) dM$. Let's represent the information matrix (17) in block form:

$$\Omega = \begin{bmatrix} \Omega_{XX} & \Omega_{XM} \\ \Omega_{XM}^T & \Omega_{MM} \end{bmatrix} \quad (19)$$

and, similarly, we represent the information vector (18):

$$\xi = \begin{bmatrix} \xi_X \\ \xi_M \end{bmatrix}. \quad (20)$$

The information matrix $\tilde{\Omega}$ and vector $\tilde{\xi}$ of the particular PDF $p(X/Y)$ can be obtained according to the relations:

$$\tilde{\Omega} = \Omega_{XX} - \Omega_{XM} \Omega_{MM}^{-1} \Omega_{XM}^T \,, \quad (21)$$

$$\tilde{\xi} = \xi_X - \Omega_{XM} \Omega_{MM}^{-1} \xi_M \,. \quad (22)$$

Considering that Ω_{MM}^{-1} is block-diagonal, due to the independence of subvectors m_j, the expressions (21), (22) can also be represented as:

$$\tilde{\Omega} = \Omega_{XX} - \sum_j \Omega_{Xm_j} \Omega_{m_j m_j}^{-1} \Omega_{Xm_j}^T \,, \quad (23)$$

$$\tilde{\xi} = \xi_X - \sum_j \Omega_{Xm_j} \Omega_{m_j m_j}^{-1} \xi_{m_j} \,. \quad (24)$$

where $\Omega_{m_j m_j}^{-1}$ are the diagonal blocks of the matrix Ω_{MM}^{-1} corresponding to the subvectors m_j, Ω_{Xm_j} are the corresponding rows of the matrix Ω_{XM} , ξ_{m_j} are the subvectors of the vector ξ_M . The posterior matrix of covariances of the vector X – P_X –and its estimate $\hat{X}$ are calculated according to

$$P_X = \tilde{\Omega}^{-1}, \ \hat{X} = P_X \tilde{\xi} \,. \quad (25)$$

Note that in order to solve the equations (25), we may also apply the methods for solving SLE (6), (9). Estimates $\hat{m}_j$ and covariance matrices P_{m_j} of state vectors for individual map objects m_j are calculated according to:

$$\hat{m}_j = P_{m_j}\left(\xi_{m_j} + \Omega_{xm_j}^T \hat{X}\right), P_{m_j} = \Omega_{m_j m_j}^{-1} \quad (26)$$

The computational complexity of the algorithm (23)–(26) is linearly dependent on the number of map objects *j*, since in (23), (24) only summation over this parameter is performed, and the estimates (26) are calculated individually. Reduction of the computational complexity in the algorithm (23)–(26) relative to the EKF is achieved due to no need of calculating and storing the correlation coefficients of the components of the vector *M* and the correlation coefficients of the vectors *M* and *X*.

At the same time, the algorithm (23)–(26) assumes non-recursive solution to the problem and a quadratic increase in computational complexity with increasing time *k*, due to the need to calculate the estimate (25). In practical problems, this is not always convenient, and that is why an incremental smoothing algorithm was developed [9, 11], which allows for relatively fast operation in real time and is widely used at present. The algorithm uses linear algebra methods for incremental calculation of the smoothed estimate, and is based on the QR-decomposition of the matrix *A* from (6), which consists in its representation in the form:

$$A = \mathbf{Q}\begin{bmatrix}\mathbf{R}\\0\end{bmatrix} \quad (27)$$

where **Q** is an orthogonal matrix, and $\mathbf{R} = \mathrm{L}^T$ is the upper triangle of the Cholesky decomposition of the matrix Ω. This allows us to reduce the solution of the system (6) to the solution of an upper-triangular system of lower dimension:

$$\mathbf{R}\mathbf{X} = d\ ,\ \begin{bmatrix}d\\e\end{bmatrix} = \mathbf{Q}^T b\,, \quad (28)$$

easily solved by back substitution. The main idea of this representation is that when new measurements appear at step *k*+1 they can be added to (28) in the form:

$$\begin{bmatrix}\mathbf{R}\\S_{k+1}H_{k+1}\end{bmatrix}\mathbf{X} = \begin{bmatrix}d\\S_{k+1}y_{k+1}\end{bmatrix}, \quad (29)$$

and after this the problem (29) can be reduced to the form (28), while the search for a new QR decomposition is simpler due to the fact that **R** is already upper-triangular. The expansion of the state vector when new state vectors of the map appear is carried out in the form:

$$\begin{bmatrix}\mathbf{R} & 0\\0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{X}\\m_{j+1}\end{bmatrix} = \begin{bmatrix}d\\0\end{bmatrix} \quad (30)$$

and at the new step of the robot in the form:

$$\begin{bmatrix}\mathbf{R} & 0\\0\ldots -S_k\Phi_k & S_k\end{bmatrix}\begin{bmatrix}\mathbf{X}_{k-1}\\x_k\\x_{k+1}\end{bmatrix} = S_k\tilde{u}_{k+1}\,. \quad (31)$$

Thus, only a few rows will be added to the matrix *A* at each moment in time, which will allow us to use the previously calculated QR decomposition of the matrix *A*.

The key aspect of the algorithm (28)–(31) from the viewpoint of its computational complexity is the sparsity of the matrix **R**, since this matrix, along with the vector *b*, is the only thing that is stored in the algorithm and requires their recalculation at each step. The possibility of obtaining a sparse matrix **R** is physically determined by the independence of the subvectors m_j, and the fact that only a small part of them is subject to measurement at each moment in time, i.e. the features of the SLAM problem indicated earlier. However, the search of QR decomposition that ensures the maximum sparsity of **R** is a non-trivial problem solved using graph models [9, 11, 22].

The relationship between SLAM algorithms and methods for solving sparse SLE and probabilistic graph models along with the possibility of their application to reduce the computational complexity of the algorithm are also discussed in the paper.

IV. CONCLUSION.

Factor-graph optimization methods are aimed at solving the Bayesian smoothing problem (in the Gaussian case, the equivalent problem of maximizing the posterior density), as a result of which the opposition of Bayesian algorithms, in particular Kalman filters, to these methods is often incorrect.

The SLAM problem has a specific feature from the viewpoint of estimation theory, which consists in the presence in the state vector, along with the Markovian sequence describing the robot motion, of a constant vector of large dimension representing a map, which consists of a large number of independent subvectors. Such specificity makes the use of EKF and a number of other recursive algorithms computationally ineffective for its solution. At the same time, it can be taken into account using probabilistic graphical models, and the calculation of the optimal Bayesian estimate can be obtained using computational methods with the corresponding sparse matrices. This allows us to talk about the combined use of the Bayesian approach and factor-graph optimization methods. In traditional problems, such algorithms usually do not have a significant advantage over recursive methods for solving the smoothing problem.

V. REFERENCES

[1] Stepanov, O.A., Zolotarevich, V.P., Litvinenko, Yu.A., Isaev, A.M., and A.V. Motorin, Recursive and nonrecursive algorithms applied to navigation data processing: Differences and interrelation with factor graph optimization algorithms, Proceedings of the 31st St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, 2024, pp. 336-344.

[2] Meditch, J.S., Stochastic optimal linear estimation and control, New York, McGraw-Hill, 1969. 384 p. (Russ. ed.: Medich. Dzh. Statisticheski optimal'nye otsenki i upravlenie. Translated from English, Moscow: Energia Publ., 1973, 440 p.).

[3] Stepanov, O.A., Primenenie teorii nelineinoi fil'tratsii v zadachakh obrabotki navigatsionnoi informatsii (Application of Nonlinear Filtering Theory for Processing Navigation Information), St. Petersburg: CSRI Elektropribor, 2003, 370 p.

[4] Gibbs, B.P. Advanced Kalman Filtering, Least-Squares and Modeling: A Practical Handbook, John Wiley&Sons, Inc., 2011. https://doi.org/10.1002/9780470890042.

[5] Särkkä, S., Bayesian Filtering and Smoothing, Cambridge University Press, 2013.
https://doi.org/10.1017/CBO9781139344203

[6] Stepanov, O.A., Osnovy teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoi informatsii. Part 1. Vvedenie v teoriyu otsenivaniya (Fundamentals of the Estimation Theory with Applications to the Problems of Navigation Information Processing. Part 1. Introduction to the Estimation Theory), St. Petersburg: CSRI Elektropribor, 2017, 509 p.

[7] Stepanov, O.A., Osnovy teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoi informatsii. Part 2. Vvedenie v

teoriyu fil'tratsii (Fundamentals of the Estimation Theory with Applications to the Problems of Navigation Information Processing. Part 2. Introduction to the Filtering Theory), St. Petersburg: CSRI Elektropribor, 2017, 428 p.

[8] Thrun, S. and Montemerlo, M., The GraphSLAM algorithm with applications to large-scale mapping of urban structures, The International Journal of Robotics Research, vol. 25, no. 5–6, 2006, pp. 403-429. https://doi.org/10.1177/0278364906065387

[9] Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J., and Dellaert, F., iSAM2: Incremental smoothing and mapping using the Bayes tree, International Journal of Robotics Research, vol. 31, no. 2, 2011, pp. 216-235

https://doi.org/10.1177/0278364911430419

[10] Wen, W., Pfeifer, T., Bai, X., and Hsu, Li-Ta, Factor graph optimization for GNSS/INS integration: A comparison with the extended Kalman filter, NAVIGATION, 2021, no. 68, pp. 315–331. https://doi.org/10.1002/navi.421

[11] Dellaert, F. and GTSAM Contributors, Georgia Tech Smoothing and Mapping library, https://github.com/borglab/gtsam, 2022.

[12] Quan, S.H, Chen, S.H, Zhou, Y.L, Zhao, S., Hu, H.Z, and Zhu, Q., A robust IMU/GNSS/ODO integrated navigation system based on factor graph, Proceedings of the 31st Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2024.

[13] Doucet, A., Freitas, N., and Gordon, N., Sequential Monte Carlo Methods in Practice, Springer, 2001, 590 p.

[14] Stepanov, O.A., Toropov, A.B., Application of stepwise Monte Carlo methods using analytical integration procedures in processing navigation information, Proceedings of the 12th All-Russia Conference on Control Problems VSPU-2014, Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, 2014, pp. 3324-3337.

[15] George, A. and Liu, G., Numerical Solution of Large Sparse Systems of Equations [Russian translation], Moscow: Mir, (1984), 334 p.

[16] Aravkin, A, Bell, B, Burke, J., and Pillonetto, G., Kalman smoothing and block tridiagonal systems: new connections and numerical stability results, Journal of Machine Learning Research, no.14, 2013. https://doi.org/10.48550/arXiv.1303.5237

[17] Dissanayake, G., Newman, P., Clark, S., Durrant-Whyte, H.F., and Csorba, M., A solution to the simultaneous localization and map building (SLAM) problem, IEEE Transactions on Robotics and Automation, vol. 17, no. 3, 2001, pp. 229-241. https://doi.org/10.1109/70.938381.

[18] Montemerlo, M., Thrun, S., Koller, D., and Wegbreit, B., FastSLAM: A factored solution to the simultaneous localization and mapping problem, Proceedings of AAAI National Conference on Artificial Intelligence, 2002.

[19] Ozisik, O. and Yavuz, S., An occupancy grid based SLAM method, Proceedings of 2008 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2008, pp. 117-119. https://doi.org/10.1109/CIMSA.2008.4595844

[20] Rosinol, A., Leonard, J.J., and Carlone, L., NeRF-SLAM: Real-time dense monocular SLAM with neural radiance fields, Proceedings of 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Detroit, MI, USA, 2023, pp. 3437-3444. https://doi.org/10.1109/IROS55552.2023.10341922

[21] Keetha, N. et al., SplaTAM: Splat, track & map 3D Gaussians for dense RGB-D SLAM, Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2024, pp. 21357-21366.

[22] Davis, T., Gilbert, J., Larimore, S., and Ng, E., A column approximate minimum degree ordering algorithm, ACM Trans. Math. Software 2004, vol. 30, no. 3, pp. 353–376. https://doi.org/10.1145/1024074.1024079

Cooperative Image Enhancement for Vision-Based MASS Navigation: A MADRL Approach

Zhuolin Wang
Navigation College
Dalian Maritime University
Dalian 116026, China
wangzhuol@dlmu.edu.cn

Yingjun Zhang
Navigation College
Dalian Maritime University
Dalian 116026, China
zhangyj@dlmu.edu.cn

Haoze Zhang
Navigation College
Dalian Maritime University
Dalian 116026, China
zhzll@dlmu.edu.cn

***Abstract*—This paper presents a MADRL-based image enhancement paradigm with multiple visual metrics. The objective is to address the vision-based navigation challenges faced by MASS in nighttime. Specifically, the paradigm consists of two major modules, scene feature analyzer and image enhancement decoder. In the scene feature analyzer, a CNN is constructed to extract key image features from four dimensions, which are used to assess image quality. Additionally, low-light images of are classified into different levels. Depending on the classification results, images are input into a RL environment for varying degrees of enhancement. In image enhancement decoder, the iterative enhancement process of degraded images is modeled as Markov Decision Process (MDP). Agents interact with the environment to iteratively optimize joint strategy for image enhancement. For training, low-light images are synthesized by combining nonlinear and linear transformations, using images from public datasets as references. Results demonstrate that the proposed paradigm achieves significant enhancement in both qualitative and quantitative evaluations.**

Keywords—marine navigation, adaptive Image Enhancement, distributed multi-agent system, D2QN with PER

I. INTRODUCTION

Perception accuracy is crucial for ensuring the safe navigation of Maritime Autonomous Surface Ships (MASS). Perception systems is composed of various sensors, each with limited application scenarios. Among them, visible light cameras have emerged as one of the most widely used devices on MASS due to cost-effectiveness and broad applicability. Such characteristics enable it to provide unrestricted visual information across diverse navigation conditions, which is particularly critical for ensuring the safe navigation of MASS.

Maritime navigation systems frequently encounter compromised visual quality during twilight conditions, where suboptimal illumination and environmental factors adversely affect imaging performance. These challenges manifests as color distortion, underexposure, and reduced contrast. Such degradation not only impacts the clarity and detail of images but also poses severe challenges to downstream visual tasks, particularly in target detection. Therefore, the development of effective image enhancement algorithms holds significant research importance and practical value for improving vision-based navigation systems of MASS. The main goal is to improve overall brightness and contrast, restore details and enhance visual quality, while emphasizing content that is critical for decision-making.

To address the challenges, researchers are exploring machine learning methods to adapt to the complex navigation environments of MASS for high-quality image enhancement. Notably, reinforcement learning (RL) offers a novel approach by learning optimal strategies through interactions. Through iterative attempts of enhancement strategies and feedback-driven parameter adjustments, agent groups dynamically optimize enhancement processes to address diverse demands across varying scenarios of MASS. This approach bridges adaptive decision-making with perceptual refinement, ensuring robust environmental awareness under challenging visual conditions.

Accordingly, an image enhancement paradigm is proposed based on multi-agent distributed reinforcement learning (MADRL). The architecture processes low-light images as input, employing a collaborative multi-agent optimization method to overcome the generalization limitations inherent in conventional deep learning approaches. Within the framework, agents dynamically learn optimal enhancement strategies through interaction with the environment, simultaneously achieving multiple visual quality improvements including luminance optimization, contrast enhancement, noise suppression, and detail restoration. The proposed multi-agent collaborative learning mechanism demonstrates significant improvements in visual perception capabilities for MASS under low-light conditions. This advancement provides robust technical support for enhancing the reliability of vision-based maritime navigation.

The manuscript is organized as follows. Section II describes the current state of research in low-light image enhancement. In section III, the proposed MADRL paradigm is introduced, including scene feature analyzer and image enhancement decoder. In Section IV, proposed method is verified through the execution of a series of simulations and experiments. The conclusions of our study are summarized in Section V.

II. ANALYSIS OF CURRENT STATE OF LOW-LIGHT IMAGE ENHANCEMENT

Current low-light image enhancement approaches can be categorized into two categories, physical model-based approaches and learning-based approaches.

A. Physical model-based approaches

Traditional low-light image enhancement methods primarily rely on physical models and prior knowledge, employing mathematical frameworks to restore degraded images. Such can be classified into two principal categories, (1) pixel distribution mapping techniques, exemplified by histogram equalization (HE) and gamma correction [1], and (2) mathematical optimization-based methods, particularly those grounded in Retinex theory [2]. These methods operate

*Research supported by National Natural Science Foundation of China [Grant No.52231014] and the Team Project of Dalian Maritime University [Grant No.3132023511].

through designed feature extraction rules to enhance or recover low-quality images.

While such approaches demonstrated promising results in early research, they exhibit several limitations that have constrained practical application. The most significant constraints include: (1) heavy reliance on manual feature engineering, resulting in limited generalization capabilities across diverse scenarios; (2) rigid enhancement pipelines that cannot adapt to varying environmental conditions; and (3) inherent issues with detail loss and artifact generation during enhancement process. These limitations have become particularly pronounced in complex real applications, leading to the gradual displacement by more adaptive, data-driven approaches.

B. Learning-based approaches

Deep learning, with its powerful feature extraction capabilities, has achieved significant progress in the field of image enhancement. The majority research employs supervised learning paradigms, where neural networks learn complex mapping functions between input low-quality images and their corresponding high-quality counterparts.

The application in low-light image enhancement was pioneered by LLNet [3], which introduced an encoder-decoder architecture incorporating residual learning principles. This work demonstrated the potential of deep neural networks in addressing illumination-related image degradation. RetinexNet [4] established a novel framework that combines Retinex theory with CNN. The architecture employs two specialized sub-networks, Decom-Net for reflectance-illuminance decomposition while Enhance-Net for component-specific processing, ultimately synthesizing illuminated images through learned feature recombination. EnlightenGAN [5] employs a conditional Generative Adversarial Network (cGAN) to address the degradation of low-light images. It employs a generator based on the U-Net to map low-light images into a visual space with normal illumination. Meanwhile, a discriminator is integrated to guide the generator in learning to produce realistic enhanced images. In maritime image enhancement, researchers have proposed various solutions to address the unique challenges of marine environments. Liu et al. [6] developed AiOENet, which incorporates Transformer modules to enhance the network's capability for detailed feature extraction in complex maritime scenes. Lu Li [7] introduced an attention-based residual module specifically designed to handle low-light image enhancement in port environments, demonstrating improved performance in challenging lighting conditions. Wenbo Zhou [8] proposed MFF-Net, which employs cross-attention modules and residual convolutional layers to simultaneously extract global and local features while integrating color restoration capabilities.

While these supervised learning approaches have demonstrated significant improvements in maritime image enhancement, they share common limitations. These methods suffer from a high dependence on high-quality labeled data. And the performance of these models rely heavily on both the quantity and quality of annotated data. However, acquiring high-quality labeled data is particularly problematic in low-light conditions, where environmental factors and equipment limitations often compromise data collection efforts.

In response to the limitations inherent in supervised learning approaches, researchers have increasingly focused on developing unsupervised and weakly-supervised methods for enhancement. A significant advancement in this direction is Zero-DCE [9], which implements a novel deep curve estimation framework that eliminates the need for reference images. This approach learns a series of best-fitting light-enhancement curves through a non-reference training process, demonstrating remarkable adaptability to various illumination conditions. Another innovative solution is MBLLEN [10], which addresses the challenge through a specialized multi-branch architecture. The network employs dedicated sub-networks to independently process different enhancement factors, followed by an intelligent fusion mechanism that optimally combines the outputs from each branch.

Despite the remarkable progress achieved by deep learning methods in image enhancement, several critical challenges persist. First, deep learning models are often criticized for their poor interpretability. These models are frequently regarded as “black boxes,” with internal mechanisms and decision-making processes that are difficult to explain. Second, generalization remains a significant challenge. Models tend to perform poorly when confronted with unseen images, particularly those that deviate from the training data distribution. This limitation is particularly evident in low-light conditions, where high-quality labeled data are scarce. To address these limitations, it is imperative to develop paradigms that can dynamically adapt to varying scenarios.

RL offers a promising paradigm shift in addressing these problems through its capability for dynamic enhancement strategy optimization. Unlike traditional methods with fixed pipelines, RL-based approaches can adaptively select the most appropriate operations based on the specific features of input images, enabling more personalized and context-aware enhancement effects. This paradigm enables the system to automatically adjust to the variable navigation environment by tuning enhancement strategies. Ultimately, it aims to optimize the perception capabilities of MASS with respect to their surrounding environment.

Reference [11] marked the first successful application of RL to image enhancement. This research enables agent to learn optimal global enhancement sequences through iterative interactions with the environment, demonstrating the potential of RL in adaptive image processing. The methodology represents a significant departure from conventional enhancement techniques by introducing a dynamic, learning-based optimization paradigm. Building upon the foundation, [12] proposed an innovative pixel-level agent architecture that revolutionizes the enhancement approach. In this work, each pixel is treated as an agent, enabling localized action selection and quality optimization

III. METHODOLOGY

This study proposes a novel MADRL paradigm to address the enhancement of degraded image in MASS, as shown in Fig. 1. The framework conceptualizes low-light images as the environment and agents serves as the subject of image quality enhancement actions. The core components of MADRL comprises of observation $o \in O$, action $a \in A(o)$, reward $R(o, a)$ and policy $\pi(a|o)$. Agent groups Ω_θ interact with the environment I_t to obtain the current observation $o_t^{(i)}$, perform

joint actions $\Pi a_t^{(i)}$, and optimize the policy $\pi_t(a_t|o_t)$ based on new observation $o_{t+1}^{(i)}$ and reward r_{t+1}, with the goal of maximizing the cumulative reward. Through collaborative learning, the framework dynamically adjusts the enhancement strategy, progressively adapting to diverse low-light image scenarios. The whole procedure of the proposed method is summarized in Algorithm 1.

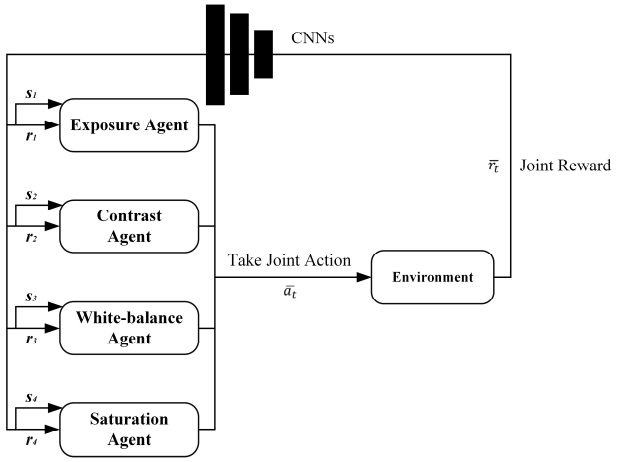


Fig. 1. Agent-environment interaction of MADRL for image enhancement

Algorithm 1 Image Enhancement with MADRL

procedure ENHANCEIMAGE($I_{degraded}$, $I_{enhanced}$)
  Initialization: MADRL Agent groups Ω_Θ
  for each agent $i \in \Omega$ **do**, initialization:
    Policy network $\pi^{(i)}$
    Value function $Q^{(i)}$
    Action space $a^{(i)}$
  end for
  repeat
    $t \leftarrow 0$
    while not termination_condition **do**
      $t \leftarrow t + 1$
      $s_t \leftarrow$ extract_features(image, CNN_model)
      $o_t \leftarrow s_t$
      for each agent $i \in \Omega$ **do**
        $a_t^{(i)} \leftarrow$ choose_action($o_t^{(i)}, \pi_t^{(i)}$)
      end for
      image $\leftarrow$ perform_actions(image, $\prod a_t^{(i)}, \Omega_t$)
      $r_t \leftarrow$ compute_reward(original_image, image, reward_function)
      $o_{t+1} \leftarrow$ extract_features(image, CNN_model)
      for each agent $i \in \Omega$ **do**
        update_policy($o_t, a_t, r_t, o_{t+1}, \pi_t, Q_t$)
      end for
      $o_t \leftarrow o_{t+1}$
    end while
  until termination_condition
  Return enhanced_image $I_{enhanced}$
end procedure

A. Features of Degraded Images as Observations

Previous research has demonstrated that synthetic data can serve as an effective supplement to real data in image enhancement tasks. Low-light images in maritime scenes are synthesized based on the method proposed in [13], and the generation process can be represented as Eq. (1).

$$I_{l(x,y)} = \beta(\alpha \cdot e_{R(x,y)})^{\Gamma} + \vartheta \quad (1)$$

where α and β are linear transformation parameters, X^{Γ} is the gamma transformation parameter. All of the parameters follow uniform distributions: $\alpha \sim U(0.8, 1)$, $\beta \sim U(0.5, 1)$, $\Gamma \sim U(1.5, 3)$. ϑ controls the brightness offset of the output image and takes values of either 0 or negative values.

In this study, degraded images of the sea surface are defined as the observation of the environment, and high-dimensional features are extracted through a customized scene feature analyzer. The analyzer is based on the VGG16 architecture, with adjustments made for maritime scenes.

In feature extraction classifier, images are resized to 224×224 and preprocessed with normalization, after which features are extracted through convolutional layers and fully connected layers. Additionally, during training, the MSE loss function and Adam optimizer are used, with a learning rate of 0.0001 and a batch size of 32, for a total of 200 epochs.

After feature extraction, each image outputs a 1x4-dimensional feature vector $F = [f_1, f_2, f_3, f_4]$, representing exposure, contrast, saturation, and white balance respectively. The vector directly reflects key attributes influencing the visual quality of the image. Finally, the vector F is mapped to the observation $O = [o_1, o_2, o_3, o_4]$ of the environment I_t in MADRL.

B. Image Enhancement Methods as Agent Groups

To achieve multi-dimensional optimization for low-light image enhancement, we propose a learning framework based on multi-agent collaborative. The agent groups consist of four specialized agents, responsible for exposure adjustment, contrast optimization, white balance correction, and hue adjustment. Each agent focuses on optimizing a specific visual attribute. Through collaborative work, they achieve fine-grained control over the image quality.

Action space of the agent groups is composed of an extendable set of image enhancement algorithms, as shown in Table I. The architecture allows each agent to independently choose the most appropriate enhancement operation based on specific requirements and image characteristics. To maintain stability during enhancement process and prevent excessive alterations from single action, the magnitude of actions is restricted. Each agent's action is limited to adjusting the relevant visual attribute by ±2% or ±5%, which not only ensures the stability of the enhancement process but also avoids irreversible outcomes caused by overly large actions.

TABLE. I CANDIDATE ACTION SPACE FOR LOW-LIGHT IMAGE ENHANCEMENT

Action	Multi-agent			
	Exposure	***Contrast***	***Color-temp***	***Saturation***
1	↑ 5%	↑ 5%	+5%	↑ 5%
2	↑ 2%	↑ 2%	+2%	↑ 2%
3	↓ 2%	↓ 2%	−2%	↓ 2%
4	↓ 5%	↓ 5%	−5%	↓ 5%

Based on the design of action space, agent groups achieve image enhancement through collaborative learning. In each iteration, agent selects an action based on current observation,

performs joint enhancement operation, evaluates the improvement in image quality, and then adjusts its strategy accordingly. Through repeated iterations, the agent groups gradually learn the optimal image enhancement strategy that adapts to diverse maritime conditions.

C. Image Quality Increments as Reward

In the low-light image enhancement task, the image visual perception characteristics and objective quality indicators after the multi-modal joint enhancement strategy is implemented by the multi-agent group in each iteration are used as the dynamic reward signals of reinforcement learning. In addition, a closed-loop feedback mechanism is formed by constructing a comprehensive evaluation system of multi-dimensional enhancement performance to drive the agents group to continuously optimize the image enhancement strategy. Each component of the reward function is set as follow.

Brightness Gain: It is used to measure the change in brightness levels before and after enhancement. The calculation formula is as Eq. (2).

$$\{\cdot EQ\cdot Br\cdot=\cdot \backslash F(1,N)\cdot\}\{\sum_{p\in P}\}\{\cdot \mathrm{EQ}\cdot \backslash \mathrm{B}(\backslash \mathrm{F}(I'(p),\cdot I(p)))\cdot\} \quad (2)$$

where $I(p)$ and $I'(p)$ represent the grayscale values of the degraded image and the enhanced image at pixel p respectively, and N is the total number of pixels.

Contrast Gain：It measures the change in local contrast before and after image enhancement. The Laplacian operator is used to be represented, as Eq. (3).

$$\{\cdot EQ\cdot Cr\cdot=\cdot \backslash F(\sum_{p\in P}|\nabla^2\cdot I'(p)|,\cdot \sum_{p\in P}|\nabla^2\cdot I(p)|)\cdot\} \quad (3)$$

where ∇^2 is the Laplacian operator, which is used to detect edges and details in the image.

In addition, several constraint terms have been added to reward function to limit actions that may lead to undesirable enhancement results.

Texture Quality：The component is measured by Gray-Level Co-occurrence Matrix (GLCM). The change in entropy values before and after enhancement is calculated to reflect the variation in texture complexity, as Eq. (4).

$$\{ EQ\ TQ = \backslash F(1,N)\ \}\sum_{p\in P}\{ EQ\ \backslash B(1 - \backslash F(|T(I'(p)) - T(I(p))|, max(T(I(p)), T(I'(p))) + \varepsilon))\ \} \quad (4)$$

Naturalness Penalty： The constraint is introduced to prevent image distortion caused by excessive enhancement. It can be calculated by computing the Bhattacharyya distance between the enhanced and the degraded, as Eq. (5).

$$\{ EQ\ NP = \backslash F(1,N)\ \}\sum_{p\in P} max(0,\ \lambda\cdot(\Delta E(p) - \delta) \quad (5)$$

Finally, the reward function is delineated as Eq. (6).

$$R = \phi\cdot(Br + Cr) - \psi\cdot(TQ + NP) \quad (6)$$

Here, ϕ and ψ are weighting coefficients used to balance the influence of each evaluation metric.

D. Value-based DRL

To achieve efficient collaboration among agent groups, a distributed learning architecture is adopted, allowing information to be shared and coordinated among agents. Action decisions of each agent are not only based on own observation but also consider the actions of other agents. Ultimately, a joint enhancement action $\Pi a_t^{(i)}$ is generated, and an intermediate state image is output.

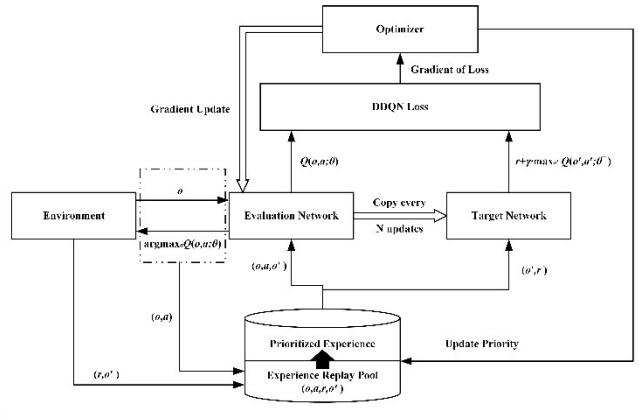


Fig. 2. Learning strategy for DDQN with PER

The action value function, which is denoted by $Q_\pi(o_t, a_t)$, is introduced as a mapping function to evaluate the expected return of joint action a_t with current observation o_t. Q-function predicts the expected return of joint action a_t given the observation o_t, and is defined as the cumulative discounted return, as Eq. (7).

$$U_t = R_t + \gamma R_{t+1} + \gamma^{\prime} R_{t+2} + \ldots = \{ \mathrm{EQ}\ \backslash \mathrm{i}\backslash \mathrm{su}(\mathrm{k}=0,\infty,\gamma^{\mathrm{K}})\ \} R_{t+k+1} \quad (7)$$

where γ is discount factor，$\gamma\in[0,1]$, indicating the importance of future rewards for immediate rewards.

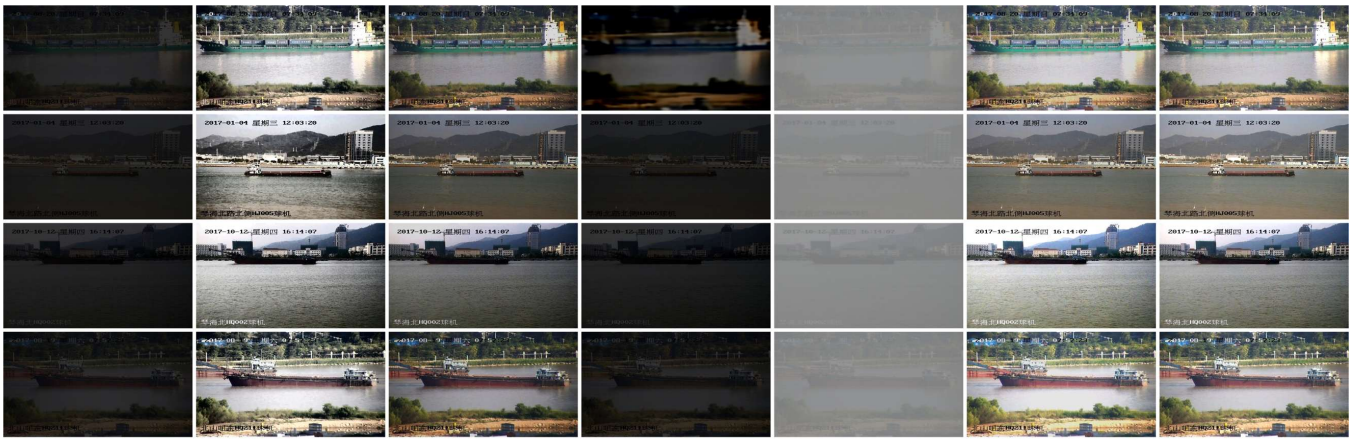
Optimal action value function $Q^*(o_t, a_t)$ is obtained by maximizing policy π, eliminating its influence and relying solely on current observation and action. It is determined by maximizing the expected return, as Eq. (8).

$$Q^*(o_t, a_t) = max(E[U_t|O_t = o_t, A_t = a_t]) \quad (8)$$

Based on the estimated values of Q-function, agent groups adopt an ε-greedy strategy for optimal joint action a_t^*, gradually improving the visual quality of the image, as Eq. (9).

$$a_t^* = \underset{a\in A}{\mathrm{argmax}}\ Q^*(o_t, a_t^{(i)}) \quad (9)$$

To stabilize the estimation of $Q^*(o_t, a_t)$ value and reduce overestimation issues, a neural network $Q(o, a; \omega)$ is used to approximate Q^*. Specifically, a Double Deep Q-Network (DDQN) is employed as the core learning method, updating



(a) Low-light (b) HE (c) MSR (d) LLNet (e) Zero-DCE (f) Proposed (g) Ground Truth

Fig. 1 Visual comparisons of low-light enhancement results from the Seaships synthetic dataset. From left to right: (a) low-light images, restored images, generated by (b) HE, (c) MSR, (d) LLNet, (d) Zero_DCE, (e) MADRL, and (f) Ground Truth, respectively.

the parameters ω to evaluate the policy. The target value $y_t^{(i)}$ is computed as Eq. (10).

$$y_t^{(i)} = r_{t+1} + \gamma \cdot Q(o_{t+1}, \underset{a}{\text{argmax}}, \ Q(o_{t+1}, a; \omega_t) ; \omega_t^-) \quad (10)$$

After collecting experience (*s, a, r, s'*), a prioritized experience replay is introduced to improve learning efficiency and model generalization. Agent stores experiences with larger temporal difference (TD) errors and assigns them higher priority for resampling.

The loss function is used to measure the difference between the predicted Q-values and the target. Based on this, the parameters of the neural network are updated as Eq. (11).

$$Loss = E_{(s,a,r,s')\sim U(D)}[(y_t^{(i)} - Q(o, a; \omega))^2] \quad (11)$$

where $U(D)$ represents the uniformly sampled distribution from the experience replay pool D. $Q(o, a; \omega)$ is the estimate of the current state-action pair by the policy network.

E. Training

The approach was implemented on a workstation equipped with an Intel(R) Core(TM) i9-10900K CPU @ 3.70GHz, 128GB RAM, and an NVIDIA GeForce RTX 3070Ti GPU, ensuring efficient computational power. The development environment is built on Python 3.12.4 and relies on PyTorch 2.5.1 deep learning framework for model training and testing.

Experimental dataset is constructed from three public sources, COCO, PASCAL VOC 2012, and Seaships (7000). Specifically, 144 ship images were selected from COCO dataset, 401 ship images from PASCAL VOC 2012, and 1973 images were randomly selected from Seaships, resulting in a total of 2518 normal light images. Using the method outlined in Section A of Chapter III, low-light images of different degradation levels were generated by adjusting the parameters α, β, Γ and ε. Ultimately, a dataset consisting of 10,072 images was constructed.

In the construction of enhancement model training pipeline, dataset was randomly splitt into in an 8:2 ratio. This partitioning resulted in a training set containing 8058 samples and test set comprising 2014 images. During model training phase, hyperparameters was set to ensure effectiveness and robustness. Specifically, agent groups are trained with a batch size of 16 for 1000 epochs. The network learning rate is set to 0.005, the reward decay is 0.95, and the memory size is 4096. The maximum enhancement adjustment step is set to max_step=10. The enhancement process for each image ends when max_step>10 or the observation $o_t^{(i)}$ of the degraded images exceeds a given threshold. Additionally, the training of the agent groups is based on DDQN, with the network consisting of multiple fully connected layers and using LeakyReLU activation functions.

IV. Experimental Evaluations

In this section, the experiments are described and both qualitative and quantitative evaluations of the proposed MADRL and other representative algorithms are performed. The effectiveness of the proposed method is further verified through the YOLO-v8 object detection framework from multiple perspectives.

A. Qualitative Evaluations

To validate the enhancement performance of algorithms under low-light conditions, two groups of synthetic low-light images have been selected with varying darkness levels for analysis, as shown in Fig. 3 and Fig. 4. Fig. 3 demonstrates images derived from the Seaships dataset, containing randomly selected synthetic ultra-dark images that simulate extreme low-light maritime observation conditions. Fig. 4 contains images synthesized from the COCO and PASCAL VOC datasets with relatively lower darkness levels, representing more common low-light scenarios. And a qualitative comparative analysis of visual enhancement effects is conducted across algorithms using images with varying darkness levels and diverse dataset origins.

While enhancing the global brightness, HE exhibits notable limitations: its nonlinear stretching mechanism leads to serious overexposure in highlight areas, coupled with conspicuous texture distortion caused by excessive contrast enhancement. Severe overexposure artifacts were persistently observed across both ultra-dark and moderately low-light image conditions. Although the Multi-Scale Retinex (MSR) algorithm demonstrates advantages in color restoration, its constrained dynamic range compression strategy fails to prevent highlight oversaturation during low-light

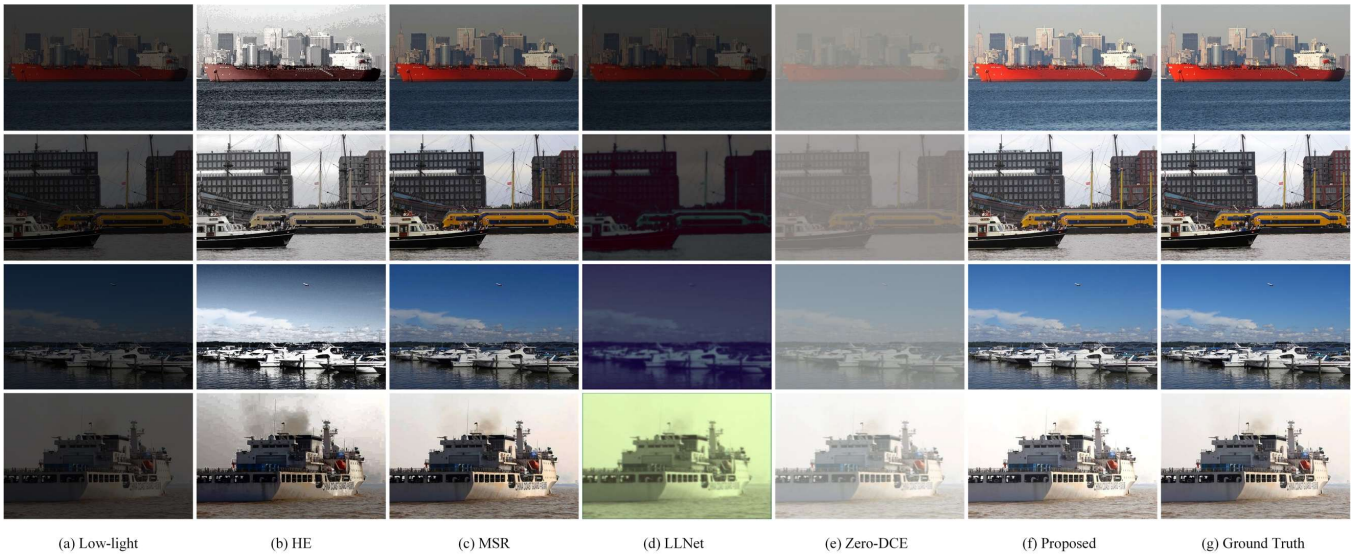


Fig. 2 Visual comparisons of low-light enhancement results from the synthetic dataset, i.e., COCO and PASCAL VOC. From left to right: (a) low-light images, restored images, generated by (b) HE, (c) MSR, (d) LLNet, (d) Zero_DCE, (e) MADRL, and (f) Ground Truth, respectively

enhancement, accompanied by chromatic aberration issues. The LLNet model manifests two critical limitations in enhancement process. Its end-to-end mapping mechanism inadequately captures dark channel prior information from low-light images, resulting in insufficient visual enhancement gains with minimal brightness improvement compared with degraded inputs. Besides, the absence of effective feature fusion module results in obvious blurring of the edges and contents of the output images. Constrained by its single-scale feature extraction architecture, Zero-DCE network demonstrates suboptimal brightness gains as evidenced in Fig. 3 and Fig. 4. , failing to achieve anticipated enhancement effects in subjective visual perception metrics. Moreover, inadequate noise suppression capability leads to salt-and-pepper noise artifacts.

In contrast, our MADRL framework achieves coordinated optimization through multi-agent collaboration, effectively improves holistic visual perception quality while maintaining brightness improvement and visual fidelity. Our approach avoids overexposure issues and demonstrates superior noise suppression capability.

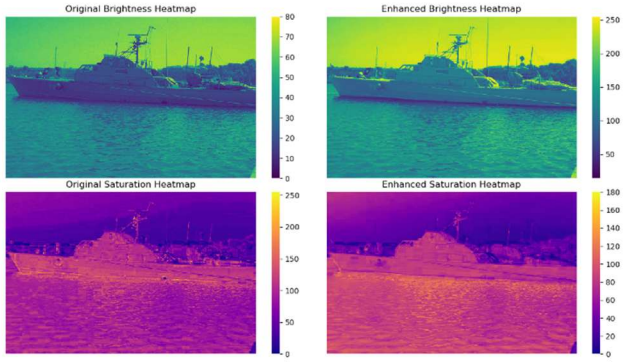


Fig. 3 Heatmap Visualization of Enhancement Intensity for Different Algorithms on Low-Light Maritime Images

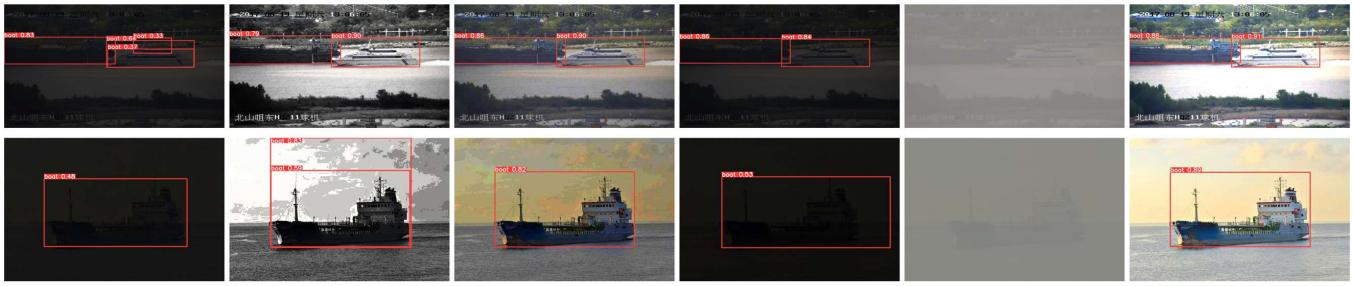
Additionally, Heatmap-based visualization comparing pre- enhancement and post-enhancement image was used to investigate the spatial distribution of enhancement effects and multi-agent collaboration. Derived from brightness, contrast and saturation measurements, these heatmaps utilize color gradients to indicate the intensity of respective visual attributes. Fig. 5 shows the comparison of the heat map before and after enhancement of a typical low-light image. The analysis reveals the dual capability of MADRL paradigm to substantially increasing brightness in dark regions while preventing highlight overexposure. For example, the hull of the ship is visualized as an area of higher volume in the brightness heat map, while its color gradient is lower in the low-light source image. Saturation heatmaps evidence well-balanced color restoration without oversaturation, with enhanced images exhibiting smoother saturation gradients, thereby validating the white-balance agent's adaptive regulation proficiency.

B. *Quantitative Evaluations*

This subsection presents the quantitative evaluations. To address the challenge of lacking high-quality reference images in real-world low-light scenarios, we employ a no-reference image quality assessment (NR-IQA) system for multi-dimensional quantitative analysis of various enhancement results. An evaluation matrix is established using four complementary metrics: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Feature Similarity Index (FSIM), and Visual Information Fidelity (VIF). PSNR focuses on evaluating brightness restoration accuracy, SSIM and FSIM analyze structural similarity and feature preservation respectively, while VIF quantifies information loss by simulating characteristics of the human visual system. Table II summarizes the quantitative evaluation results obtained from the test set images.

Table. II QUANTITATIVE RESULTS OF ENHANCEMENT METHOD

Methods	**Evaluation Indicators**			
	SSIM (↑)	***PSNR*** (↑)	***FSIM*** (↑)	***VIF*** (↑)
HE	0.5974	10.2025	0.7486	0.3213
Retinex	0.8156	17.5404	0.9126	0.5983
LLNet	0.8228	28.1698	0.7969	0.3029
Zero-DCE	0.8955	**32.9540**	0.9349	0.5215
MADRL	**0.8975**	29.0553	**0.9353**	**0.5987**



(a) Low-light (b) HE (c) MSR (d) LLNet (e) Zero-DCE (f) MADRL

It can be seen from Table II that the HE algorithm exhibits inferior performance in SSIM and FSIM metrics due to limitations in its global histogram equalization mechanism, confirming its deficiencies in structural preservation. MSR algorithm, although it improves the VIF value through dynamic range compression, its reliance on local contrast enhancement strategies results in substantially lower PSNR compared to deep learning-based approaches. The deep learning-based LLNet model outperforms traditional algorithms in preserving structural features (SSIM and FSIM). However, its end-to-end mapping mechanism encounters inherent bottlenecks in brightness compensation capability under low-light conditions, resulting in a PSNR of only 28.17 dB. Zero-DCE network improves brightness gain through a zero-reference learning framework, yet it exhibits inferiority in VIF metrics when compared with the method outlined, owing to the absence of a multi-scale feature fusion module.

Our proposed MADRL paradigm achieves state-of-the-art performance across all metrics. SSIM and FSIM reach 0.89 and 0.93 respectively, with VIF (0.59 ± 0.02) demonstrating 86.33% and 97.66% enhancement compared to conventional methods and deep learning LLNet methods respectively. This comprehensive performance superiority stems from the synergistic optimization of low-light enhancement, noise suppression, and feature preservation through its multi-agent collaborative mechanism, coupled with effective modeling of subjective perceptual quality via a reinforcement learning-based dynamic reward strategy. The results conclusively validate the framework's dual capability to enhance both naturalness and authenticity of image restoration under complex illumination conditions.

C. *Yolo-v8 based Target Dectection Results*

The study employs the YOLOv8 object detection architecture to evaluate images enhanced through various algorithms. It specifically assesses the operational effectiveness of MADRL-enhanced imagery in supporting MASS navigation systems under low-light marine conditions.

Specifically, by feeding low-light images, those enhanced by our proposed MADRL-based method, and those processed by other benchmark algorithms into the YOLOv8 model, we quantitatively evaluated the improvement in object detection performance through comparative analysis of confidence scores across different image groups. The comparative results are illustrated in Fig. 6. The results demonstrate that images enhanced by our method not only provide clearer maritime scene representations for MASS intelligent navigation systems, but also significantly improve object detection accuracy and reliability, evidenced by higher average precision and lower miss detection rates. This confirms the effectiveness of MADRL-based image enhancement in improving object detection accuracy under low-light maritime conditions, while providing a reliable visual perception framework for the MASS navigation system.

V. CONCLUSIONS

This study proposes a MADRL-based framework for low-light image enhancement in marine environments, aiming at improving target detection capabilities of MASS during night navigation. By incorporating a multi-agent system, the image enhancement task is decomposed into multiple visual feature dimensions, where each agent independently explores and learns using a distributed reinforcement learning algorithm. Specifically, double deep Q-learning with prioritized experience replay (PER). Meanwhile, through inter-agent information exchange and collaborative decision-making, the agent groups achieve comprehensive low-light image enhancement, effectively addressing the complex visual characteristics of visible-light images in low-light maritime environments. The results demonstrate that compared to traditional image enhancement algorithms, the proposed paradigm significantly improves both image quality and MASS nighttime target detection performance. In simulated low-light scenarios, images enhanced by our framework not only maintain high visual quality but also effectively highlight detected targets. Moreover, the proposed method achieves an average increase in target detection accuracy of 6.5%, further validating its effectiveness and superiority. Given its potential for integration into various shipborne nighttime navigation systems, this framework holds promise for enhancing the safety and reliability of MASS navigation.

ACKNOWLEDGMENT

The authors would like to thank the editors and anonymous reviewers for their valuable comments and suggestions to improve the presentation of this paper. This work was supported by the National Natural Science Foundation of China [Grant No.52231014] and the Team Project of Dalian Maritime University [Grant No.3132023511].

REFERENCES

[1] M. Ekstrom, “Digital image processing,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 28, no. 4, pp. 484–486, 1980, doi: 10.1109/TASSP.1980.1163437.

[2] S. Wang, J. Zheng, H.-M. Hu, and B. Li, “Naturalness Preserved Enhancement Algorithm for Non-Uniform Illumination Images,” *IEEE Trans. Image Process.*, vol. 22, pp. 3538–3548, 2013.

[3] K. G. Lore, A. Akintayo, and S. Sarkar, “LLNet: A deep autoencoder approach to natural low-light image enhancement,” *Pattern Recognit.*, vol. 61, pp. 650–662, 2017.

[4] C. Wei, W. Wang, W. Yang, and J. Liu, “Deep Retinex Decomposition for Low-Light Enhancement,” *ArXiv*, vol. abs/1808.04560, 2018, [Online].Available:https://api.semanticscholar.org/CorpusID:52008443

[5] Y. Jiang *et al.*, "EnlightenGAN: Deep Light Enhancement Without Paired Supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021, doi: 10.1109/TIP.2021.3051462.

[6] R. W. Liu, Y. Lu, Y. Guo, W. Ren, F. Zhu, and Y. Lv, "AiOENet: All-in-One Low-Visibility Enhancement to Improve Visual Perception for Intelligent Marine Vehicles Under Severe Weather Conditions," *IEEE Trans. Intell. Veh.*, vol. 9, no. 2, pp. 3811–3826, 2024, doi: 10.1109/TIV.2023.3347952.

[7] L. Li *et al.*, "Attention-oriented residual block for real-time low-light image enhancement in smart ports," *Comput. Electr. Eng.*, vol. 120, p. 109634, 2024,.

[8] W. Zhou, B. Li, and G. Luo, "Multi-Feature Fusion-Guided Low-Visibility Image Enhancement for Maritime Surveillance," *J. Mar. Sci. Eng.*, vol. 11, no. 8, 2023, doi: 10.3390/jmse11081625.

[9] C. Guo *et al.*, "Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 1777–1786. doi: 10.1109/CVPR42600.2020.00185.

[10] F. Lv, F. Lu, J. Wu, and C. Lim, "MBLLEN: Low-Light Image/Video Enhancement Using CNNs," presented at the British Machine Vision Conference (BMVC), 2018, p. 4. [Online]. Available: http://bmvc2018.org/contents/papers/0700.pdf

[11] J. Park, J.-Y. Lee, D. Yoo, and I. S. Kweon, "Distort-and-Recover: Color Enhancement Using Deep Reinforcement Learning," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5928–5936. doi: 10.1109/CVPR.2018.00621.

[12] R. Zhang, L. Guo, S. Huang, and B. Wen, "ReLLIE: Deep Reinforcement Learning for Customized Low-Light Image Enhancement," in *PROCEEDINGS OF THE 29TH ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA, MM 2021*, Assoc Comp Machinery; ACM SIGMM, 2021, pp. 2429–2437. doi: 10.1145/3474085.3475410.

[13] F. Lv, Y. Li, and F. Lu, "Attention Guided Low-Light Image Enhancement with a Large Scale Low-Light Simulation Dataset," *Int. J. Comput. Vis.*, vol. 129, no. 7, pp. 2175–2193, Jul. 2021, doi: 10.1007/s11263-021-01466-8.

Adaptive Estimator of Coordinates and Motion Parameters of a Maneuvering Object Based on Data from a Fixed Bearing Sensor

Khmarski P.A.

Microwave electronics and radio tomography laboratory
Institute of Applied Physics of the National Academy of Sciences of Belarus
Minsk, Republic of Belarus
https://orcid.org/0000-0003-3404-3917

Naumov A.O.

Department of Automation, Radar, and Transceiver Devices
Military Academy of the Republic of Belarus
Minsk, Republic of Belarus
https://orcid.org/0009-0006-0080-4713

***Abstract*—An adaptive trajectory estimator based on an IMM filter with a bank of five parallel filters has been developed to track maneuvering objects using data from a fixed bearing sensor. Simulation results indicate a 24–65% improvement in tracking accuracy compared to a third-order Kalman filter.**

Keywords— bearing sensor, trajectory tracking, IMM filter, maneuvering object, Kalman filter, simulation, data processing

I. INTRODUCTION

Bearing-only tracking represents a challenging scientific and engineering problem that arises in various practical applications including passive localization, acoustics, and navigation [1–5]. The main difficulties are due to the nonlinear nature of bearing measurements, strong dependence of estimation accuracy on the mutual geometry of the sensor and object, the need to adapt algorithms to changes in the motion dynamics, and the lack of direct range information. Traditional tracking algorithms employing single-model Kalman filters of various orders result in significant transients and higher overall measurement errors when the observed object performs a maneuver [6–10]. To compensate for dynamic errors, a promising approach is the use of multi-hypothesis estimators with inter-scan memory, known as IMM (Interacting Multiple Model) filters, which account for different motion models (MM) and adaptively estimate their posterior probabilities [8–14]. This work is devoted to the development of an adaptive trajectory estimator of coordinates and motion parameters for a maneuvering object based on an IMM filter, and presents results of its effectiveness under various observation conditions.

II. PROBLEM STATEMENT

A fixed bearing sensor provides primary bearing measurements β with a standard deviation not exceeding σβ and a data update period Δt. The object being observed may perform several types of motion: straight-line movement with constant velocity, uniformly accelerated motion, or execute horizontal turns with different overloads. It is required to estimate the state vector $\hat{\boldsymbol{\alpha}}_k$ of the observed object, including filtered estimates of bearing, velocities, accelerations, and the corresponding covariance matrices of estimation errors $\mathbf{R}_k$. The estimation must be performed under model uncertainty. The association of tracks with bearing measurements is assumed to be solved, and the object is assumed to be detected and observed at every data update interval.

III. ALGORITHM AND STRUCTURE OF THE TRAJECTORY ESTIMATOR

The use of a trajectory estimator considering only a single motion model inevitably causes transient errors and an increase in overall estimation error at the moments of model changes. To compensate for dynamic errors, an adaptive estimator based on an IMM filter is proposed, accounting for several motion models with probabilities of their mutual transitions [8–15]. To cover the most widespread types of motion, models of stationary object (0th-order polynomial), uniform motion (1st order), and uniformly accelerated motion (2nd order) were chosen [9,10,14,16]. Based on these models, Kalman filters of the corresponding orders were synthesized. To further minimize filtration errors during transients and maneuvers, Singer-model Kalman filters of 0th and 1st order are also included in the IMM structure [14]. The algorithm is as follows: the current bearing estimate enters all IMM channels tuned to different motion models. Each channel outputs conditional state estimates, which, after probability computations, are combined using the posterior model probabilities to form the unconditional state estimate and covariance matrix for the current step (see the structure in Figure 1).

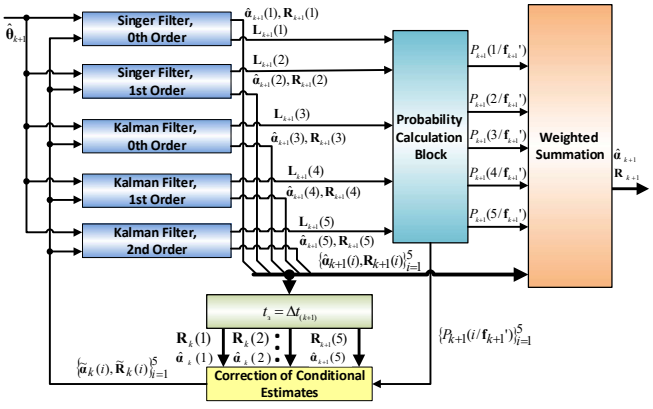


Fig. 1. Block diagram of the adaptive trajectory estimator

IV. COMPUTER SIMULATION RESULTS

To compare the overall estimation errors of different trajectory estimators, simulated primary measurements (σβ = 0.66°, Δt = 10 s) were used for an object moving at 220 m/s

at an altitude of 1 km and executing horizontal turns of 180° (scenarios 1–3), 270° (4–6), and 360° (7–9) with overloads of 1.1g, 1.5g, and 2.0g. Each scenario was repeated in 1000 trials. Kalman filters of 1st and 2nd order experienced loss of tracking during maneuvers. The 3rd-order Kalman filter operated stably but with dynamic estimation errors exceeding the single-measurement error by 2–3 times. The proposed IMM filter provided a significant increase in overall accuracy: a gain of 24.4–44.1% for 180° turns, 25.9–59.6% for 270° turns, and 10–65% for 360° turns compared to the 3rd-order Kalman filter (see Fig. 2).

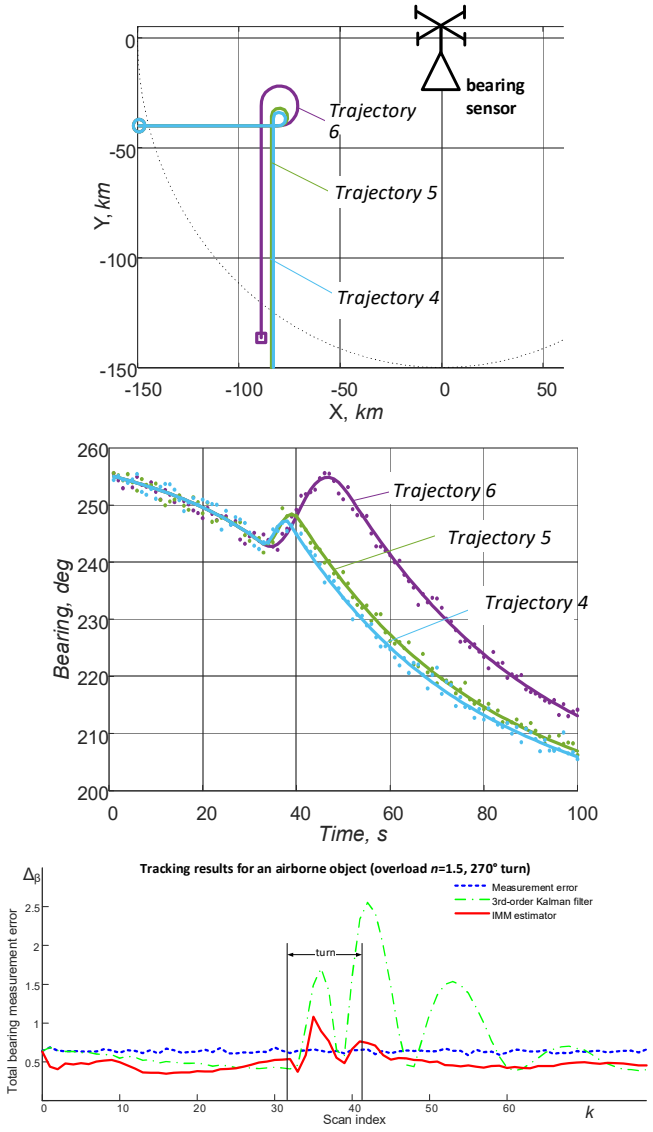


Fig. 2. Example of computer simulation results (trajectories 4–6)

V. Conclusion

An adaptive estimator of coordinates and motion parameters of a maneuvering object using data from a fixed bearing sensor is proposed and analyzed. The structure of the estimator includes five parallel filters of different orders. Computer simulation results for various object trajectories, including 180°, 270°, and 360° turns with overloads from 1.1g to 2.0g, provided quantitative assessments of its efficiency. The IMM-based estimator demonstrated sustainable tracking both on straight segments and during intensive maneuvers up to 2g overload. Promising future directions include optimization of the estimator parameters for specific object types, expansion of the motion model set, adaptation for use on a moving bearing sensor and in multiposition systems, and investigation of the estimator's applicability under various interference conditions.

References

[1] Stepanov O.A., Isaev A.M., Motorin A.V., Litvinenko Yu.A., Zolotarevich V.P. Recursive and Nonrecursive Algorithms Applied to Navigation Data Processing: Differences and Interrelation with Factor Graph Optimization Algorithms // Proc. 31st Saint Petersburg Int. Conf. ICINS 2024, St. Petersburg, Russia, 2024, pp. 450–452.

[2] Voronina N.G. Solving the Problem of Tracking Marine Object Trajectories Using Angular Information // Underwater Research and Robotics. 2019. No. 3(29). P. 13–23.

[3] Voronina N.G., Shafranyuk A.V. Algorithm for constructing trajectories of a maneuvering object based on bearing-only information using the Basis Pursuit method // J. Phys.: Conf. Ser. 2021. Vol. 1864.

[4] Voronina N.G., Shafranyuk A.V. Piecewise-stationary filtering in trajectory analysis tasks // Proc. Conf. "Control in Marine Systems". 2020. P. 30–33.

[5] Solonar A.S., Khmarski P.A., Naumov A.O., Juraev D.A., Muhammedov B.M. The use of numerical Monte Carlo integration to verify the physical feasibility of a trajectory based on surveillance radar data // Stochastic Modelling and Computational Sciences. 2023. 3(1). P. 59–73.

[6] Khmarski P.A., Solonar A.S. Features of implementing adaptive discrete quasi-linear filters for trajectory parameters // BGUIR Reports. 2012. No. 8(70). P. 57–64.

[7] Gorshkov S.A., Solonar A.S. Comparison of methods of adaptive discrete filtering of coordinates of maneuvering targets // Information-Measurement and Control Systems. 2006. Vol. 4. No. 6. P. 14–30.

[8] Dmitriev S.P., Stepanov O.A. Multi-alternative filtering in navigation information processing problems // Radiotekhnika. 2004. No. 7. P. 11–17.

[9] Bar-Shalom Y., Li X.R. Tracking and Data Association. Part 2. Moscow: Bauman Moscow State Technical University, 2011. 239 p.

[10] Blackman S., Popoli R. Design and analysis of modern tracking systems. Boston: Artech House, 1999. 1230 p.

[11] Solonar A.S., Khmarskiy P.A., Mihalkovskiy A.A., Tsuprik S.V. Features of trajectory measurement of coordinates and motion parameters of ground objects in on-board optical-location systems // 25th St. Petersburg Int. Conf. on Integrated Navigation Systems (ICINS), St. Petersburg, 2018, pp. 1–5. doi:10.23919/ICINS.2018.8405853.

[12] Solonar A.S., Khmarski P.A., Mihalkovskiy A.A. Features of filtering coordinates and motion parameters of an object during a stabilized turn // BGUIR Reports. 2013. No. 4(74). P. 67–73.

[13] Blom H.A.P., Bar-Shalom Y. The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients // IEEE Trans. Autom. Control. 1988. Vol. 33, No. 8. Aug. P. 780–783. doi:10.1109/9.1299

[14] Solonar A.S., Khmarski P.A., Tsuprik S.V. Tracking estimator of ground target coordinates and motion parameters using onboard optical location system data // Gyroscopy Navig. 14, 244–258 (2023).

[15] Solonar A.S., Khmarski P.A. Main problems of trajectory processing and approaches to their solution within multitarget tracking frameworks // J. Phys.: Conf. Ser. 2021. Vol. 1864. Article ID 012004. https://doi.org/10.1088/1742-6596/1864/1/012004

[16] Li X.R., Jilkov V.P. Survey of maneuvering target tracking. Part I: dynamic models // IEEE Trans. Aerospace Electron. Syst. 2003. Vol. 39, No. 4, pp. 1333–1364. DOI:10.1109/TAES.2003.1261132

Astronomical Calibration of a Strapdown Astroinertial Navigation System using a Simple Testbed

N.N. Vasilyuk

SPC Elektrooptika, LLC
Moscow, Russia
ORCID: 0000-0003-2317-8066

***Abstract* — Astronomical calibration is understood as the use of stellar observations to determine the geometric parameters of a strapdown astroinertial navigation system that systematically affect the errors in determining its geographic coordinates. The design of the astronomical calibration testbed is quite simple: it only needs to rotate the calibrated system and maintain its attitude between rotations. The testbed is not required to measure the attitude of the system being held. The set of calibrated parameters is divided into three groups that are calibrated sequentially, using a common set of astronomical observations. That is, one observation session is sufficient for a complete astronomical calibration of the system. The first and second groups of parameters are calibrated based on the results of comparing the coordinates of stars detected in digital images and calculated from star catalog data. The third group of parameters is calibrated based on the results of comparing the system attitude increments calculated from inertial and astronomical measurements. The errors in astronomical calibration of parameters within each group are quite small and amount to units of arc seconds.**

Keywords—star tracker, astroorientation, astroinertial navigation system, astronomical calibration, distortion calibration, camera intrinsic elements

I. Introduction

The strapdown astroinertial navigation system (AINS) consists of an astro-vision unit (AVU), a strapdown inertial measurement unit (IMU) and a time keeper. All AINS components are assembled into a single structure that has no moving parts (Fig. 1). AINS consists of $I \geq 1$ identical digital cameras with numbers $i = 1…I$, rigidly fixed on a common base. The field of view of each camera is at least 10°. The exposure start times of all cameras are synchronized with each other. The mechanical connection of AINS and IMU ensures that their mutual attitude remains unchanged throughout the AINS operation period.

The AVU receives images of the starry sky, detects and recognizes the observed stars and determines its astro-orientation relative to the Geocentric Celestial Reference System (GCRS). For recognition with an acceptable error probability, it is necessary to observe at least four stars. From the astro-orientation of the AVU, the astro-orientation of the AINS is calculated, which is used to correct the growing errors of inertial navigation.

Each digital camera of the AVU receives an image of its section of the starry sky, in which images of individual stars are detected. From these images, the coordinates of the direction vectors of the detected stars are calculated in the camera frame. To calculate the coordinates of the direction vectors, the camera intrinsic parameters are used: focal length, coordinates of the principal point, distortion coefficients. The actual values of these parameters differ from the design ones and must be determined as a result of individual calibration of the camera.

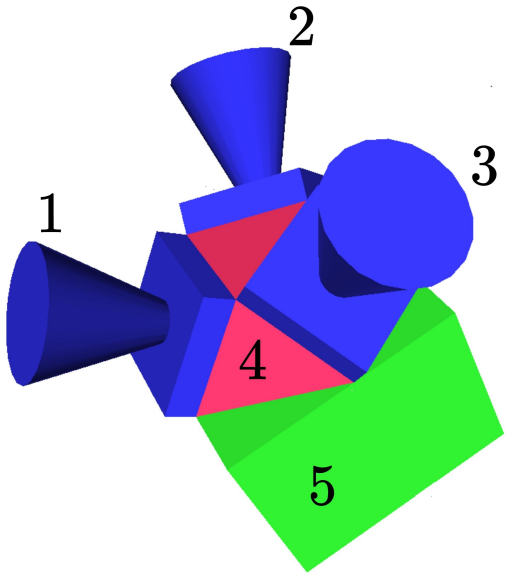


Fig. 1. Arrangement of the strapdown AINS. 1,2,3 – AVU cameras; 4 – AVU base; 5 – IMU.

Star recognition is performed using a constellation with extended angular dimensions, assembled from stars detected by all cameras. This approach allows determining the astro-orientation of the AVU even in cases where each camera separately observes less than four stars. In addition, expanding the angular dimensions of the constellation improves the accuracy of determining the astro-orientation of the AVU. To assemble an extended constellation, it is necessary to determine the mutual attitudes of the cameras within the AVU. In this case, the AVU can be considered as a single virtual camera with a complex field of view, composed of several real cameras.

The AINS must determine its coordinates, velocity vector, and attitude relative to the Earth's surface. This set of AINS navigation parameters is defined for the IMU measurement frame. The AVU determines the astro-orientation of its virtual camera frame. To determine the AINS astro-orientation from the AVU astro-orientation, it is necessary to determine the attitude of the virtual camera relative to the IMU. This attitude differs from the design attitude due to AINS assembly errors.

The dense arrangement of the AINS structure prevents direct measurement of the relative attitudes of the cameras in

the AVU and the attitude of the AVU relative to the IMU. Therefore, these attitudes are determined as a result of calibration. Here, a difficulty arises due to the different physical nature of the quantities measured by the calibrated devices. The AVU determines the directions to infinitely distant light sources. The IMU measures the components of the angular rate and the specific force vectors of the moving AINS. To calibrate the relative attitude of the AVU and the IMU, the reference light sources must be observed during all calibration turns of the AINS. The directions to these sources must be specified relative to the rotation axes of the calibration testbed with an error of units of arc seconds. The mechanical turns of the AINS must also be measured with "arc seconds errors". Such initial data require the construction of a fairly complex testbed for calibrating the assembled AINS.

The requirements for the design of the AINS calibration testbed can be significantly simplified if we take into account that navigation-class IMUs are always calibrated during manufacture. For their calibration, high-precision rotary tables are used, which are not subject to requirements for binding to optical radiation sources. The calibrated IMU itself is a device for accurately measuring the components of the angular rate vector relative to its measurement frame.

On the other hand, stars can be considered as infinitely distant point light sources. The angular coordinates of these sources are known with very high accuracy, and their concentration in the upper hemisphere is unattainable by any technical means. The algorithms for detecting and recognizing stars are standard algorithms for the operation of the AVU. Therefore, for the calibration of the AINS, it makes sense to consider astronomical calibration, when the rotation of the AINS is measured by the IMU gyroscopes, and the observed stars are used as point light sources. In this case, the testbed is needed only to ensure the convenience of turning and fixing the AINS under the starry sky in the desired direction. There are no requirements for the accuracy of angular rotation or positioning.

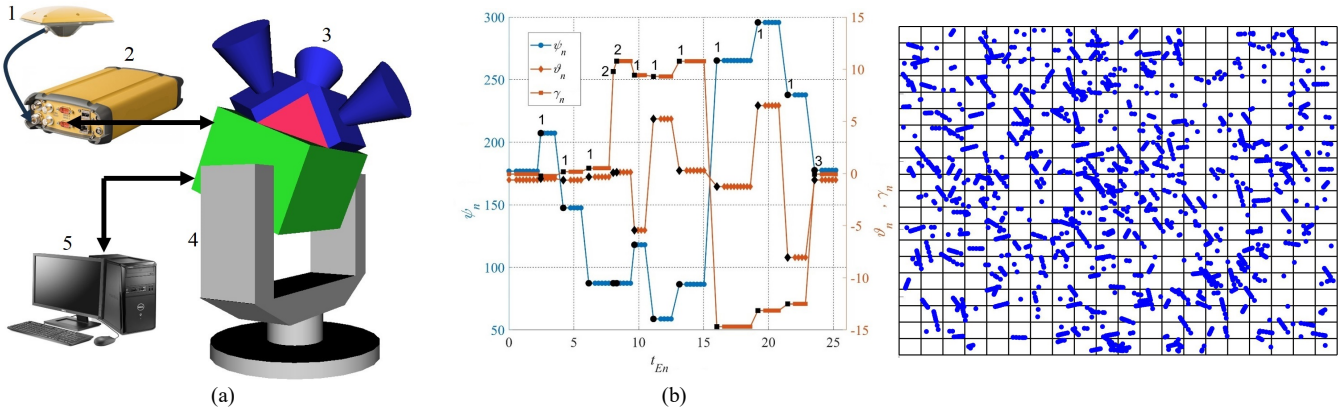
Astronomical calibration is divided into groups of parameters: intrinsic elements of individual cameras (1st group), parameters of relative attitude of cameras within the AVU (2nd group), parameters of AVU attitude relative to the IMU (3rd group). The values of the parameters in each group are determined in separate algorithmic blocks, which are executed sequentially and process a common set of raw calibration data. That is, for a complete astronomical calibration of the AINS, one calibration session is sufficient.

II. Receiving and Preprocessing of Calibration Data

Raw data for the AINS calibration are obtained on a simple testbed. Coordinates and time at the calibration location are obtained from a global navigation satellite system (GNSS) receiver (Fig. 2.a). The AINS calibration trajectory contains several successive static positions connected by turns. Certain restrictions are imposed on the sequence of AINS turns, which will be discussed below.

During the calibration rotations, N frames with images of the starry sky are recorded from each camera to the external storage medium. All cameras start to expose the next frame with the number $n = 1…N$ synchronously, at the same exposure start moment t_{En}, determined on the UTC(SU) time scale. The AINS attitude at the time t_{En} is described by the unknown angles of the yaw ψ_n, pitch ϑ_n and roll γ_n (Fig. 2.b.). Frames from the cameras are recorded simultaneously with the IMU measurements. Images of stars in each frame are first detected and then recognized (Fig. 2.c.).

After passing the calibration trajectory, N frames with exposure start times t_{En} have recorded for each camera of the AVU. Each frame contains R_n^i recognized stars, $i = 1…I$, $n = 1…N$. The total numbers of stars R_Σ^i and R_Σ, recognized in all frames of camera i and in all frames of all cameras: $R_\Sigma^i = \sum_{n=1}^{N} R_n^i$, $R_\Sigma = \sum_{i=1}^{I} R_\Sigma^i$. Within frame n of camera i, the stars are assigned serial numbers $r_n^i = 1...R_n^i$. With continuous numbering of the stars detected in all frames of camera i, the star with number r_n^i is assigned serial number $r^i = r_n^i + \sum_{q=1}^{n-1} R_q^i$: first, the stars of the first frame are numbered, then the second, etc. With continuous numbering of the stars detected in all frames of all cameras, the star with number r^i is assigned serial number $r = r^i + \sum_{q=1}^{i-1} R_\Sigma^q$: first, the stars detected in all frames of camera 1 are numbered, then those of camera 2, etc. The position of the brightness center of the star with the number r_n^i is specified by raster coordinates $[\tilde{h}^i_{n,r_n^i}, \tilde{w}^i_{n,r_n^i}]^\mathsf{T}$ containing the lens distortion [1].



(a) (b)

Fig.2. Receiving and preprocessing data for AISN astronomical calibration. (a) Calibration testbed arrangement: 1,2 – GNSS antenna and receiver; 3 – AINS; 4 – mount; 5 – data storage and mount control. (b) Typical AINS calibration trajectory: t_{En}, ; ψ_n, ϑ_n, γ_n, arc degrees. (c) Positions of brightness centers of stars in the image plane, detected and recognized in all frames obtained by single camera during the moving through the calibration trajectory.

Each detected star with number r_n^i is assigned the record in the star catalog, which allows calculating the coordinates $\mathbf{b}^i_{BCRS\,n,r_n^i}$ of its direction vector relative to the Barycentric Celestial Reference System (BCRS). In these coordinates, it is necessary to take into account the relativistic aberration of light arising from the orbital motion of the Earth. This is done by converting the vector $\mathbf{b}^i_{BCRS\,n,r_n^i}$ into the vector $\mathbf{g}^i_{GCRS\,n,r_n^i}$ specified in the Geocentric Celestial Reference System (GCRS). The arrays of raster coordinates $\{[\tilde{h}^i_{n,r_n^i}, \tilde{w}^i_{n,r_n^i}]^\mathsf{T}\}_{r_n^i=1}^{R_n^i}$ and catalog coordinates $\{\mathbf{g}^i_{GCRS\,n,r_n^i}\}_{r_n^i=1}^{R_n^i}$ obtained for all frames of all cameras are the initial data for the astronomical calibration of the AINS [1].

III. Processing of Calibration Data

A. Calibration of the 1st group of parameters

The vector $\mathbf{p}^i = [F^i\ h^i_O\ w^i_O\ k^i_1\ k^i_2]^\mathsf{T}$ of intrinsic elements is calibrated for each camera $i = 1\ldots I$ separately, where F^i is the focal length; h^i_O, w^i_O are the raster coordinates of the principal point of the image; k^i_1, k^i_2 - are the radial distortion coefficients [1].

Stars are observed through the atmosphere and atmospheric refraction must be taken into account in the observation results. To calibrate the 1st group of parameters, refraction is taken into account in the catalog coordinates $\{\mathbf{g}^i_{GCRS\,n,r_n^i}\}_{r_n^i=1}^{R_n^i}$. For this purpose, each vector is first projected onto the International Terrestrial Reference System (ITRS): $\mathbf{g}^i_{ITRS\,n,r_n^i} = \mathbf{S}^{GCRS}_{ITRS}(t_{En})\mathbf{g}^i_{GCRS\,n,r_n^i}$, where $\mathbf{S}^{GCRS}_{ITRS}(t_{En})$ is the matrix of the Earth's rotation relative to GCRS, calculated for the observation epoch t_{En}. Then, the vector $\mathbf{g}^i_{ITRS\,n,r_n^i}$ is projected onto the local East-North-Up (ENU) coordinate system: $\mathbf{g}^i_{ENU\,n,r_n^i} = \mathbf{S}^{ITRS}_{ENU}(B,L)\mathbf{g}^i_{ITRS\,n,r_n^i}$, where $\mathbf{S}^{ITRS}_{ENU}(B,L)$ is the matrix of coordinate transformation from ITRS to ENU; B, L are the latitude and longitude of the calibration location. Finally, the vector $\mathbf{g}^i_{ENU\,n,r_n^i}$ is transformed into the vector $\mathbf{p}^i_{ENU\,n,r_n^i}$, which takes into account astronomical refraction on a plane-parallel model of the standard atmosphere. As a result of this chain of transformations, the array $\{\mathbf{p}^i_{ENU\,n,r_n^i}\}_{r_n^i=1}^{R_n^i}$ of reference values of the direction vectors of the detected stars is obtained from the array $\{\mathbf{g}^i_{GCRS\,n,r_n^i}\}_{r_n^i=1}^{R_n^i}$ of the catalog coordinates [1].

Each camera i has its own Camera Frame CF^i, onto which the vectors $\{\mathbf{p}^i_{ENU\,n,r_n^i}\}_{r_n^i=1}^{R_n^i}$ must be projected. The design of the calibration testbed is very simple, so the attitude of CF^i relative to ENU at the time of shooting frame n is unknown. To find this attitude, it is necessary to include the rotation angles $\boldsymbol{\psi}^i_n = [\psi^i_n\ \vartheta^i_n\ \gamma^i_n]^\mathsf{T}$ of CF^i relative to ENU at the time t_{En} in the vector of estimated parameters, together with $\mathbf{p}^i$. Thus, the entire vector $\mathbf{x}^i_{\{1\}}$ of estimated parameters for calibrating the 1st group of AINS parameters will contain only 5 "useful" parameters and $3N$ "interfering" parameters that need to be estimated to obtain the useful parameters: $\mathbf{x}^i_{\{1\}} = [\boldsymbol{\psi}^{i\,\mathsf{T}}_1\ \boldsymbol{\psi}^{i\,\mathsf{T}}_2\ \ldots\boldsymbol{\psi}^{i\,\mathsf{T}}_N\ \mathbf{p}^{i\,\mathsf{T}}]^\mathsf{T}$.

To calculate the vector $\mathbf{x}^i_{\{1\}}$, the vector of residuals $\mathbf{e}^i_{\{1\}}(\mathbf{x}^i_{\{1\}}) = [\mathbf{e}^{i\,\mathsf{T}}_{1,1}\ldots\mathbf{e}^{i\,\mathsf{T}}_{1,R^i_1}\ldots\mathbf{e}^{i\,\mathsf{T}}_{N,1}\ldots\mathbf{e}^{i\,\mathsf{T}}_{N,R^i_N}]^\mathsf{T}$ of dimension $2R^i_\Sigma \times 1$ is constructed (Fig.3.a.), where $\mathbf{e}^i_{n,r_n^i} = [\hat{x}^i_{n,r_n^i} - x^i_{n,r_n^i}\ \ \hat{y}^i_{n,r_n^i} - y^i_{n,r_n^i}]^\mathsf{T}$; $[\hat{x}^i_{n,r_n^i}\ \ \hat{y}^i_{n,r_n^i}]^\mathsf{T}$ are vector coordinates calculated from the vector $\mathbf{p}^i_{ENU\,n,r_n^i}$ using the values ψ^i_n, ϑ^i_n, γ^i_n, $\mathbf{p}^i$; $[x^i_{n,r_n^i}\ \ y^i_{n,r_n^i}]^\mathsf{T}$ are vector coordinates calculated from $[\tilde{h}^i_{n,r_n^i}, \tilde{w}^i_{n,r_n^i}]^\mathsf{T}$ using the value $\mathbf{p}^i$ [1].

The results of calibration of the 1st group of parameters are the set $\{\hat{\mathbf{p}}^i\}_{i=1}^I$ of vectors and their covariance matrices $\{\mathbf{P}^i_{\delta\mathbf{p}}\}_{i=1}^I$ contained in vectors $\hat{\mathbf{x}}^i_{\{1\}} = \arg\min_{\mathbf{x}}(\mathbf{e}^{i\mathsf{T}}_{\{1\}}(\mathbf{x}^i_{\{1\}})\mathbf{e}^i_{\{1\}}(\mathbf{x}^i_{\{1\}}))$ and covariance matrices $\mathbf{P}^i_{\delta\mathbf{x}}$, $i = 1\ldots I$ [1]. As a result of calibration, it is possible to

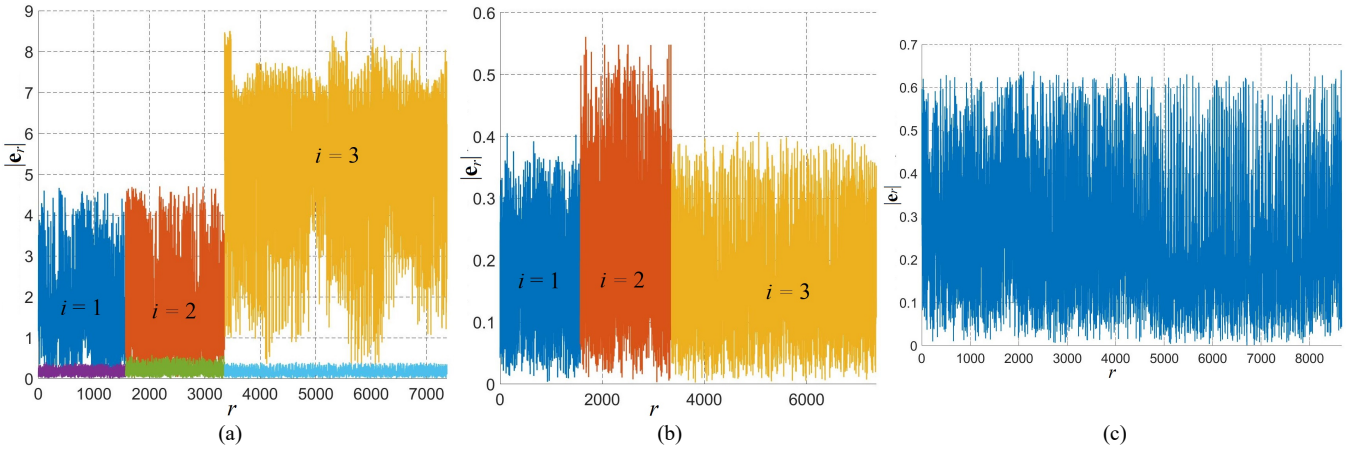


Fig.3. Residuals modules $|\mathbf{e}_r| = |\mathbf{e}^i_{n,r_n^i}|$ in separate stars with serial numbers $r = 1\ldots R_\Sigma$, expressed in units of pixel size, when calibrating the 1st and 2nd groups of AINS parameters for $I = 3$. (a) Residuals before (blue, orange, yellow) and after (violet, green, light blue) calibration of the 1st group of parameters. (b) Residuals after calibration of the 1st group of parameters on an enlarged scale. (c) Residuals after calibration of the 2nd group of parameters.

obtain the root-mean-square errors in determining useful parameters σ_F = 2 μm, $\sigma_{\hbar O}$ = 0.7 pixels, σ_{wO} = 0.6 pixels and interfering parameters σ_ψ = 11″ , σ_ϑ = 8″ , σ_γ = 11″ with the angular pixel size of 13″ and the final residuals $|\mathbf{e}^i_{n,r^i_n}| \le 0,5$ pixel (Fig. 3.b.).

B. Calibration of the 2nd group of parameters

The second group of calibrated parameters consists of sets $\boldsymbol{\alpha}^i = [\alpha^i_\psi\ \alpha^i_\vartheta\ \alpha^i_\gamma]^\mathsf{T}$ of attitude angles of the camera i = 2…I relative to camera 1. In this case, the attiude angles of camera 1 are considered zero $\boldsymbol{\alpha}^1 = [0\ 0\ 0]^\mathsf{T}$ [2]. By the time the calibration of the 2nd group of parameters begins, the 1st group has already been calibrated and arrays $\{\{\boldsymbol{\xi}^i_{n,r^i_n}\}_{r^i_n=1}^{R^i_n}\}_{n=1}^{R^i_\Sigma}$, $\boldsymbol{\xi}^i_{n,r^i_n} = [x^i_{n,r^i_n}\ \ y^i_{n,r^i_n}]^\mathsf{T}$, i = 1…I of vector coordinates of the brightness centers of star images have been obtained, which take into account the calibrated values of the intrinsic elements of individual cameras.

To calibrate the 2nd group of AINS parameters, it is sufficient to estimate only the attitude angles $\{\boldsymbol{\psi}^1_n\}_{n=1}^N$ of camera i = 1 as "inferring" parameters. The attitude of camera with number i = 2…N relative to ENU at time t_{En} is found through six successive rotations: first three rotations at variable "inferring" angles $\boldsymbol{\psi}^1_n$, and then three rotations at constant "useful" angles $\boldsymbol{\alpha}^i$. Thus, the entire vector $\mathbf{x}_{\{2\}}$ of estimated parameters for calibrating the 2nd group of AINS parameters contains 3N "inferring" parameters and 3(I−1) "useful" parameters: $\mathbf{x}_{\{2\}} = [\boldsymbol{\psi}^{1\mathsf{T}}_1\ \boldsymbol{\psi}^{1\mathsf{T}}_2 \ldots \boldsymbol{\psi}^{1\mathsf{T}}_N\ \boldsymbol{\alpha}^{2\mathsf{T}} \ldots \boldsymbol{\alpha}^{I\mathsf{T}}]^\mathsf{T}$ [2].

The residuals vector $\mathbf{e}_{\{2\}}(\mathbf{x}_{\{2\}})$ is constructed for all detected stars in all frames of all cameras and has the dimension $2R_\Sigma \times 1$. The vector $\mathbf{e}_{\{2\}}(\mathbf{x}_{\{2\}})$ consists of pairs of coordinate differences $\mathbf{e}^i_{n,r^i_n} = \hat{\boldsymbol{\xi}}^i_{n,r^i_n}(\boldsymbol{\psi}^1_n, \boldsymbol{\alpha}^i) - \boldsymbol{\xi}^i_{n,r^i_n}$, where $\hat{\boldsymbol{\xi}}^i_{n,r^i_n}(\boldsymbol{\psi}^1_n, \boldsymbol{\alpha}^i) = [\hat{x}^i_{n,r^i_n}\ \ \hat{y}^i_{n,r^i_n}]^\mathsf{T}$ are the vector coordinates of the central projection of the direction specified by the vector $\mathbf{p}^i_{ENU\ n,r^i_n}$, onto the plane of frame n of camera i. The value of the vector of calibrated parameters is found by minimizing the modulus of the residual vector: $\hat{\mathbf{x}}_{\{2\}} = \arg\min_{\mathbf{x}}(\mathbf{e}^\mathsf{T}_{\{2\}}(\mathbf{x}_{\{2\}})\mathbf{e}_{\{2\}}(\mathbf{x}_{\{2\}}))$. The results of calibration of the 2nd group of parameters are sets of vectors $\{\hat{\boldsymbol{\alpha}}^i\}_{i=2}^I$ and their covariance matrices $\{\mathbf{P}^i_{\boldsymbol{\alpha}}\}_{i=2}^I$, contained in the vector $\hat{\mathbf{x}}_{\{2\}}$ and its covariance matrix $\mathbf{P}_{\delta\mathbf{x}}$. As a result of calibration, it is possible to obtain the root-mean-square errors of determining the "useful" parameters $\sigma_{\alpha\psi}$, $\sigma_{\alpha\vartheta}$, $\sigma_{\alpha\gamma}$ < 1.5″ for i = 2.3 and the "interfering" parameters σ_ψ= 1.1″ , σ_ϑ, σ_γ < 0.6″ with the angular pixel size of 13″ and the final residuals $|\mathbf{e}^i_{n,r^i_n}| \le 0,6$ pixel (Fig. 3.c.). The final residuals after calibration of the 2nd group of parameters include the final residuals after calibration of the 1st group of the parameters.

When calibrating the 2nd group of parameters, the errors in determining the "interfering" parameters decreased. This is due to the increase in the angular size of the constellation being processed. To calibrate the 1st group of parameters, the constellations are observed within the fields of view of individual cameras and processed separately. To calibrate the 2nd group, several constellations observed by individual cameras are computationally combined into single constellation. The effective angular sizes of the constellation being processed increase with such a combination.

C. Calibration of the 3rd group of parameters

To determine the 3rd group of calibrated parameters, the AVU is considered as a virtual camera with a complex field of view obtained by combining the fields of view of real cameras [3]. The Virtual Frame (VF) is associated with the virtual camera. VF can either be rotated relative to CF[1] by fixed, preset angles $\boldsymbol{\psi}^{VF} = [\psi^{VF}\ \vartheta^{VF}\ \gamma^{VF}]^\mathsf{T}$, or identically coincide with CF[1]. In this case, $\boldsymbol{\psi}^{VF} = [0\ 0\ 0]^\mathsf{T}$. The angular coordinates of the stars observed by individual physical cameras i = 1…I and measured relative to CF^i are transformed to VF through the transformations specified by the direction cosines matrices $\mathbf{S}^{CF^i}_{CF^1}(\boldsymbol{\alpha}^i)$ and $\mathbf{S}^{CF^1}_{VF}(\boldsymbol{\psi}^{VF})$. The IMU measures the components of the angular rate and specific force vectors in its Measurement Frame (MF). Calibration of the 3rd group of parameters is the determination of the attitude of the VF relative to the MF, which is specified by three angles $\boldsymbol{\beta} = [\beta_\psi\ \beta_\vartheta\ \beta_\gamma]^\mathsf{T}$ [3]

When the AINS turns, the VF and MF make the same turn. The simple calibration testbed does not measure the magnitude and direction of this turn. The calibrated AVU determines its attitude relative to the stars with high accuracy. The attitude of the IMU relative to external reference points on the simple testbed can only be determined by gyrocompassing. Gyrocompassing is a very rough method, the root-mean-square error of the yaw is about 2′, and of the roll and pitch about 10″- 15″.

However, the navigation-class IMU is a high-precision angular rate vector meter. By integrating this rate, it is possible to obtain the IMU attitude increment over the time interval between the astronomical measurements of the AVU. The AVU attitude increment over the same time interval is calculated from its own measurements. Thus, on a simple testbed, the AVU attitude relative to the IMU should be calibrated not by attitude measurements, but by their increments.

The AVU determines the attitude of the VF relative to the GCRS at time t_{En} based on the observations of stars by all cameras. This orientation is represented as the matrix $\mathbf{S}^{VF,n}_{GCRS}$. The increment in the AVU attitude over the time interval $[t_{E\,n\text{-}1},\ t_{En}]$ is $\Delta\mathbf{S}^{avy}_n = (\mathbf{S}^{VF,n-1}_{GCRS})^\mathsf{T}\mathbf{S}^{VF,n}_{GCRS}$. The increment in the attitude of the IMU is calculated by integrating the measurements $\boldsymbol{\omega}(t)$, $t \in [t_{En-1}, t_{En}]$, over the same interval: $\Delta\mathbf{S}^{imu}_n = \mathbf{I}_3 + \int_{t_{En-1}}^{t_{En}} \Delta\mathbf{S}^{imu}_n(t)\,\boldsymbol{\omega}(t)\,dt$.

To find the calibrated values of the angles $\boldsymbol{\beta}$, the matrix $\mathbf{C}^{VF}_{MF}(\boldsymbol{\beta})$ of coordinate transformation from VF to MF is constructed. This matrix allows one to write down an equation for finding $\boldsymbol{\beta}$, which relates the increments $\Delta\mathbf{S}^{avy}_n$ and $\Delta\mathbf{S}^{imu}_n$:

$$\mathbf{C}^{VF}_{MF}(\boldsymbol{\beta})\Delta\mathbf{S}^{avy}_n = \Delta\mathbf{S}^{imu}_n\mathbf{C}^{VF}_{MF}(\boldsymbol{\beta})\ ,\ n = 2...N\ . \qquad (1)$$

The solution to this equation is ambiguous. The matrix $\Delta\mathbf{S}_n^{avy}$ describes the rotation of the AVU around its eigenvector $\mathbf{k}_{VF,n}$: $\Delta\mathbf{S}_n^{avy}\mathbf{k}_{VF,n} = \mathbf{k}_{VF,n}$. Therefore, if the matrix $\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta})$ satisfies this equation, another matrix $\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta})\mathbf{Q}$ will also satisfy this equation, where $\mathbf{Q}$ is the matrix of rotation by an arbitrary angle around the vector $\mathbf{k}_{VF,n}$. Therefore, the attitude of the AVU relative to the IMU cannot be calibrated using observations recorded by the AINS stationary relative to the Earth. The calibration trajectory of the AINS (Fig. 2.b.) must contain at least two rotations with different vectors $\mathbf{k}_{VF,n} \neq \mathbf{k}_{VF,q}$, where n, q are the numbers of different AVU frames (Fig. 4.a).

Repeating equation (1) for all frames $n = 2\ldots N$ allows us to write an overdetermined system of equations for finding $\boldsymbol{\beta}$:

$$\begin{cases} \mathbf{I}_3 = (\Delta\mathbf{S}_2^{avy})^{\mathsf{T}}(\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta}))^{\mathsf{T}}\Delta\mathbf{S}_2^{imu}\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta}) \\ \vdots \\ \mathbf{I}_3 = (\Delta\mathbf{S}_N^{avy})^{\mathsf{T}}(\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta}))^{\mathsf{T}}\Delta\mathbf{S}_N^{imu}\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta}) \end{cases},$$

where $\mathbf{I}_N$ is the identity matrix of size $N\times N$.

The residual of each equation of this system is a matrix

$$\begin{bmatrix} f_{11n} & f_{12n} & f_{13n} \\ f_{21n} & f_{22n} & f_{23n} \\ f_{31n} & f_{32n} & f_{33n} \end{bmatrix} = (\Delta\mathbf{S}_N^{avy})^{\mathsf{T}}(\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta}))^{\mathsf{T}}\Delta\mathbf{S}_N^{imu}\mathbf{C}_{MF}^{VF}(\boldsymbol{\beta}) - \mathbf{I}_3 .$$

From the matrix elements for each value $n = 2\ldots N$, the residual vector is constructed in the form of a small rotation vector $\mathbf{f}_n(\boldsymbol{\beta}) = [f_{n,x}\ f_{n,y}\ f_{n,z}]^{\mathsf{T}} = 0{,}5[f_{32n} - f_{23n}\ \ f_{13n} - f_{31n}\ \ f_{21n} - f_{12n}]^{\mathsf{T}}$. The entire residual vector for finding the values of the 3rd group of calibrated parameters is assembled from the vectors $\mathbf{f}_n(\beta)$: $\mathbf{e}_{\{3\}}(\boldsymbol{\beta}) = [\mathbf{f}_2^{\mathsf{T}}(\boldsymbol{\beta})\,\mathbf{f}_3^{\mathsf{T}}(\boldsymbol{\beta})\ldots\mathbf{f}_N^{\mathsf{T}}(\boldsymbol{\beta})]^{\mathsf{T}}$. The desired value of the vector of calibrated parameters is found by minimizing the modulus of the residual vector $\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}}(\mathbf{e}_{\{3\}}^{\mathsf{T}}(\boldsymbol{\beta})\mathbf{e}_{\{3\}}(\boldsymbol{\beta}))^{1/2}$.

The values of the rotation angles φ_n^{avu} , φ_n^{imu} of the AVU and IMU contained in the increments $\Delta\mathbf{S}_n^{avy}$ and $\Delta\mathbf{S}_n^{imu}$ must be equal to each other:

$$\varphi_n^X = \arccos\frac{\mathrm{Tr}(\Delta\mathbf{S}_n^X) - 1}{2},\ X = avu, imu .$$

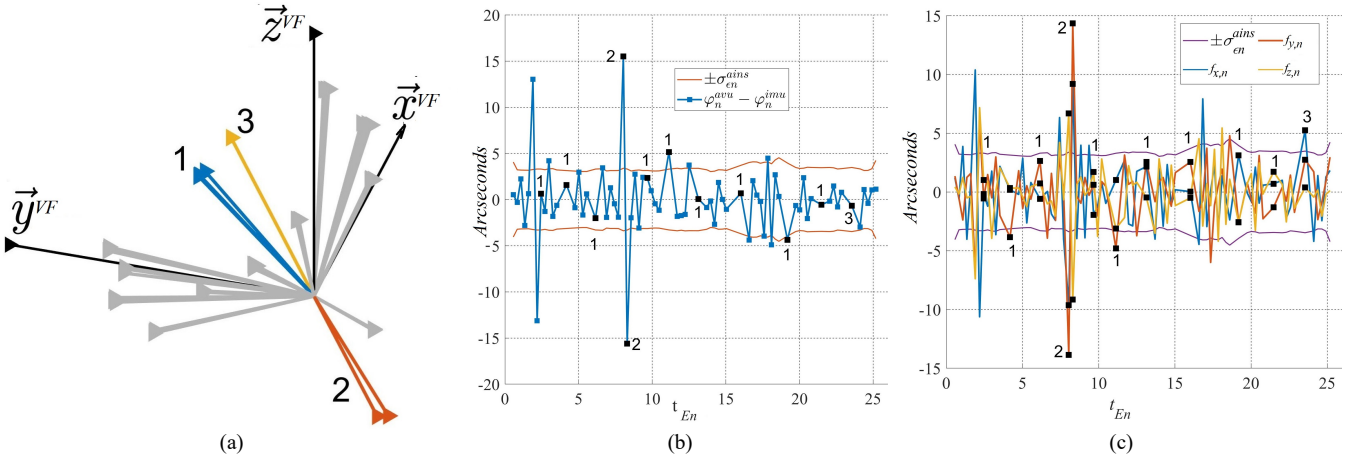
However, the real difference $\varphi_n^{avu} - \varphi_n^{imu} \neq 0$ is caused by measurement errors, refraction accounting errors and deformations of the AINS structure (Fig. 4.b.). This difference can be calculated before calibration, immediately after receiving the attitude increments. The difference limits the residual discrepancy after calibration – the components of the residual discrepancy vector will be distributed within the difference in rotation angles (Fig. 4.c.).

The calibration results are the vector $\hat{\boldsymbol{\beta}}$ and its covariance matrix $\mathbf{P}_{\delta\boldsymbol{\beta}}$. As a result of calibration, it is possible to obtain the root-mean-square errors in determining the angles of the AVU attitude relative to the IMU $\sigma_{\beta\psi} = 6.0''$, $\sigma_{\beta\vartheta} = 0.6''$, $\sigma_{\beta\gamma} = 0.6''$ with the angular pixel size of $13''$. As in the case of a real camera, here the rotation angle around the optical axis (angle β_ψ) was calibrated with a greater error than the tilt angles of the optical axis (angles β_ϑ and β_γ).

IV. Conclusions

Calibration of the strapdown AINS is a mandatory technological procedure that must be performed after manufacturing of each instance. Astronomical calibration of the AINS does not require complex testbed. A simple mount capable to hold the mass of the AINS and fixing its attitude is sufficient.

Astronomical calibration allows obtaining numerical



(a) (b) (c)

Fig.4. Calibration of the 3rd group of AINS parameters; t_{En} in time minutes; the numbers of arrows and values on the graphs correspond to the numbers of turns in Fig. 2.b.. (a) Unit vectors $\mathbf{k}_{VF,n}$ for different n values; colored arrows – increments in the attitude of the AINS due to the rotation of the mount, gray arrows are increments in the attitude of the AINS due to the rotation of the Earth; $\vec{x}^{VF}$, $\vec{y}^{VF}$, $\vec{z}^{VF}$ are VF coordinate axes. (b); $\sigma_{\varepsilon n}^{ains}$, arc seconds is the standard deviation of the difference $\varphi_n^{avu} - \varphi_n^{imu}$, calculated according to the specification of errors of the AVU and IMU. (c) Graphs of the components of the residual vectors $\mathbf{f}_n(\hat{\boldsymbol{\beta}})$.

values of all geometric parameters of the AINS that affect the error in determining its geographic coordinates (latitude and longitude) based on star observations. These parameters are divided into three groups that are calibrated sequentially, based on the same set of astronomical observations. That is, for complete calibration of the AINS using the astronomical method, one session of astronomical observations is sufficient.

The first group of parameters consists of intrinsic elements of each digital camera separately: focal length, coordinates of the principial point, radial distortion coefficients. These elements set the law of the central projection of the objects space onto the image plane.

The second group of parameters determines the relative attitude of individual digital cameras in the AVU. Using these parameters, the results of astronomical observations performed by several physical cameras are computationally combined and presented as the results of observations by one virtual camera. The size of the angular field of view of the virtual camera exceeds the size of the fields of view of real cameras. Due to this, the error in determining the attitude of the AVU relative to the stars is reduced.

The third group of parameters defines the constant attitude of the AVU relative to the IMU. This attitude is necessary for projecting the local vertical direction obtained from the IMU sensors measurements onto the VF. As a result of joint processing of this direction and the AVU attitude relative to the stars, the desired geographic coordinates of the AINS are obtained.

REFERENCES

[1] N.N. Vasilyuk, G.A. Nefedov, E.A. Sidorova, N.O. Shagimuratova, “Calibration of the intrinsic parameters of the digital camera of a star tracker based on ground-based observations of stars, taking atmospheric refraction and aberration of light into account,” Measurement Techniques, vol.66, pp.593–609, November 2023. https://doi.org/10.1007/s11018-023-02272-z.

[2] N.N. Vasilyuk, G.A. Nefedov, E.A. Sidorova, N.O. Shagimuratova, “Astronomical Calibration of a Strapdown Astroinertial Navigation System: Part 1. Calibration of Mutual Attitude of Digital Cameras,” Gyroscopy and Navigation, vol.15, no.2, pp.147–158, June 2024. https://doi.org/10.1134/S2075108724700251.

[3] N.N. Vasilyuk, G.A. Nefedov, E.A. Sidorova, N.O. Shagimuratova, “Astronomical Calibration of a Strapdown Astroinertial Navigation System. Part II. Calibration of Mutual Attitude of Inertial and Astronomical Sensors,” Gyroscopy and Navigation, vol.15, no.3, pp 256–268, September 2024. https://doi.org/10.1134/S2075108724700354

The Principle of Building a Tightly Coupled Inertial Optical Navigation System

V.A. Pogorelov
Aircraft engineering
Don State Technical University
Rostov-on-Don, Russia
vadim.pogorelov.rnd@gmail.com

S. V. Sokolov
Computer technology and information security
Rostov State University of Economics,
Rostov-on-Don, Russia
s.v.s.888@yandex.ru

Abstract — **One of the most highly accurate methods for solving the problem of autonomous navigation is processing the optical flow captured during the movement of an object. The existing optical flow processing methods based on the determination of the velocity field make it possible to estimate only projections of the linear and angular velocities of an object. In this regard, it is proposed to integrate an optical navigation system (ONS) with strapdown inertial navigation systems (SINS). The advantage of this approach is the autonomous evaluation of all navigation parameters that is stable over an arbitrary time interval. Due to the unavoidable interference of various physical nature, the synthesis of the considered integrated inertial-optical navigation system was carried out using the methods of the modern theory of nonlinear stochastic filtering. An extended Kalman filter, modified to take into account the correlation of object and observer noise, was chosen as an algorithm for estimating the vector of navigation parameters.**

Keywords*: optical flow, strapdown inertial navigation systems, velocity field, inertial optical navigation system, extended Kalman filter*

I. INTRODUCTION

In the last decade, unmanned aerial vehicles (UAV) have developed significantly in the military and civilian fields. The main areas of their use include the tasks of monitoring extended facilities such as oil and gas pipelines, industrial enterprises, etc.; cargo delivery, irrigation of agricultural land and pest control, search for people in remote areas and their rescue, exploration, targeting, etc. To successfully solve the tasks listed above, modern UAV are equipped with automatic control systems (ACS) that ensure their movement along a given route without operator involvement. The central problem of creating such ACS is the synthesis of their navigation systems (NS). Modern NS for advanced UAV must meet a number of requirements.

Firstly, they must ensure the required accuracy of movement along a given route, secondly, they must be resistant to interference of various physical nature, and thirdly, they must have low mass and low energy consumption. Also, the cost of NS is very relevant for small UAV produced in large batches.

To date, it is not possible to ensure full compliance with all the listed requirements within the framework of one type of NS. The main way to resolve this contradiction is to integrate NS based on various physical principles. The most common option for such integration SINS and global navigation satellite systems (GNSS) [1-4]. Despite the fact that various integration schemes have been developed and are widely used (separate, loosely coupled, tightly coupled and deeply integrated), the problem of ensuring the required navigation accuracy of UAV remains unresolved to the end. Moreover, the problem of protecting GNSS is significantly aggravated if they are exposed to spoofing interference. Protection of the receiver of signals from GNSS from distracting interference has significantly increased in recent years due to the development of a number of new methods for creating intelligent interference.

An analysis of the literature allows us to identify the following main ways of intellectual substitution of true signals of GNSS [6-9]:

The emission of rangefinder codes similar to the signals of GNSS, leading to incorrect estimates of the consumer's location and time;

- emission of signals from non-existent satellites;
- overlap of the visible satellite signal with noise from the false satellite signal;
- jamming, the signal-code structure of which coincides with the structure of the true signal of GNSS.

A natural alternative to integrating signals from SINS-GNSS, is to create a fully autonomous NS. The errors of such a NS do not depend on the time of its operation and the effects of external disturbing influences. Terrain Referenced Navigation (TRN) are currently the most widely used [10-13]. However, their use in UAV inevitably faces a number of challenges.

If TRN is used, the main problem is the impossibility of its application during the whole flight. This is due to the necessity of analysis of standard maps of large dimension, and so the inevitable long search of the extremum of the chosen cost function; as a result, the obtained coordinates of the mobile object become out of date. Besides, if the object moves over extended tundra, steppe, and deserted lands with regions of relatively flat surface, and over sea and ocean areas, it is characteristics. Moreover, if TRN is used, it is necessary for the object to keep constant angular orientation during the whole period of taking external navigation data. Obviously, during the flight, due to disturbances of various nature, it is not possible to accurately maintain the specified spatial orientation of the object. It is also necessary to take into account the high cost of TRN, which makes their use in small UAV impractical.

An alternative direction for the creation of fully autonomous high-precision NS, which has been applied in practice, is the integration of SINS with vision systems. This type of NS is called "optical navigation systems" (ONS) [14-19].

ONS installed on UAV are divided into three types: using a reference map, creating a map while moving, and not using a map. The latter include systems that implement visual odometry algorithms. A special but promising case of using visual odometry algorithms is the use of ONS that use information about changes in the intensity of the optical flow recorded during the movement of UAV.

Currently, the application of these systems faces two problems. The first is the high computational cost of determining the parameters of the velocity field, which carries information about the linear and angular velocities of UAV.

The work was carried out in accordance with the State Task №1023080200012-3-2.3.4.

The second is the impossibility of positioning UAV and determining its angular orientation based on the found velocity field, which is noisy with unavoidable measurement interference with unknown probabilistic characteristics.

Thus, the development of a new method for integrating SINS and ONS that takes into account the effects of external and internal measurement interference and does not impose restrictions on the movement of UAV is an urgent task.

II. STOCHASTIC EQUATIONS OF THE STATE VECTOR OF THE NAVIGATION SYSTEM IN THE FORM OF AN "OBJECT"

To solve this problem, we introduce the following coordinate systems (CS):

- $O\xi\eta\zeta$ – the Earth-centered earth-fixed (ECEF) CS I, the origin of which coincides with the center of the Earth, the axis $O\eta$ is directed along the axis of rotation of the Earth; $O\zeta$ coincides at the initial moment of time with the line of intersection of the plane of the equator and the Greenwich meridian, $O\xi$ complements the CS to the right;
- the instrument CS (ICS) J $Oxyz$ with the origin in the center of mass (CM) of the object, the axes of which are directed along the mutually orthogonal measuring axes of the sensitive elements (CE) of SINS;
- the accompanying geocentric CS (ACS) S $OXYZ$ with the origin in the CM of the object, the OY axis of which lies in the plane of the local meridian, the OZ axis is directed from the center of the Earth along the geocentric vertical, and the OX axis complements the coordinate system to the right.

The measuring complex of the SINS includes three orthogonal accelerometers and three orthogonal angular velocity sensors (AVS), the sensitivity axes of which are directed along the corresponding axes of the ICS.

As a mathematical model of a SINS, we use the most general nonlinear form of equations of a SINS [22]:

$$\dot{\mathbf{Y}}=\mathbf{F}(\mathbf{Y},t)+\mathbf{F_I}(\mathbf{Y},t)\mathbf{W_Y}, \qquad (1)$$

where $\mathbf{Y}=[\alpha\ \beta\ \gamma\ \lambda\ \varphi\ V_X\ V_Y\ V_Z\ h]^{\mathrm{T}}$, $\mathbf{Y}(0)=\mathbf{Y}_0$

$\mathbf{W_Y}=\left|\mathbf{W_d}^{\mathrm{T}}\ \ \mathbf{W_a}^{\mathrm{T}}\right|^{\mathrm{T}}$,

$$\mathbf{F}(\mathbf{Y},t)=\begin{bmatrix} \dfrac{\boldsymbol{\Phi}(\beta,\gamma)(\mathbf{Z_d}-\mathbf{m_d})}{\begin{bmatrix}0 & (\cos\varphi)^{-1}\\ -1 & 0\end{bmatrix}\begin{bmatrix}-V_Y\\ V_X\end{bmatrix}(r+h)^{-1}} \\ \mathbf{C}^{\mathrm{T}}(\alpha,\beta,\gamma,\lambda,\varphi)\mathbf{Z_a}+ \\ +\left(\left(2\begin{bmatrix}0\\ \Omega\cos\varphi\\ \Omega\sin\varphi\end{bmatrix}+(r+h)^{-1}\begin{bmatrix}-V_Y\\ V_X\\ V_X\mathrm{tg}\varphi\end{bmatrix}\right)\times\begin{bmatrix}V_X\\ V_Y\\ V_Z\end{bmatrix}\right)- \\ -\begin{bmatrix}0\\ -\Omega^2(r+h)\cos\varphi\sin\varphi\\ \Omega^2(r+h)\cos^2\varphi+g\end{bmatrix} \\ \hline V_Z \end{bmatrix},$$

$$\mathbf{F_I}(\mathbf{Y},t)=\begin{bmatrix}-\boldsymbol{\Phi}(\beta,\gamma) & 0\\ 0 & 0\\ 0 & -\mathbf{C}^{\mathrm{T}}(\alpha,\beta,\gamma,\lambda,\varphi)\\ 0 & 0\end{bmatrix},$$

$$\Phi(\beta,\gamma)=\begin{bmatrix}\sin\gamma/\cos\beta & \cos\gamma/\cos\beta & 0\\ \cos\gamma & -\sin\gamma & 0\\ \sin\gamma\,\mathrm{tg}\beta & \cos\gamma\,\mathrm{tg}\beta & 1\end{bmatrix},$$

where α,β,γ - Euler-Krylov angles, which determine the orientation of the ICS relative to the ECEF, $\mathbf{Z_d}=[Z_x\ \ Z_y\ \ Z_z]^{\mathrm{T}}$ - measurement vector of three AVS, $\mathbf{W_d}=[W_x\ W_y\ W_z]^{\mathrm{T}}$ – the additive interference vector of the AVS measurement, further approximated by white Gaussian noise (WGN) with zero mathematical expectation and an intensity matrix $\mathbf{D_d}$, $\mathbf{m_d}=[m_{dx}\ \ m_{dy}\ \ m_{dz}]^{\mathrm{T}}$ - vector of constant zero displacement AVS, $\mathbf{V_S}=[V_X\ \ V_Y\ \ V_Z]^{\mathrm{T}}$ – the vector of the object's velocity relative to the Earth, r – the radius of the Earth, h – object height, g – gravitational acceleration, $\mathbf{Z_a}=[Z_{ax}\ \ Z_{ay}\ \ Z_{az}]^{\mathrm{T}}$ - measurement vector of three accelerometers, $\mathbf{W_a}=[W_{a_x}\ \ W_{a_y}\ \ W_{a_z}]^{\mathrm{T}}$ – additive interference vector of accelerometer measurements approximated by WGN with zero mean and intensity matrix $\boldsymbol{D_a}$, $\mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi)=\mathbf{D}(\alpha,\beta,\gamma)\mathbf{B}^{\mathrm{T}}(\lambda,\varphi)$ – the rotation matrix ICS relative to ACS, $\mathbf{D}(\alpha,\beta,\gamma)$ – the rotation matrix ICS with respect to ECEF K, $\mathbf{B}(\lambda,\varphi)$ – the rotation matrix ACS relative to ECEF:

$$\mathbf{B}(\lambda,\varphi)=\begin{bmatrix}\cos(\lambda+\Omega t) & 0 & -\sin(\lambda+\Omega t)\\ -\sin(\lambda+\Omega t)\sin\varphi & \cos\varphi & -\cos(\lambda+\Omega t)\sin\varphi\\ \sin(\lambda+\Omega t)\cos\varphi & \sin\varphi & \cos(\lambda+\Omega t)\cos\varphi\end{bmatrix}.$$

Model (1) fully satisfies the possibilities of further solving the problem of stochastic estimation of the vector of navigation parameters of UAV. However, it is necessary to take into account the fact that measurements of all the parameters of the SINS have already been used in the synthesis of stochastic equations of the vector of state of the SINS (1). At the same time, the use of modern methods of nonlinear stochastic filtering also involves the construction of an observer of the state vector of a SINS - measurements containing information about the estimated variables of the model (1). In this regard, we will further consider the possibility of using OSN readings for its synthesis.

III. INFORMATION MODEL OF AN OPTICAL NAVIGATION SYSTEM

The estimation of the motion parameters of UAV in modern ONS is based on the processing by various methods of information contained in changes in the intensity, contrast or color gamut of light reflected from the scanned surface [13-18].

One of the main approaches to estimating the navigation parameters of an object based on the time sequence of images is the use of methods based on calculating the field of optical flow velocities [17].

These methods allow us to estimate the motion parameters based on the movement of the brightness pattern, which consists in a change in the illumination of the earth's surface. In this case, it is not necessary to establish correspondences between features in consecutive images. The essence of these methods is that the movement of the "scene" relative to the observer generates a velocity distribution on the picture plane (*x, y*) ("velocity field").

This is expressed in the time variation of brightness values on the image plane. In this case, the velocities of the image points are functions of the parameters of the linear and angular motion of the object.

To build a measurement model for the ONS, we additionally introduce the *OX*Y*Z** CS associated with the video camera (VCS)

In the VCS, the equations relating the optical flow parameters u, v to the projections of the linear and angular velocities of an object have the form [17]:

$$\begin{bmatrix} u \\ v \end{bmatrix} = f / Z^* \begin{bmatrix} -1 & 0 & x^* \\ 0 & -1 & y^* \end{bmatrix} \begin{bmatrix} V_{X^*} \\ V_{Y^*} \\ V_{Z^*} \end{bmatrix} + f \begin{bmatrix} y^* x^* & -(x^{*2}+1) & y^* \\ (y^{*2}+1) & -y^* x^* & -x^* \end{bmatrix} \begin{bmatrix} \omega_{X^*} \\ \omega_{Y^*} \\ \omega_{Z^*} \end{bmatrix} \quad (2)$$

where $x^* = fX^*/Z^*, y^* = fY^*/Z^*$ – coordinates of the projection of the scan point P onto the image plane, f - focal length, z^* - image depth (projections of the scan point P), $V_{X^*}, V_{Y^*} V_{Z^*}$ – projections of the linear velocity vector CM of an object on the VCS axis, $\omega_{X^*}, \omega_{Y^*} \omega_{Z^*}$ – projections of the angular velocity vector of the object on the VCS axis, the OZ* axis coincides with the line of sight of the video camera.

The use of an ONS involves the removal of optical information from the Earth's surface, while the beginning of the VCS coincides with the beginning of the ICS, at the initial moment of time the axis *OX** VCS coincides with the direction of the axis *Oy* of the ICS, the axis *OY** VCS - with the direction of the axis Ox of the ICS and the axis *OZ** VCS is directed opposite to the axis *Oz* of the ICS.

This allows us to write down the equations of the optical flow parameters in projections on the axis of the ICS in the following form:

$$\begin{bmatrix} u \\ v \end{bmatrix} = f / Z \begin{bmatrix} 0 & 1 & -y \\ 1 & 0 & -x \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} + f \begin{bmatrix} -(y^2+1) & yx & x \\ -yx & x^2+1 & -y \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}, \quad (3)$$

where x, y – coordinates of the projection of the scan point P onto the image plane P на плоскость изображения, $Z = Z(x,y)$ - scan point depth P, V_x, V_y, V_z – projections of the linear velocity vector CM of an object on the ICS axis,, $\omega_x, \omega_y \omega_z$ – projections of the vector of the absolute angular velocity of an object on the axis ICS..

For the possibility of forming an observer of the vector Y based on equations (3), we express the variables included in them in ACS.

In the absence of rotational motion of the video camera when scanning the Earth's surface, the image depth of the scanning point P is defined as $Z(x,y) = \sqrt{x^2 + y^2 + h^2}$, therefore, when rotating it, the depth of the image is equal to:

$Z(x,y) = \sqrt{x^2 + y^2 + h^2 / c_{33}^2(\alpha,\beta,\gamma,\lambda,\varphi)}$, where $C_{33}(\alpha,\beta,\gamma,\lambda,\varphi)$ - the corresponding element of the matrix $C(\alpha,\beta,\gamma,\lambda,\varphi)$. Because linear and angular velocity vectors in ACS can be represented as:

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = \mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi) \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix}, \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \mathbf{Z_d - m_d - W_d},$$

that is, finally, the equations of the optical flow velocity field, expressed in the estimated navigation parameters, take the form:

$$\begin{bmatrix} u \\ v \end{bmatrix} = f\left(x^2 + y^2 + h^2 / c_{33}^2(\alpha,\beta,\gamma,\lambda,\varphi)\right)^{-\frac{1}{2}} \times$$
$$\times \begin{bmatrix} 0 & 1 & -y \\ 1 & 0 & -x \end{bmatrix} \mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi) \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix} +$$
$$+ f \begin{bmatrix} -(y^2+1) & yx & x \\ -yx & x^2+1 & -y \end{bmatrix} (\mathbf{Z_d - m_d - W_d}). \quad (4)$$

Due to the fact that calculating the optical flux vector using the basic optical flux equation is a complex computational task – unlike measuring the intensity of illumination $I(x, y, t)$ and calculating its partial derivatives $\partial I/\partial t$, $\partial I/\partial x$, $\partial I/\partial y$, we will form the observer vector of the object's navigation parameters as follows.

We transform the optical flow equation to the form

$$\frac{\partial I}{\partial t} = -\begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

and then consider the current value of the rate of change of the measured illumination intensity as the observer's output signal $I(x,y,t)$ in time $\partial I/\partial t$, considering the obvious possibility of determining partial derivatives in real time $\partial I/\partial t$, $\partial I/\partial x$, $\partial I/\partial y$ in the onboard computer. In this case, the output signal of the observer $\boldsymbol{Z_{It^*}}$ of the vector of navigation parameters $\boldsymbol{Y}$, taking into account (4), takes the form:

$$Z_{It^*} = \frac{\partial I}{\partial t} = -f \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} \left[\left(x^2 + y^2 + h^2 / c_{33}^2(\alpha,\beta,\gamma,\lambda,\varphi)\right)^{-\frac{1}{2}} \begin{bmatrix} 0 & 1 & -y \\ 1 & 0 & -x \end{bmatrix} \mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi) \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix} + \right.$$
$$\left. + \begin{bmatrix} -(y^2+1) & yx & x \\ -yx & x^2+1 & -y \end{bmatrix} (\mathbf{Z_d - m_d - W_d}) \right]. \quad (5)$$

Excluding the deterministic component from the right-hand side of (5), we finally obtain the information model of the observer of the state vector ***Y:***

$$Z_{It} = -f \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} \left[\left(x^2 + y^2 + h^2 / c_{33}^2(\alpha,\beta,\gamma,\lambda,\varphi)\right)^{-\frac{1}{2}} \begin{bmatrix} 0 & 1 & -y \\ 1 & 0 & -x \end{bmatrix} \mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi) \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix} - \right.$$
$$\left. - \begin{bmatrix} -(y^2+1) & yx & x \\ -yx & x^2+1 & -y \end{bmatrix} \mathbf{W_d} \right]. \quad (6)$$

Very often, the probabilistic interference characteristics of $\boldsymbol{W_{It}}$ optical flux intensity measurements differ from Gaussian

ones, which makes it impossible to apply the well-developed apparatus of the modern theory of nonlinear filtering. In such cases, we use the following technique for the possibility of using this device. Let's average the observation signals $\boldsymbol{Z_{It_i}}$ (7) obtained at all N scan points *(xi, yi), i=1,..,N:*

$$Z_{\Sigma}=\frac{1}{N}\sum_{i=1}^{N}Z_{It_i}=\frac{f}{N}\sum_{i=1}^{N}\left[\frac{\partial I}{\partial x}\ \frac{\partial I}{\partial y}\right]_{\substack{x=x_i\\y=y_i}}\left\{\left(x_i^2+y_i^2+h^2/c_{33}^2(\alpha,\beta,\gamma,\lambda,\varphi)\right)^{-\frac{1}{2}}\begin{bmatrix}0&1&-y_i\\1&0&-x_i\end{bmatrix}\mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi)\begin{bmatrix}V_X\\V_Y\\V_Z\end{bmatrix}\right\}+ \tag{7}$$

$$+\frac{f}{N}\sum_{i=1}^{N}\left[\frac{\partial I}{\partial x}\ \frac{\partial I}{\partial y}\right]_{\substack{x=x_i\\y=y_i}}\begin{bmatrix}-(y_i^2+1)&y_ix_i&x_i\\-y_ix_i&x_i^2+1&-y_i\end{bmatrix}\mathbf{W_d}+\frac{1}{N}\sum_{i=1}^{N}W_{It_i}.$$

Since the interference of light intensity measurements is equally powerful for each scanning point, by virtue of the central limit theorem, the interference distribution

$$W_{It^*}=\frac{1}{N}\sum_{i=1}^{N}W_{It_i}$$

in the new observer (7) already at $N\geq 3$ it will be close to the Gaussian, constantly approaching it with increasing N .This, in turn, makes it possible to approximate the interference of $\boldsymbol{W_{It^*}}$ centered WGN with the intensity of $\boldsymbol{D_{It}*}$. As a result, the desired observer of the navigation parameters of a moving object, which provides the possibility of using the developed nonlinear filtering methods, has the form:

$$\mathbf{Z_\Sigma}=\mathbf{Q_1}(\alpha,\beta,\gamma,\lambda,\varphi,h,t)\mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi)\begin{bmatrix}V_X\\V_Y\\V_Z\end{bmatrix}+\begin{bmatrix}\mathbf{Q_2(t)}&1\end{bmatrix}\begin{bmatrix}\mathbf{W_d}\\\mathbf{W_{It^*}}\end{bmatrix},$$

$$\mathbf{Q_1}(\alpha,\beta,\gamma,\lambda,\varphi,h,t)=-\frac{f}{N}\sum_{i=1}^{N}\left[\frac{\partial I}{\partial x}\ \frac{\partial I}{\partial y}\right]_{\substack{x=x_i\\y=y_i}}\left(x_i^2+y_i^2+h^2/c_{33}^2(\alpha,\beta,\gamma,\lambda,\varphi)\right)^{-\frac{1}{2}}\begin{bmatrix}0&1&-y_i\\1&0&-x_i\end{bmatrix},$$

$$\mathbf{Q_2(t)}=\frac{f}{N}\sum_{i=1}^{N}\left[\frac{\partial I}{\partial x}\ \frac{\partial I}{\partial y}\right]_{\substack{x=x_i\\y=y_i}}\begin{bmatrix}-(y_i^2+1)&y_ix_i&x_i\\-y_ix_i&x_i^2+1&-y_i\end{bmatrix}$$

or

$$\mathbf{Z_\Sigma}=\mathbf{H}(\mathbf{Y},t)+\mathbf{H_0}(t)\mathbf{W_\Sigma},\mathbf{W_\Sigma}=\begin{bmatrix}\mathbf{W_d}^{\mathrm{T}}&\mathbf{W_{It^*}}\end{bmatrix}^{\mathrm{T}},$$

where

$$\mathbf{H}(\mathbf{Y},t)=\mathbf{Q_1}(\alpha,\beta,\gamma,\lambda,\varphi,h,t)\mathbf{C}(\alpha,\beta,\gamma,\lambda,\varphi)\begin{bmatrix}V_X\\V_Y\\V_Z\end{bmatrix},(8)$$

$$\mathbf{H_0}(t)=\begin{bmatrix}\mathbf{Q_2}(t)&1\end{bmatrix}.$$

The predominant feature of the observer (8) is the obvious dependence of its information part on all components of the estimated vector Y, which positively affects the convergence and accuracy of the filtering process.

IV. Stochastic Estimation of Navigation Parameters in an Inertial Optical Navigation System

Equations (1) and (8) obtained in the standard "observer object" form provide the possibility of solving the problem of stochastic estimation of the vector Y based on an extended Kalman filter with correlated object and observer noise [19-20]. Since in the case of correlated noise, when minimizing the a posteriori covariance matrix, the filtering coefficient is defined as

$$\mathbf{K}(\hat{\mathbf{Y}},t)=\left(\mathbf{R}(\hat{\mathbf{Y}},t)\frac{\partial\mathbf{H}^{\mathrm{T}}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}+\mathbf{S}(\hat{\mathbf{Y}},t)\right)\mathbf{D}_{\Sigma}^{-1}(\hat{\mathbf{Y}},t),$$

where correlation is taken into account by introducing a matrix S that depends on the parameters of the object and observer, and the characteristics of their noise:

$$\mathbf{S}(\hat{\mathbf{Y}},t)=\mathbf{F}_1(\hat{\mathbf{Y}},t)\begin{bmatrix}\mathbf{D_d}\vdots\mathbf{0}\\\cdots\cdots\cdots\\\mathbf{0}\end{bmatrix}\mathbf{H}_0^{\mathrm{T}}(t),$$

then the final filter you are looking for takes the form:

$$\dot{\hat{\mathbf{Y}}}=\mathbf{F}(\hat{\mathbf{Y}},t)+\mathbf{K}(\hat{\mathbf{Y}},t)(\mathbf{Z_\Sigma}-\mathbf{H}(\hat{\mathbf{Y}},t)),$$

$$\dot{\mathbf{R}}(\hat{\mathbf{Y}},t)=\left(\frac{\partial\mathbf{F}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}-\mathbf{S}(\hat{\mathbf{Y}},t)\mathbf{D}_{\Sigma}^{-1}(\hat{\mathbf{Y}},t)\frac{\partial\mathbf{H}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}\right)\mathbf{R}(\hat{\mathbf{Y}},t)+$$

$$+\mathbf{R}(\hat{\mathbf{Y}},t)\left(\frac{\partial\mathbf{F}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}-\mathbf{S}(\hat{\mathbf{Y}},t)\mathbf{D}_{\Sigma}^{-1}(\hat{\mathbf{Y}},t)\frac{\partial\mathbf{H}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}\right)^{\mathrm{T}}-$$

$$-\mathbf{R}(\hat{\mathbf{Y}},t)\frac{\partial\mathbf{H}^{\mathrm{T}}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}\mathbf{D}_{\Sigma}^{-1}(\hat{\mathbf{Y}},t)\frac{\partial\mathbf{H}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}\mathbf{R}(\hat{\mathbf{Y}},t)+$$

$$+\mathbf{F}_1(\hat{\mathbf{Y}},t)\mathbf{D}_{\mathbf{Y}}\mathbf{F}_1^{\mathrm{T}}(\hat{\mathbf{Y}},t)-\mathbf{S}(\hat{\mathbf{Y}},t)\mathbf{D}_{\Sigma}^{-1}(\hat{\mathbf{Y}},t)\mathbf{S}^{\mathrm{T}}(\hat{\mathbf{Y}},t),$$

where $\hat{\mathbf{Y}}$ - estimation of the vector of navigation parameters $\mathbf{Y}(t)$, $\mathbf{R}(\hat{\mathbf{Y}},t)$ - a posteriori covariance matrix,

$$\mathbf{D_Y}=\begin{vmatrix}\mathbf{D_d}&\mathbf{0}\\\mathbf{0}&\mathbf{D_a}\end{vmatrix},$$

$$\mathbf{D}_{\Sigma}(\hat{\mathbf{Y}},t)=\mathbf{H_0}(t)\begin{vmatrix}\mathbf{D_d}&\mathbf{0}\\\mathbf{0}&\mathbf{D_{It^*}}\end{vmatrix}\mathbf{H}_0^{\mathrm{T}}(t)$$

As a rule, the use of the extended Kalman filter (9) provides a compromise on the criterion of "accuracy - computational cost" when implemented in general and special purpose computers.

V. Conclusion

A comparison of the results obtained with the known ones [14-19] allows us to conclude that using an ONS as an observer of the navigation parameters of an object when integrating a SINS is more effective than directly determining the linear and angular velocity projections from the parameters of the optical flow and then integrating them, both in terms of computational costs and by the accuracy of estimating the object's motion parameters. Other advantageous features of the considered inertial-optical NS are its autonomy and the ability to observe all parameters of the object's movement, which is extremely important to ensure the stability and accuracy of the process of estimating navigation parameters. The simplicity of the technical implementation of this system allows it to be used for navigation of UAV for various purposes.

References

[1] Krasilshchikov MN, Sebrryakov GG, eds, Modern information technique applying to maneuverable unmanned vehicles guidance, navigation and control problems. Moscow: Publ. Fizmatlit; 2009.

[2] I. N. Rozenberg, S.V. Sokolov, V.I. Umanskij and V.A. Pogorelov, “Theoretical Basis of Close Integration of Inertial Satellite Navigation Systems”, Moscow: Publ. Fizmatlit, 2018

[3] G. I. Emel'yantsev and A. P. Stepanov, Integrated Inertial-Satellite Orientation and Navigation Systems, Ed. By V. G. Peshekhonov, Elektropribor, St. Petersburg, 2016.

[4] G. Schmidt, R.Phillips, INS/GPS Integration Architectures, NATO RTO Lecture Series, RTO-EN-SET-116, Low-Cost Navigation Sensors and Integration Technology, Mar. 2011

[5] B. L. Stevens and F. L. Lewis, Aircraft Control and Simulation, John Wiley & Sons, Second edition, 2003.

[6] D. Shepard, J. Bhatti, T. Humphreys, "Drone Hack", GPS World Aug. 2012, pp. 30-33.

[7] Congressional Budget Office. The Global Positioning System for Military Users: Current Modernization Plans and Alternatives., Oct. 2011.

[8] A. R. Baziar, M. Moazedi, M. R Mosavi, Analysis of single frequency GPS receiver under delay and combining spoofing algorithm, Journal of Wireless Personal Communications, 2015, vol. 83, no. 3, pp. 1955–1970.

[9] S. Mahmood, Analysis of Differential Global Positioning System (DGPS) Techniques and GPS Jamming on Precision Guided Munition (PGM), Performance, NATO/AGARD MSP Meeting, Technologies for Precision Air Strike Operations in Rapid Reaction and Localized Conflict Scenarios, Seville, Spain, AGARD CP-576, 1996.

[10] Pritchett J.E., Pue A.J, Robust Guidance and Navigation for Airborne Vehicles Using GPS/ Terrain Aiding // Position Location and Navigation Symposium, IEEE 2000, San Diego, CA, USA, pp.457-463.

[11] Lee J., Chang-Ky Sung, Oh J., Han K., Lee S., Myeong-Jong Yu. A Pragmatic Approach to the Design of Advanced Precision Terrain-Aided Navigation for UAVs and Its Verification, Remote Sens. 2020, no. 12, p. 1396.

[12] Wang K, Zhu T, Wang J, Impact of terrain factors on the matching performance of terrain-aided navigation. NAVIGATION. 2019; no 66, pp 451–462.

[13] Pogorelov V.A., Multi-structural estimation of an integrated navigation system of a reusable spacecraft, Cosmic Research, 2008, no 46(3), pp.238–243.

[14] Mittal M., Mohan R., Burgard W., Valada A. Vision-Based Autonomous UAV Navigation and Landing for Urban Search and Rescue, Proceedings of the International Symposium on Robotics Research (ISRR), 2019.

[15] Geiger A., Lenz P., Stiller C., Urtasun R. Vision meets Robotics: The KITTI Dataset // International Journal of Robotics Research (IJRR), http://www.cvlibs.net/datasets/kitti/ (2013)

[16] Raudies F., Neumann H. A review and evaluation of methods estimating ego-motion // Computer Vision and Image Understanding, 2012, vol. 116, pp.606–633.

[17] Baker S., Scharstein D., Lewis J.P., Roth S., Black M.J., Szeliski R. A Database and Evaluation Methodology for Optical Flow, International Journal of Computer Vision, 2011, pp.1–31.

[18] Fleet D.J., Weiss Y. Optical Flow Estimation. Mathematical Models in Computer Vision: The Handbook, 2005, pp.239-258.

[19] Pogorelov, V. A., Sokolov S.V., Algorithmic support for integrated navigation systems, Journal of Computer and Systems Sciences International. 2008, vol. 47, no. 2. pp 308-320.

[20] S.V.Sokolov, V.A. Pogorelov, "Stochastic Evaluation, Control and Identification in High-precision Navigation Systems", Moscow: Publ. Fizmatlit, 2016.

[21] I.Rusnak, Optimal State Estimation of Nonlinear Dynamic Systems, Nonlinear Systems - Modeling, Estimation, and Stability, InTech, 2018, pp 159-178.

Evaluation of the Reliability of the Results of a Flight Experiment with an Inertial Navigation System

M.S. Selezneva
Department of Automatic Control Systems
Bauman Moscow State Technical University
Moscow, Russia
m.s.selezneva@mail.ru

K.A. Neusypin
Department of Automatic Control Systems
Bauman Moscow State Technical University
Moscow, Russia
neysipin@mail.ru

Abstract— The paper presents a methodology for assessing the reliability of the results of a flight experiment with a navigation complex. The navigation complex, which includes an inertial navigation system and a satellite navigation system receiver, is investigated. The navigation complex is installed on an airship. Joint signal processing is performed using linear and nonlinear adaptive Kalman filters. When using the reliability assessment methodology, it is proposed to use a nonlinear self-organizing model of inertial navigation system errors.

Keywords— navigation complex, airship, navigation system errors, Kalman filter, self-organization, assessment methodology.

I. INTRODUCTION

To determine the navigation parameters of the airship, navigation systems of various compositions are used. The basic system of the navigation system is the inertial navigation system. An inertial navigation system built on micro-electro-mechanical (MEMS) elements is often used [1, 2]. MEMS elements are low cost and are characterized by instability of characteristics. Therefore, compensation for errors in the inertial navigation system is key to improving the reliability and accuracy of the low-flying airship. When using an airship to conduct flight tests to determine the accuracy characteristics of the navigation system or individual navigation systems, it is necessary to accurately know the coordinates of the airship and the orientation angles in two modes: hovering and movement along a given route. For this purpose, an accurate inertial navigation system on fiber-optic gyroscopes and five GNSS signal receivers are installed on board the airship. The inertial navigation system and GNSS signal receivers are combined into a high-precision NC, which is the standard of navigation definitions [3, 4]. During the flight experiment, information was recorded from a rough inertial navigation system, which was subsequently processed using an adaptive estimation algorithm to determine the errors of the system for their subsequent compensation in the output signal of the inertial navigation system. An adaptive algorithm, which is a direct modification of the Kalman filter (KF) with a rigid connection according to an updated sequence [4, 5], was used as an estimation algorithm. To assess the reliability of the obtained results, a technique was used that involves comparing the exact information obtained from the inertial navigation system built on fiber-optic gyroscopes with parameter estimates obtained after processing the calculated values of drift velocities calculated by the estimation algorithm. Comparative analysis showed a fairly high accuracy of the adaptive estimation algorithm in the hovering mode of the airship. In the study of the accuracy characteristics in the methodology, a linear mathematical model of inertial navigation system errors was used.

To simulate the KF, data from a semi-naturalistic experiment with a real INS were used. When the INS is installed on a fixed base, the output information on the location and speed are INS errors. INS errors are a discrete analogue of white Gaussian noise. At the output of the KF, we obtain estimates of INS errors in determining the speed, the angles of deviation of the gyrostabilized platform (GSP) from the horizon plane and the horizontal drifts of the GSP.

INS errors in determining the location obtained during the laboratory experiment were compared with the calculated errors in determining the latitude and longitude, which were calculated using the KF.

Errors of a real INS installed on a fixed base are mainly due to GSP drifts, therefore, to calculate the INS errors in determining the path traveled, we use the well-known formula [6, 7]:

$$\delta x = -\varepsilon^* R_z(\mathrm{t} - \frac{\sin vt}{v}) \tag{1}$$

where δx – estimates of the northern components of errors in determining the INS; ε^* – constant drift of GSP; R_z – radius of the earth; $v^2 = g/R_z$.

The drift value estimates obtained by the KF must be averaged, since the formula was obtained under the condition of a constant GSP drift velocity. The obtained nature of the change in the horizontal GSP drifts over the interval of laboratory experiments allows averaging over three intervals: 0 - 100; 100 - 200; 200-325 min.

Then we get:

$$\delta x = \varepsilon^* R_z\left(\mathrm{t} - \frac{\sin vt}{v}\right) + \left(\varepsilon_2^* - \varepsilon_1^*\right)\left[\left(t - t_1\right) - \frac{\sin v\left(t - t_1\right)}{v}\right] + \tag{2}$$

$$+\left(\varepsilon_3^* - \varepsilon_2^*\right)\left[\left(t - t_2\right) - \frac{\sin v\left(t - t_2\right)}{v}\right]$$

The formula [7] is used to determine latitude:

$$\delta L = \frac{\delta x}{R_z} \tag{3}$$

where L – latitude of location.

The calculated values of the errors in latitude were compared with the errors of the real INS during the laboratory experiment. The values of the errors in longitude are calculated similarly.

During the flight experiment during the movement along the airship route, the assessment of the accuracy characteristics using the adopted method demonstrated low accuracy of the error estimates, which is due to the inadequacy of the used linear mathematical model and linear adaptive estimation algorithm. When estimating the errors of the inertial navigation system during the flight along the route, it is advisable to use a nonlinear adaptive Kalman filter, which uses a nonlinear model of errors of the inertial navigation system. The mathematical model of the inertial navigation system in the algorithm is built by the self-organization method, which ensures the adequacy of the model at each flight interval.

The accuracy of the navigation definitions of the aircraft is increased algorithmically. Usually, various estimation algorithms are used to correct the inertial navigation system, for example, a nonlinear Kalman filter. The nonlinear Kalman filter uses covariance matrices of input and measurement noise. The accuracy of estimation depends on the reliability of these matrices, so the study of noise and its impact on the accuracy of estimation algorithms is a pressing problem. There are approaches within which the implementation of the nonlinear Kalman filter is reduced to solving a stochastic partial differential equation written in the Ito form or in the Stratonovich form [8, 9]. However, such practical implementation is also difficult because when integrating these equations, it is necessary to apply special rules that do not coincide with the usual rules of mathematical analysis. The most complete consideration of all the features of the nature of the change in errors of the inertial navigation system and, what is especially important, of a specific inertial navigation system under the conditions of each specific flight is possible by constructing a nonlinear model using one of the evolutionary algorithms.

II. SELF-ORGANIZATION ALGORITHM

We assign a parameter vector to each basis function $(a, f)^{\mathrm{T}}$ and examine the two-dimensional vector , where a is the amplitude, f is the frequency. The parameterized set of basis functions has the form:

$$F_{\mathrm{n}} = \{a_{ii}(f_i x) \mid i = 1, \ldots, N\}. \tag{4}$$

The self-organization algorithm is based on the hypothesis of model selection using an ensemble of criteria. The resulting model has the form [10, 11]

$$M(x) = \sum_{i=1}^{N} a_i \mu_{ni}(f_i x), \tag{5}$$

where N – number of basis functions in the model; μ_{ni} – basis functions from F_p.

Accordingly, the data reliability assessment technique requires using constructed nonlinear self-organizing models at each calculation interval. An inertial navigation system corrected by GNSS signals is considered. A method for algorithmic correction of an integrated inertial navigation system using an adaptive nonlinear Kalman filter is investigated. In order to improve the accuracy of navigation determinations, full-scale tests of navigation equipment were conducted in order to more accurately determine the statistical characteristics used in the nonlinear Kalman filter. To assess the reliability of the conducted flight experiments, a technique was used using linear and nonlinear mathematical models of inertial navigation system errors, as well as an adaptive linear estimation algorithm, which is a direct modification of the Kalman filter and the nonlinear adaptive Kalman filter. In the reliability assessment technique, linear and nonlinear models of inertial navigation system errors are used, respectively. As a nonlinear mathematical model, models constructed by the self-organization method based on the measured data of a rough inertial navigation system are used.

Using the idea of selection in the development of mathematical algorithms, the hypothesis is accepted that ineffective combinations discarded in the first rows of self-selection could not produce optimal combinations of the next row if we had passed them further.

The model constructed using the self-organization algorithm has the form:

$$y_1(\delta V) = b_{01} + b_{11}x_1 + b_{21}x_3 + b_{31}x^2{}_1 + b_{41}x^2{}_3 + b_{51}x_1x_3$$
$$y_2(\varphi) = b_{02} + b_{12}x_1 + b_{22}x_4 + b_{32}x^2{}_1 + b_{42}x^2{}_4 + b_{52}x_1x_4$$
$$y_3(\varepsilon) = b_{03} + b_{13}x_3 + b_{23}x_4 + b_{33}x^2{}_3 + b_{43}x^2{}_4 + b_{54}x_3x_4$$

During the flight experiment, coefficients were obtained that are calculated using the Gauss method

$$b_{01} = 5,86; b_{11} = -3,27; b_{21} = 4,21;$$
$$b_{31} = 0,0013; b_{41} = -0,0804; b_{51} = 0,08.$$

The presented model is used in the methodology for assessing the accuracy of the INS.

The model is calculated anew at each stage of operation, since the type of model at the output of the self-organization algorithm depends on the specific measurement sample used.

Here errors in determining the speed (δV); GSP deviation angles (φ); drift velocity (ε).

The results of modeling based on the flight experiment data demonstrated the high efficiency of the developed methodology for assessing the accuracy of INS error models. The accuracy of navigation determinations when using self-organizing models in the algorithmic support of the NC increases: the airship speed increases by an average of 8%; the GSP deviation angle - 10-12%; the GSP drift velocity - 15%.

When conducting the modeling, it was taken into account that the nature of the process under study is approximately known. Therefore, as a reference function, which will be complicated at subsequent stages, we choose a harmonic (sinusoidal) function of the form:

$$Z = p\sin(\psi_k + wt).$$

At the preliminary selection stage, we determine the value of the amplitude and phase of the support function as follows:

- as the initial value of the amplitude, we select the average value of the measurements obtained by discarding the maximum and minimum measurements from 20 values of the training sequence;

Note that the initial data for the training and verification sequences will be taken in a 2:1 ratio. We assume that as the initial data we have 30 values obtained during the operation of the system in the correction mode from the satellite

system immediately before switching it to the autonomous mode.

- during the preliminary selection, the amplitude varies up to 14 values. In this case, the value changes by up to 100%;
- the phase varies from 0 to ±90%. As a result, we also obtain 14 values: 7 positive and 7 negative.

When checking the operability and quality of the developed algorithms, the results of semi-naturalistic modeling of a real INS were used.

Determination of the accuracy characteristics of the system in statics was carried out when installing the INS on the basis of the stand. Since the system is installed on a fixed base, the output information on the location and speed are INS errors. Navigation information was removed from the output and indication unit visually after 5 minutes.

When modeling based on laboratory experiment data, the system errors in speed δV_x, δV_y, were used as measurements in the estimation algorithm and the self-organization algorithm.

At the output of the estimation algorithm, we obtain estimates of the INS errors in speed, the angles of deviation of the GSP from the horizontal plane, and the horizontal drifts of the GSP.

Fig. 1 shows the result of estimating the angle of deviation of the gyrostabilized platform relative to the accompanying trihedron of the selected coordinate system.

The assessment was carried out by modeling the nonlinear Kalman filter and its modification.

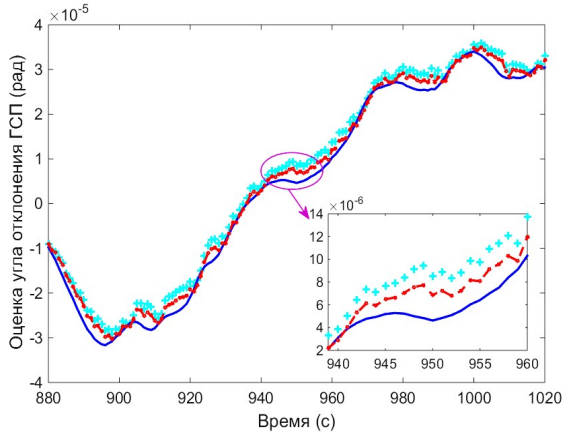


Fig. 1. Estimate of the GSP deviation angle of a real INS system obtained using a nonlinear Kalman filter and its modification

Fig. 1 shows: the azure dotted line is the GSP deviation angle obtained during a laboratory experiment; the blue solid line is the estimate of the GSP deviation angle obtained using a classic nonlinear Kalman filter; the red dotted line is the estimate of the GSP deviation angle obtained using a modified nonlinear Kalman filter.

The modified Kalman filter uses a self-organizing model.

III. CONCLUSIONS

The paper presents a methodology for assessing the reliability of the results of a flight experiment with a navigation complex that includes an inertial system and a satellite navigation system receiver. Joint signal processing was performed using linear and nonlinear adaptive kalman filters. The results of the semi-natural experiment demonstrated the advantage of the modified kalman filter. Under flight experiment conditions, this advantage will be even greater, since the a priori model will reflect the real process of changing ins errors worse.

ACKNOWLEDGEMENTS

The study was supported by the grant of the Russian Science Foundation No. 23-79-10028, https://rscf.ru/project/23-79-10028/

REFERENCES

[1] Goloshchapov A. Primenenie MEMS-tekhnologii v navigacii //Komponenty i tekhnologii. – 2014. №. 4 (153). S. 65-69.

[2] Neusypin K. A., Selezneva M. S. Korrekciya navigacionnoj sistemy letatel'nogo apparata s pomoshch'yu adaptivnyh fil'trov Kalmana //Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki. 2022. №. 11. S. 25-36.

[3] Dzhandzhgava G. I. i dr. Navigacionnyj kompleks s povyshennymi harakteristikami nablyudaemosti i upravlyaemosti //Aviakosmicheskoe priborostroenie. 2016. №. 6. S. 18-24.

[4] Shen' K., Proletarskij A. V., Neusypin K. A. Issledovanie algoritmov korrekcii navigacionnyh sistem letatel'nyh apparatov //Vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta im. NE Baumana. Seriya «Priborostroenie». 2016. №. 2 (107). – S. 28-39.

[5] Kaj Sh. i dr. Algoritmicheskie metody korrekcii navigacionnoj informacii s ispol'zovaniem sputnikovoj radionavigacionnoj sistemy v usloviyah anomal'nyh izmerenij //Radiotekhnika i elektronika. – 2019. – T. 64. – №. 1. – S. 31-37.

[6] Pupkov K. A., Neusypin K. A. Voprosy teorii i realizacii sistem upravleniya i navigacii. 1997.

[7] Selezneva M. S. i dr. Algoritmy obrabotki informacii navigacionnyh sistem i kompleksov letatel'nyh apparatov. – 2018.

[8] Stratonovich R. L. Uslovnye markovskie processy i ih primenenie k teorii optimal'nogo upravleniya. Ripol Klassik, 1966.

[9] Shahtarin B. I., Shen' K., Neusypin K. A. Modifikaciya nelinejnogo fil'tra Kalmana v skheme korrekcii navigacionnyh sistem letatel'nyh apparatov //Radiotekhnika i elektronika. 2016. T. 61. №. 11. S. 1065-1072.

[10] Mueller J. A., Ivachnenko A. G., Lemke F. GMDH algorithms for complex systems modelling //Mathematical and Computer Modelling of Dynamical Systems. 1998. T. 4. №. 4. S. 275-316.

[11] Onwubolu G. C. et al. Self-organizing data mining for weather forecasting //IADIS European conference data mining. 2007. S. 81-88.

Increasing the Degree of Parametric Identifiability of an Error Model for a UAV Navigation System

M.S. Selezneva
Department of Automatic Control Systems
Bauman Moscow State Technical University
Moscow, Russia
m.s.selezneva@mail.ru

K.A. Neusypin
Department of Automatic Control Systems
Bauman Moscow State Technical University
Moscow, Russia
neysipin@mail.ru

Abstract—A method for improving the quality characteristics of models used in the algorithmic support of aircraft is investigated. The efficiency of many objects depends on the accuracy of the algorithmic support. The algorithms use mathematical models with different quality characteristics. The article investigates a quality characteristic of models - the degree of parametric identifiability. To determine the quality characteristic, a numerical criterion for the degree of parametric identifiability of one class of nonlinear models is used. This class includes nonlinear models represented by the SDC method

Keywords— unmanned aerial vehicle; navigation system; model; SDC representation; criterion for the degree of parametric identifiability.

I. INTRODUCTION

Unmanned aerial vehicles (UAVs) have found wide application in various practical applications.

All over the world, UAVs play an increasingly important role in defense programs and defense strategy. Technological advances have made it possible to develop both large and small UAVs with ever-increasing capabilities.

Increasing the accuracy of UAVs can be achieved by using new, more advanced devices and systems, but this requires large financial and time costs. A more promising way to increase accuracy with minimal costs involves the use of algorithmic approaches [1, 2, 3]. Since the primary source of errors is the UAV navigation complex (NC), it is advisable to adjust the NC in order to reduce errors in determining the navigation parameters of the UAV.

Modern algorithmic support for NC includes algorithms for control, evaluation, identification, model construction and integration [4, 5]. One of the areas of improving the algorithmic support of the UAV NC is the use of models with improved characteristics, in particular, models with increased degrees of identifiability. Modern UAV NC of the class under study consist of an inertial navigation system (INS), a GNSS signal receiver and a special computer, which implements algorithms for processing INS and GNSS information, as well as UAV control. NC has errors associated with the insecurity of the radio channel and accumulating INS errors.

It is advisable to improve the accuracy of the existing NC by algorithmic means [6, 7].

To improve the accuracy of UAV navigation definitions, joint processing of several navigation systems combined into NC is used. One of the configurations of the UAV NC is a structure that includes INS, GPS and a nonlinear Kalman filter (NFK). With the help of the NFK, INS errors are estimated and subsequently compensated for in the output signal of the system.

Under UAV maneuvering conditions, a priori nonlinear models of INS error changes become inadequate to the real process, therefore, various NFK modifications are used [8, 9]. Model construction algorithms allow constructing a model of the process under study with high accuracy, but these models are non-physical, their structure is unknown in advance and constantly changes during the process of model construction. Model construction algorithms require increased performance of the onboard computer, so their use in a number of problems and in particular in the problem under consideration is difficult.

Simplification of the onboard implementation of estimation algorithms is achieved by identifying models of the process under evaluation, for example, by using the least squares method, but the accuracy of the estimation is reduced compared to modifications of the NFC.

Thus, when maneuvering the UAV, in order to maintain the adequacy of the INS error model, it is advisable to carry out parametric identification of the model. Identification algorithms allow determining the individual parameters of the model matrix, the structure of which is specified a priori.

The accuracy of determining the model parameters depends on the properties of the model matrix and the accuracy of the identification algorithm. The choice of the identification algorithm used is determined from considerations of the possibilities of implementation in the existing special computer of the UAV.

In practical applications, some parameters of the matrices of these models change and are reliably unknown, so they are identified. Determining the quality of identification of model parameters allows selecting the most accurate models, assessing the accuracy of correction of measuring systems, i.e., it is an important and urgent task.

One of the promising approaches to solving the problem of increasing the accuracy of identification is the use of models with improved identifiability characteristics.

The degree of identifiability is understood as the time interval during which it is possible to identify a parameter with a given accuracy. To calculate the degree of identifiability of the parameters of a nonlinear model, various linearization approaches are used, in particular, the SDC representation [10, 11].

To decrease the reduced measurement noise, as well as in the linear formulation of the problem, it is proposed to increase the degree of parametric identifiability of the parameter under study.

For linear stationary and non-stationary systems, numerical criteria for the degree of identifiability have been developed. When describing a dynamic object with nonlinear equations, various approaches to the linearization of the

model under study are used to calculate the degree of parametric identifiability. However, when performing the linearization procedure, the accuracy of the representation of a dynamic object is significantly reduced, with the exception of some cases when linearization is carried out without loss of accuracy. In this regard, it is advisable to develop a criterion for the degree of identifiability of the parameters of nonlinear models, which would be an indicator of the quality of identification.

II. DEVELOPMENT OF A NUMERICAL CRITERION FOR THE DEGREE OF IDENTIFIABILITY OF PARAMETERS OF NONLINEAR MODELS OF ONE CLASS

Well-known criteria for the degree of identifiability [12, 13, 14] assume the analysis of objects described by linear equations. However, there are often cases when the equations of an object have a clearly expressed nonlinear character. Let the equation for the state vector have the form:

$$\mathbf{x}_k = \mathbf{\Phi}_k\left(\mathbf{x}_{k-1}\right) + \mathbf{w}_k \tag{1}$$

where $\mathbf{x}_k$ is the state vector, $\mathbf{\Phi}_k\left(\mathbf{x}_{k-1}\right)$ is the vector of the nonlinear model characterizing the dynamics of the process under study. Part of the state vector is measured using the INS and GPS navigation systems:

$$\mathbf{z}_k = \mathbf{H}_\mathrm{k}\mathbf{x}_k + \mathbf{v}_k \tag{2}$$

where $\mathbf{z}_k$ is the measurement vector, $\mathbf{H}_k$ is the measurement matrix, $\mathbf{w}_k$ and $\mathbf{v}_k$ are discrete analogues of Gaussian white noise with zero mathematical expectations and covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively, uncorrelated with each other. The equation of the object in the SDC representation and the equation of measurements have the form (1), (2). In this case, the state vector $\mathbf{x}_{k+n}$ can be expressed by its value at the initial moment of time in the form

$$\begin{aligned}\mathbf{x}_{k+n} &= \mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_k,\mathbf{x}_k)\mathbf{x}_k \\ &+\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_{k+1},\mathbf{x}_{k+1})\mathbf{G}(t_k,\mathbf{x}_k)\mathbf{w}_k \\ &+\cdots+\mathbf{G}(t_{k+n-1},\mathbf{x}_{k+n-1})\mathbf{w}_{k+n-1}.\end{aligned} \tag{3}$$

Substituting the expression for $\mathbf{x}_{k+n}$ to the equation of measurements $\mathbf{y}_{k+n}$ yields

$$\begin{aligned}\mathbf{y}_{k+n} &= \mathbf{H}_{k+n}\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_k,\mathbf{x}_k)\mathbf{x}_k \\ &+\mathbf{H}_{k+n}\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_{k+1},\mathbf{x}_{k+1})\mathbf{G}(t_k,\mathbf{x}_k)\mathbf{w}_k \\ &+\cdots+\mathbf{H}_{k+n}\mathbf{G}(t_{k+n-1},\mathbf{x}_{k+n-1})\mathbf{w}_{k+n-1}+\mathbf{v}_{k+n}.\end{aligned} \tag{4}$$

Substituting the expression $\mathbf{x}_k$ into this equation, we obtain

$$\begin{aligned}\mathbf{y}_{k+n} &= \mathbf{H}_{k+n}\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_k,\mathbf{x}_k)\mathbf{O}_k^{+}\mathbf{y}_k^{*} \\ &-\mathbf{H}_{k+n}\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_k,\mathbf{x}_k)\mathbf{O}_k^{+}\mathbf{v}_k^{*} \\ &+\mathbf{H}_{k+n}\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_{k+1},\mathbf{x}_{k+1})\mathbf{G}(t_k,\mathbf{x}_k)\mathbf{w}_k \\ &+\cdots+\mathbf{H}_{k+n}\mathbf{G}(t_{k+n-1},\mathbf{x}_{k+n-1})\mathbf{w}_{k+n-1}+\mathbf{v}_{k+n},\end{aligned} \tag{5}$$

where $\mathbf{O}_k^{+}=\left[\mathbf{O}_k^{*T}\mathbf{O}_k^{*}\right]^{-1}\mathbf{O}_k^{*T}$ – pseudoinverse matrix $\mathbf{O}_k^{*}$.

Let's introduce the notation

$$\left[\lambda_{1,k}\quad\lambda_{2,k}\quad\cdots\quad\lambda_{n,k}\right]=\mathbf{H}_{k+n}\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_k,\mathbf{x}_k)\mathbf{O}_k^{+}. \tag{6}$$

$$\begin{aligned}\mathbf{v}_k^0 &= \gamma_{1,k}\mathbf{w}_k+\gamma_{2,k}\mathbf{w}_{k+1}+\cdots+\gamma_{n,k}\mathbf{w}_{k+n-1} \\ &-\lambda_{1,k}\mathbf{v}_k-\lambda_{2,k}\mathbf{v}_{k+1}-\cdots-\lambda_{n,k}\mathbf{v}_{k+n-1}+\mathbf{v}_{k+n} \\ &=\mathbf{H}_{k+n}\mathbf{\Phi}(t_{k+n-1},\mathbf{x}_{k+n-1})\cdots\mathbf{\Phi}(t_k,\mathbf{x}_k)\mathbf{O}_k^{+}\mathbf{v}_k^{*} \\ &+\cdots+\mathbf{H}_{k+n}\mathbf{G}(t_{k+n-1},\mathbf{x}_{k+n-1})\mathbf{w}_{k+n-1}+\mathbf{v}_{k+n}.\end{aligned}$$

Then the problem statement is reduced to determining the unknown non-stationary elements of the column vector $\left[\lambda_{1,k}\quad\lambda_{2,k}\quad\cdots\lambda_{n,k}\right]$ according to newly formed dimensions, i.e.,

$$\begin{aligned}\lambda_{1,k} &= f_{1,k}\left(y_k,\cdots,y_{k+2n-1}\right)+v_k^{00}; \\ \lambda_{2,k} &= f_{2,k}\left(y_k,\cdots,y_{k+2n-1}\right)+v_{k+1}^{00}; \\ &\cdots\quad\cdots\quad\cdots \\ \lambda_{n,k} &= f_{n,k}\left(y_k,\cdots,y_{k+2n-1}\right)+v_{k+n-1}^{00},\end{aligned} \tag{7}$$

Where

$$\begin{bmatrix}f_{1,k}\left(y_k,\cdots,y_{k+2n-1}\right)\\ f_{2,k}\left(y_k,\cdots,y_{k+2n-1}\right)\\ \cdots\\ f_{n,k}\left(y_k,\cdots,y_{k+2n-1}\right)\end{bmatrix}=\begin{bmatrix}y_k & y_{k+1} & \cdots & y_{k+n-1}\\ y_{k+1} & y_{k+2} & \cdots & y_{k+n}\\ \cdots & \cdots & \cdots & \cdots\\ y_{k+n-1} & y_{k+n} & \cdots & y_{k+2n-2}\end{bmatrix}^{-1}\begin{bmatrix}y_{k+n}\\ y_{k+n+1}\\ \cdots\\ y_{k+2n-1}\end{bmatrix}$$

$$\begin{bmatrix}v_k^{00}\\ v_{k+1}^{00}\\ \cdots\\ v_{k+n-1}^{00}\end{bmatrix}=\begin{bmatrix}y_k & y_{k+1} & \cdots & y_{k+n-1}\\ y_{k+1} & y_{k+2} & \cdots & y_{k+n}\\ \cdots & \cdots & \cdots & \cdots\\ y_{k+n-1} & y_{k+n} & \cdots & y_{k+2n-2}\end{bmatrix}^{-1}\begin{bmatrix}v_k^{0}\\ v_{k+1}^{0}\\ \cdots\\ v_{k+n-1}^{0}\end{bmatrix}$$

Therefore, the criterion of the degree of identifiability of the parameters of the model of dynamic non-stationary systems has the form:

$$DI_{Nk}^{i}=\frac{E\left[\left(\lambda_{i,k}\right)^2\right]R_0}{E\left[\left(y_{i,k}\right)^2\right]\hat{R}_k^{i}}, \tag{8}$$

where $E\left[\left(\lambda_{i,k}\right)^2\right]$ – variance of an arbitrary i-th component of the parameter vector λ; $E\left[\left(z_{i,k}\right)^2\right]$ – variance of the directly measured state vector; R_0 – variance of the original measurement noise; $\hat{R}_k^{i}$ – variance of the reduced measurement noise.

The variance of the initial measurement noise is determined for practical reasons in accordance with the operating mode of the measuring system or the value from the measuring device passport is taken. Certain difficulties arise when calculating the reduced measurement noise.

However, when using an adaptive estimation algorithm for identification, the variance of the reduced measurement noise is calculated at each step of the algorithm operation. The developed numerical criterion for the degree of identifiability has a clear physical meaning, is simple, and allows calculating the quality of parameter identification in the form of a scalar. The criterion is simple and convenient in practical applications.

To reduce computational costs when implementing algorithmic support for UAV NC, a scalar form of algorithm representation is used. Direct and reduced measurements of the system under study are used for identification. Obviously, the accuracy of identification can be increased by improving the accuracy of these measurements, i.e., reducing the measurement noise.

The use of scalar identification algorithms assumes complete parametric identifiability of the system under study. Therefore, before using such identification, it is necessary to make sure that the INS error model is identifiable using the selected measurements.

When checking the performance of the estimation algorithms, the classical INS error model is adopted. If the obtained INS error models have sufficient accuracy, they can be used to correct navigation information. A semi-naturalistic modeling was carried out based on laboratory experiment data using the classical Kalman filter and the Kalman filter with scalar parameter identification.

The use of the identification algorithm makes it possible to form models with improved parametric identifiability characteristics, the use of which in the algorithmic support of the navigation complex allows for the effective correction of navigation information.

Fig. 1 shows: 1 – RMS of estimation errors using the classical Kalman filter; 2 – RMS of estimation errors using the Kalman filter with parametric identification.

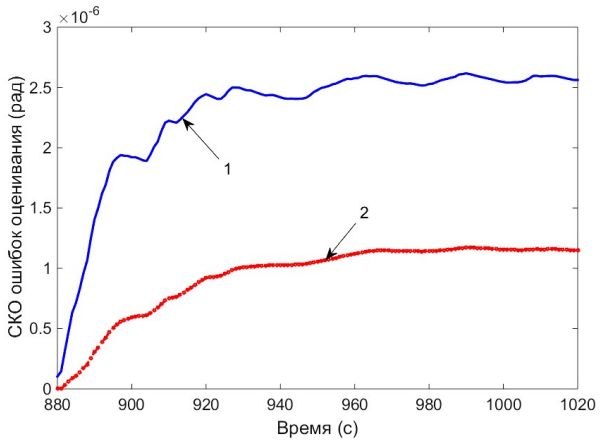


Fig. 1. Standard deviation of estimation errors using the Kalman filter and the Kalman filter with parametric identification

In real UAV operation conditions, it can be assumed that the advantage of Kalman filters with parametric identification will be even more pronounced, since the degree of adequacy of the mathematical model used in the classical Kalman filter is significantly reduced, especially during intensive UAV maneuvering.

The results of modeling the NC errors demonstrated the advantages of models with improved quality characteristics.

III. CONCLUSIONS

Well-known qualitative criteria, in particular the criterion of the degree of parametric identifiability, are widely used to evaluate the qualitative characteristics of linear models. The article presents a numerical criterion of the degree of parametric identifiability of nonlinear models represented by the SDC method.

Thus, using qualitative criteria, models with improved characteristics are created.

ACKNOWLEDGEMENTS

The study was supported by the grant of the Russian Science Foundation No. 23-79-10028, https://rscf.ru/project/23-79-10028/

REFERENCES

[1] Fam S. Ch., Neusypin K. A., Selezneva M. S. Issledovanie sistemy marshrutnoj korrekcii navigacionnoj sistemy BLA //XXVII Sankt-Peterburgskaya mezhdunarodnaya konferenciya po integrirovannym navigacionnym sistemam. 2020. S. 346-348.

[2] Nin T. Metody korrekcii navigacionnyh sistem bespilotnyh letatel'nyh apparatov //Politekhnicheskij molodezhnyj zhurnal. 2019. №. 9. S. 6.

[3] Kaj Sh. Razrabotka metodov ocenivaniya i prognozirovaniya navigacionnyh sistem letatel'nyh apparatov //Avtomatizaciya. Sovremennye tekhnologii. 2015. №. 7. S. 13-18.

[4] Shatilov A. Yu., Nagin I. A. Tesno svyazannyj algoritm kompleksirovaniya NAP SRNS i mnogocelevoj INS //Radiotekhnika. 2012. №. 6. S. 118-126.

[5] Neusypin K. A., Proletarskij A. V., Selezneva M. S. Algoritmy kompleksirovaniya inercial'noj navigacionnoj sistemy s datchikami uglovyh uskorenij //Mekhatronika, avtomatizaciya, upravlenie. 2023. T. 24. №. 2. S. 107-112.

[6] Selezneva M. S., Neusypin K. A., Proletarsky A. V. Navigation complex with adaptive non-linear Kalman filter for unmanned flight vehicle //Metrology and Measurement Systems. – 2019. – T. 26. – №. 3.

[7] Selezneva M. S. et al. Modification of the Federated Kalman Filter Using the Observability Degree Criterion of State Variables //2019 26th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS). – IEEE, 2019. – S. 1-3.

[8] Simon D. Using nonlinear Kalman filtering to estimate signals //Embedded Systems Design. – 2006. T. 19. №. 7. S. 38.

[9] Julier S. J., Uhlmann J. K. New extension of the Kalman filter to nonlinear systems //Signal processing, sensor fusion, and target recognition VI. Spie, 1997. T. 3068. S. 182-193.

[10] Afanasiev V. N. Guaranteed control of feedback linearizable nonlinear object //AIP Conference Proceedings. – American Institute of Physics, 2012. – T. 1493. – №. 1. – S. 13-19.

[11] Afanas'ev V.N. Upravlenie nelinejnymi neopredelennymi dinamicheskimi ob"ektami. M.: Librokom/URSS, 2015. 224 s.

[12] Proletarsky A. V., Neusypin K. A., Selezneva M. S. Method for Improving Accuracy of INS using Scalar Parametric Identification //2019 International Russian Automation Conference (RusAutoCon). IEEE, 2019. S. 1-4.

[13] Shen' K., Neusypin K. A. Issledovanie kriteriev stepenej nablyudaemosti, upravlyaemosti i identificiruemosti linejnyh dinamicheskih sistem //Mekhatronika, avtomatizaciya, upravlenie. 2016. T. 17. №. 11. S. 723-731.

[14] Cibizova T. Yu., P'o S., Selezneva M. S. Matematicheskoe modelirovanie dinamicheskih sistem s ispol'zovaniem parametricheskoj identificiruemosti //Sovremennye naukoemkie tekhnologii. 2018. №. 1. S. 54-60.

Performance Evaluation of Autonomous Berthing System for Unmanned Boat

Haoze Zhang
Navigation College
Dalian Maritime University
Dalian, China
zhz11@ dlmu.edu.cn

Yingjun Zhang
Navigation College
Dalian Maritime University
Dalian, China
zhangyj@dlmu.edu.cn

Zhuolin Wang
Navigation College
Dalian Maritime University
Dalian, China
wangzhuol@dlmu.edu.cn

***Abstract*—Autonomous berthing is a critical component of fully autonomous navigation technology for MASS (Maritime Autonomous Surface Ship), relying on the coordinated operation of perception, decision-making, and control modules within the autonomous berthing system. Before such systems can be deployed on MASS, they must undergo rigorous validation through virtual simulation and unmanned boat testing. However, due to current technological limitations, most autonomous berthing systems remain at the unmanned boat testing stage and are not yet mature enough for widespread application to MASS. To address this gap, this paper investigates the performance evaluation of autonomous berthing systems for unmanned boat and proposes a quantitative assessment method based on G1 method and Fuzzy Comprehensive Evaluation (FCE). Experimental results demonstrate that the proposed method can objectively assess the reliability and safety of unmanned boat autonomous berthing systems, providing theoretical support for the further development of autonomous berthing technologies.**

Keywords—unmanned boat, autonomous berthing, G1 method, fuzzy comprehensive evaluation

I. INTRODUCTION

With the rapid development of artificial intelligence and intelligent control technologies, the autonomous operation capability of unmanned systems in complex scenarios has become a research hotspot. Thorough theoretical research and experimental validation are prerequisites for the reliable application of these unmanned systems. Similarly, autonomous berthing systems must undergo virtual simulation testing, unmanned boat testing, and full-scale ship testing before they can be widely applied. Unmanned boat testing serves as a crucial link between theoretical research and engineering applications, playing an essential role in providing data support for subsequent testing and technological upgrades. Therefore, this paper conducts a study on the performance evaluation of autonomous berthing systems for unmanned boat, offering methodological support to accelerate the transition of autonomous berthing technology from experimental validation to engineering application.

In 2020, Fan [1] conducted research on the autonomous navigation capability testing of unmanned boat. They established an evaluation index system for the autonomous navigation capabilities of unmanned boat from four aspects: autonomous perception, autonomous planning, autonomous collision avoidance, and autonomous control. Based on Analytic Hierarchy Process (AHP) and FCE, they constructed an evaluation model and validated the proposed method through the collection of simulation experiment data from the "Lanxin" unmanned boat at Dalian Maritime University.

In 2021, Su [2] conducted research on the testing and evaluation of autonomous navigation capabilities for unmanned ship. They established an evaluation index system for the autonomous navigation capabilities of unmanned ship from four aspects: perception and recognition ability, path planning ability, collision avoidance decision-making ability, and navigation control ability. Based on AHP and a cloud model-improved FCE, they constructed an evaluation model and validated the proposed method through the collection of simulation experiment data from the "Lanxin" unmanned boat at Dalian Maritime University.

In 2022, Zhang [3] conducted research on the performance evaluation of the autonomous berthing system for MASS. They established a performance evaluation index system for the autonomous berthing system of MASS from four aspects: safety, efficiency, intelligence, and economy. Based on DEMATEL-ANP, they constructed an evaluation model and validated the proposed method through the collection of autonomous berthing data from an intelligent container ship.

In 2023, Fan [4] conducted research on the performance evaluation of the automatic collision avoidance system for MASS. A performance evaluation index system was established from three aspects: safety indicators, economic indicators, and efficiency indicators. Based on the simple correlation function, G1 method, and FCE, an evaluation model was constructed. The proposed method was validated using simulation test data.

In summary, current research primarily focuses on the evaluation of autonomous navigation systems, autonomous berthing systems, and automatic collision avoidance systems for MASS, as well as autonomous navigation systems for unmanned boat. However, most evaluation indicators cannot be directly applied to the performance evaluation of unmanned boat autonomous berthing systems. As a result, there is currently a lack of a scientific and accurate performance evaluation system for unmanned boat autonomous berthing systems. To address this, this paper establishes an evaluation framework with safety, intelligence, and adaptability as primary indicators. Based on this, a performance evaluation index system for unmanned boat autonomous berthing systems is constructed, with both quantitative calculations and qualitative analysis used to achieve the quantification of the indicators. The weights of the indicators at various levels are determined using G1 method, and a performance evaluation model for unmanned boat autonomous berthing systems is built by FCE. The proposed performance evaluation model for autonomous berthing systems is verified through a case study.

This work was supported by the National Natural Science Foundation of China [Grant No.52231014] and Team Project of Dalian Maritime University [Grant No.3132023511].

II. Establishment of Performance Evaluation Indicators of Autonomous Berthing System for Unmanned Boat

A. Indicator Selection

To develop a rational set of evaluation indicators, this study adheres to the principles of scientific validity, goal orientation, completeness, operability, significance, and dynamism in constructing the performance evaluation index system of autonomous berthing system for unmanned boat. As shown in Fig. 1, through literature review, acquisition of domain-specific knowledge, expert experience, and analysis of real-world cases, an evaluation framework is established with safety, intelligence, and adaptability as the primary indicators. Based on this framework, nine secondary indicators are derived, forming a comprehensive performance evaluation index system for unmanned boat autonomous berthing system.

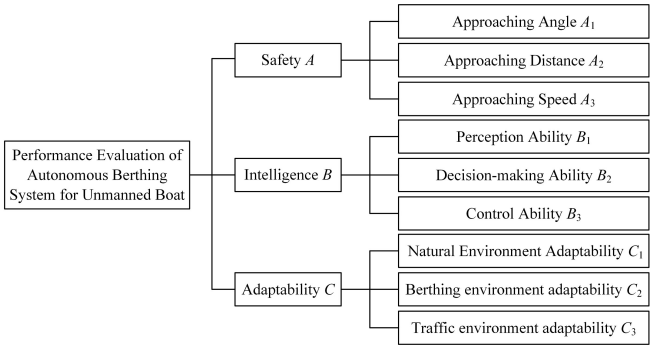


Fig. 1. Performance evaluation index system for the autonomous berthing system of unmanned boat

B. Quantification of indicators

(1) Safety Indicators

The approaching angle refers to the angle between the boat's heading line and the front edge of the berth after the boat enters the berth. According to nautical experience, the approaching angle should not be too large. The membership values of this indicator are defined using trapezoidal, triangular, and inverted trapezoidal membership functions, with the corresponding membership function diagrams shown in Fig. 2.

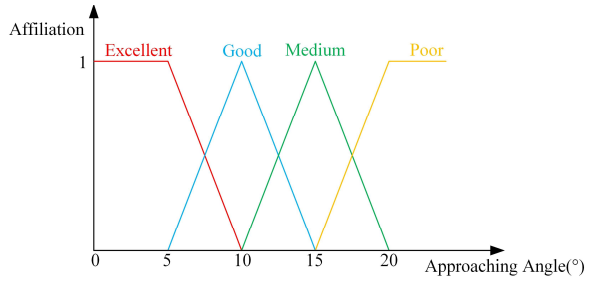


Fig. 2. Membership diagram of the approaching angle

The "Excellent" grade corresponds to an approaching angle range of $(0,\ 10)$ degrees. Its membership function is defined using a left-half trapezoidal distribution, as shown in Eq. (1).

$$\mu_{D_{11}}(x)=\begin{cases}1 & x\le 5\\ (10-x)/5 & 5\le x<10\\ 0 & x\ge 10\end{cases} \tag{1}$$

The "Good" grade corresponds to an approaching angle range of $(5,\ 15)$ degrees. Its membership function is defined using a triangular distribution, as shown in Eq. (2).

$$\mu_{D_{12}}(x)=\begin{cases}(x-5)/5 & 5\le x<10\\ (15-x)/5 & 10\le x<15\\ 0 & other\end{cases} \tag{2}$$

The "Medium" grade corresponds to an approaching angle range of $(10,\ 20)$ degrees. Its membership function is defined using a triangular distribution, as shown in Eq. (3).

$$\mu_{D_{13}}(x)=\begin{cases}(x-10)/5 & 10\le x<15\\ (20-x)/5 & 15\le x<20\\ 0 & other\end{cases} \tag{3}$$

The "Poor" grade corresponds to an approaching angle range of $(15,\ +\infty)$ degrees. Its membership function is defined using a triangular distribution, as shown in Eq. (4).

$$\mu_{D_{14}}(x)=\begin{cases}0 & x\le 15\\ (x-15)/5 & 15\le x<20\\ 1 & x\ge 20\end{cases} \tag{4}$$

The approaching distance refers to the lateral distance between the midship of the boat and the forward edge of the berth after the unmanned boat has entered the berth. According to navigational experience, it is recommended that the approaching distance be maintained at approximately 1.5 times the ship's beam when berthing. The membership values for this indicator are defined using membership functions with descending half-trapezoidal, triangular, and ascending half-trapezoidal distributions. The membership function diagram is shown in Fig. 3.

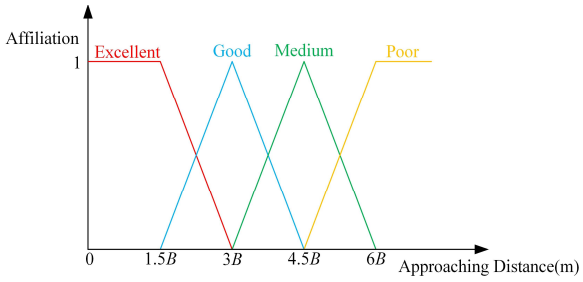


Fig. 3. Membership diagram of the approaching distance

The "Excellent" grade corresponds to an approaching distance range of $(0,\ 3B)$ degrees. Its membership function is defined using a left-half trapezoidal distribution, as shown in Eq. (5).

$$\mu_{D_{21}}(x)=\begin{cases}1 & x\le 1.5B\\ (3B-x)/1.5B & 1.5B\le x<3B\\ 0 & x\ge 3B\end{cases} \tag{5}$$

The "Good" grade corresponds to an approaching distance range of $(1.5B,\ 4.5B)$ degrees. Its membership function is defined using a triangular distribution, as shown in Eq. (6).

$$\mu_{D_{22}}(x)=\begin{cases}(x-1.5B)/1.5B & 1.5B\le x<3B\\ (4.5B-x)/1.5B & 3B\le x<4.5B\\ 0 & other\end{cases} \tag{6}$$

The "Medium" grade corresponds to an approaching distance range of $(3B,\ 6B)$ degrees. Its membership function is defined using a triangular distribution, as shown in Eq. (7).

$$\mu_{D_{23}}(x)=\begin{cases}(x-3B)/1.5B & 3B\le x<4.5B\\ (6B-x)/1.5B & 4.5B\le x<6B\\ 0 & other\end{cases} \quad (7)$$

The "Poor" grade corresponds to an approaching distance range of $(4.5B,\ +\infty)$ degrees. Its membership function is defined using a triangular distribution, as shown in Eq. (8).

$$\mu_{D_{24}}(x)=\begin{cases}0 & x\le 4.5B\\ (x-4.5B)/1.5B & 4.5B\le x<6B\\ 1 & x\ge 6B\end{cases} \quad (8)$$

The approaching speed refers to the unmanned boat's speed after entering the berth. While berthing, boats typically maintain a certain residual speed to adjust their position. However, unmanned boats differ from traditional ships in that they must reduce their speed to $0\mathrm{m/s}$ upon reaching the berth to avoid collision hazards. Therefore, if the approaching speed is $0\mathrm{m/s}$, the evaluation grade for this indicator is considered "Excellent"; if the approaching speed is not $0\mathrm{m/s}$, the evaluation grade is considered "Poor".

TABLE I. QUANTITATIVE ANALYSIS OF PERCEPTUAL ABILITY

Indicator	Weight	Definition	Score
Position and State Perception Ability	0.5	Equipped with high-precision pose perception, capable of providing accurate pose information in real-time	100
		Pose perception is relatively accurate and can provide reliable pose information in most situations	75
		The pose perception ability is average and unable to provide accurate pose information	50
		The pose perception ability is poor and cannot provide stable pose information	25
Obstacle Perception Ability	0.4	Capable of accurately identifying all static and dynamic obstacles around the unmanned boat	100
		Able to accurately identify obstacles in most situations, with occasional recognition errors	75
		The obstacle recognition ability is average, with a tendency to miss or misidentify surrounding obstacles	50
		The obstacle perception ability is extremely poor, with little to no ability to identify surrounding obstacles	25
Natural Environment Perception Ability	0.1	Capable of accurately perceiving various natural environmental factors	100
		Capable of perceiving the main natural environmental factors	75
		Capable of perceiving some natural environmental factors	50
		The natural environment perception ability is insufficient	25

(2) Intelligence Indicators

In the quantification of the three indicators: perception ability, decision-making ability, and control ability-the factors, weights, and their definitions considered comprehensively are shown in Tables 1 to 3. Among them, the evaluation grades are: 'Excellent' for a score of 80-100, 'Good' for a score of 60-80, 'Medium' for a score of 40-60, and 'Poor' for a score of 0-40.

TABLE II. QUANTITATIVE ANALYSIS OF DECISION-MAKING ABILITY

Indicator	Weight	Definition	Score
Global path planning ability	0.4	Possesses the ability to efficiently perform global path planning	100
		Possesses basic global path planning ability	75
		Possesses global path planning ability	50
		The global path planning ability is significantly insufficient	25
Local path planning ability	0.35	Possesses the ability to efficiently perform local path planning	100
		Possesses basic local path planning ability	75
		Possesses local path planning ability	50
		The local path planning ability is significantly insufficient	25
Task allocation ability	0.25	Possesses the ability to efficiently allocate tasks.	100
		Possesses a strong task allocation ability	75
		Possesses task allocation ability	50
		The task allocation ability is significantly insufficient	25

TABLE III. QUANTITATIVE ANALYSIS OF CONTROL ABILITY

Indicator	Weight	Definition	Score
Heading control accuracy	0.4	The error is less than or equal to 1°	100
		The error is less than or equal to 3°	75
		The error is less than or equal to 5°	50
		The error is greater than 5°	25
Speed control accuracy	0.2	The error is less than or equal to 0.2m/s	100
		The error is less than or equal to 0.4m/s	75
		The error is less than or equal to 0.6m/s	50
		The error is greater than 0.6m/s	25
Path tracking control accuracy	0.4	The error is less than or equal to 1B	100
		The error is less than or equal to 2B	75
		The error is less than or equal to 3B	50
		The error is greater than 3B	25

(3) Adaptability Indicators

Under the adaptability indicator, this paper establishes three secondary indicators: natural environment adaptability, berthing environment adaptability, and traffic environment adaptability. The quantification levels for each indicator are shown in Table 4.

TABLE IV. THE WEIGHTS OF EACH LEVEL OF INDICATORS

Indicator	Grade	Definition
C_1	Excellent	Autonomous berthing can be achieved under conditions of poor visibility and strong wind and current
	Good	Autonomous berthing can be achieved under conditions of good visibility and strong wind and current

Indicator	Grade	Definition
	Medium	Autonomous berthing can be achieved under conditions of poor visibility with no wind and no current
	Poor	Autonomous berthing can be achieved under conditions of good visibility with no wind and no current
C_2	Excellent	Autonomous berthing can be achieved in the presence of parked boats and other obstacles near the berth, as well as in the presence of navigating ships
	Good	Autonomous berthing can be achieved in the absence of parked boats and other obstacles near the berth, even in the presence of navigating ships
	Medium	Autonomous berthing can be achieved in the presence of parked boats and other obstacles near the berth, with no navigating ships
	Poor	Autonomous berthing can be achieved in the absence of parked boats and other obstacles near the berth, with no navigating ships
C_3	Excellent	Autonomous berthing can be achieved under conditions of high traffic flow
	Good	Autonomous berthing can be achieved under conditions of heavy traffic flow
	Medium	Autonomous berthing can be achieved under conditions of moderate traffic flow
	Poor	Autonomous berthing can be achieved under conditions of low traffic flow

III. Establishment of the Performance Evaluation Model for the Autonomous Berthing System of Unmanned Boat

A. Weight Calculation Model

The G1 method is a subjective weighting approach derived from the AHP. It calculates the weights of each indicator based on expert rankings of the importance of evaluation criteria and the importance ratio between adjacent indicators. Compared to AHP, this method does not require the establishment of a judgment matrix or consistency testing, effectively avoiding complex calculations.

The experts selected in this study include individuals involved in the development of autonomous berthing systems, scholars researching ship maneuvering, as well as experienced captains or pilots. By incorporating the advice of experts from different fields, the impact of subjectivity is minimized to the greatest extent, making the evaluation results more reasonable. The specific steps for the operation are outlined as follows.

(1) Determine the sequential relationships between the indicators

As shown in Eq. (9), experts rank the evaluation indicators U_1、$U_2 \ldots U_\vartheta$ in order of importance from high to low, forming a sequential relationship.

$$U_1^* > U_2^* > \cdots > U_t^* > U_\vartheta^* \tag{9}$$

Where, $t = \vartheta,\ \vartheta - 1,\ \cdots,\ 2$ is a positive integer, and i represents the number of indicators.

(2) Determine the importance ratio between adjacent indicators

Experts need to provide the importance ratio between adjacent indicators. Let the importance ratio between two adjacent indicators U_{t-1}^* and U_t^* in the sequential relationship be denoted as r_t, where:

$$r_t = U_{t-1}^* / U_t^* \tag{10}$$

Where, the value of r_t can be referenced from Table 5.

TABLE V. The reference value of the r_t

Value	Meaning
1.0	The indicator U_{t-1}^* is equally important as the indicator U_t^*
1.1	The indicator U_{t-1}^* is between equally important and slightly more important than the indicator U_t^*
1.2	The indicator U_{t-1}^* is slightly more important than the indicator U_t^*
1.3	The indicator U_{t-1}^* is between slightly more important and significantly more important than the indicator U_t^*
1.4	The indicator U_{t-1}^* is significantly more important than the indicator U_t^*
1.5	The indicator U_{t-1}^* is between significantly more important and extremely more important than the indicator U_t^*
1.6	The indicator U_{t-1}^* is extremely more important than the indicator U_t^*
1.7	The indicator U_{t-1}^* is between extremely more important and critically more important than the indicator U_t^*
1.8	The indicator U_{t-1}^* is critically more important than the indicator U_t^*

(3) Determine the subjective weights.

Let the weight of the indicator U_t^* be w_t, then:

$$w_t = \left(1 + \sum_{t=2}^{\vartheta} \prod_{j=t}^{\vartheta} r_t\right)^{-1} \tag{11}$$

Where j is a positive integer.

The weight of the previous indicator is calculated sequentially, and it is given by:

$$w_{t-1} = w_t \cdot r_t \tag{12}$$

After normalization, the final weight W_t of the indicator can be obtained.

$$W_t = \frac{w_t}{\sum_1^\vartheta w_t} \tag{13}$$

As shown in Table 6, based on the above method, the comprehensive weights of the primary and secondary indicators are determined by inviting five experts to rank the importance of the indicators and assign relative importance values.

TABLE VI. The weights of each level of indicators

Primary Indicator	Weight	Secondary Indicator	Weight	Comprehensive Weight
A	0.3909	A_1	0.2155	0.0842
		A_2	0.3017	0.1179
		A_3	0.4828	0.1887
B	0.3553	B_1	0.3629	0.1289
		B_2	0.4355	0.1547
		B_3	0.2016	0.0716
C	0.2538	C_1	0.3000	0.0762
		C_2	0.2500	0.0636
		C_3	0.4500	0.1142

B. Evaluation Model

FCE based on fuzzy mathematics, applies the principle of fuzzy relation composition to quantify factors that are unclear or difficult to measure, providing a comprehensive evaluation of the membership degree of the evaluated object from multiple factors [5]. The specific steps of the method are as follows.

(1) Determine the evaluation indicator set

Define the evaluation indicator set as $\boldsymbol{U} = \{U_1,\ U_2 \ldots U_\vartheta\}$.

(2) Determine the evaluation grade set

Define the evaluation grade set as $\boldsymbol{Y} = \{Y_1,\ Y_2 \ldots Y_\vartheta\}$.

(3) Determine the indicator weight set

Define the weight set as $\boldsymbol{W} = \{W_1,\ W_2 \ldots W_\vartheta\}$, and satisfy the Equation (14).

$$\sum_{n=1}^{\vartheta} W_n = 1 \tag{14}$$

(4) Construct the fuzzy evaluation matrix

Let $\boldsymbol{Q}$ be the known fuzzy evaluation matrix, which defines a unique fuzzy evaluation from $\boldsymbol{U}$ to $\boldsymbol{Y}$, and its relationship satisfies Equation (15). The fuzzy evaluation result set $\boldsymbol{H}$ can be obtained using this equation, and the evaluation result can be determined by applying the maximum membership principle.

$$\boldsymbol{H} = \boldsymbol{W} \times \boldsymbol{Q} \tag{15}$$

IV. Experiment on the Performance Evaluation of the Autonomous Berthing System for Unmanned Boat

The data from a set of autonomous berthing tests for unmanned boat are evaluated based on the method proposed in this paper [6], thereby verifying the rationality and effectiveness of the method. The data for each level of indicators and the corresponding fuzzy evaluation matrix are shown in Table 7.

TABLE VII. Data of each indicators and fuzzy evaluation matrix

Indicator	Test Data	Fuzzy Evaluation Matrix
A_1	4.01°	(0,0,0,1)
A_2	0.6353m	(0,0,0,1)
A_3	0m/s	(0,0,0,1)
B_1	Good	(0,0,1,0)
B_2	Medium	(0,1,0,0)
B_3	Excellent	(0,0,0,1)
C_1	Good	(0,0,1,0)
C_2	Medium	(0,1,0,0)
C_3	Medium	(0,1,0,0)

Based on the data obtained from the unmanned boat's autonomous berthing test, the evaluation indicator set, evaluation grade set, indicator weight set, and fuzzy evaluation set are established respectively.

Establish the evaluation indicator set:

$\boldsymbol{U} = \{A, B, C\}$;

$\boldsymbol{U_1} = \{A_1, A_2, A_3\}$;

$\boldsymbol{U_2} = \{B_1, B_2, B_3\}$;

$\boldsymbol{U_3} = \{C_1, C_2, C_3\}$.

Establish the evaluation grade set:

$\boldsymbol{Y} = \{\text{Poor},\ \text{Medium},\ \text{Good},\ \text{Excellent}\}$.

Establish the indicator weight set:

$\boldsymbol{W} = \{0.0842,0.1179,0.1887,0.1289,0.1547,0.0716,$

$0.0762,0.0636,0.1142\}$;

$\boldsymbol{W_1} = \{0.2155,\ 0.3017,\ 0.4828\}$;

$\boldsymbol{W_2} = \{0.3629,\ 0.4355,\ 0.2016\}$;

$\boldsymbol{W_3} = \{0.3000,\ 0.2500,\ 0.4500\}$.

Establish the fuzzy evaluation set:

$$\boldsymbol{Q} = \begin{vmatrix} 0.000 & 0.000 & 0.000 & 1.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \\ 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \\ 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \end{vmatrix};$$

$$\boldsymbol{Q_1} = \begin{vmatrix} 0.000 & 0.000 & 0.000 & 1.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \end{vmatrix};$$

$$\boldsymbol{Q_2} = \begin{vmatrix} 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \end{vmatrix};$$

$$\boldsymbol{Q_3} = \begin{vmatrix} 0.000 & 0.000 & 1.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 1.000 & 0.000 & 0.000 \end{vmatrix}.$$

Establish the fuzzy evaluation result set:

The fuzzy evaluation result set for the performance of the unmanned boat autonomous berthing system is:

$$\boldsymbol{H} = \boldsymbol{W} \times \boldsymbol{Q} = \{0.0000,0.3325,0.2051,0.4624\}$$

Based on the maximum membership principle, the performance evaluation grade for the unmanned boat autonomous berthing system designed in this paper is "Excellent".

The fuzzy evaluation result set for the safety indicator *A* is:

$$\boldsymbol{H_1} = \boldsymbol{W_1} \times \boldsymbol{Q_1} = \{0.0000,0.0000,0.0000,1.0000\}$$

Based on the maximum membership principle, the evaluation grade for the safety indicator *A* is "Excellent".

The fuzzy evaluation result set for the intelligence indicator *B* is:

$$\boldsymbol{H_2} = \boldsymbol{W_2} \times \boldsymbol{Q_2} = \{0.0000,0.4355,0.3629,0.2016\}$$

Based on the maximum membership principle, the evaluation grade for the intelligence indicator *B* is "Medium".

The fuzzy evaluation result set for the adaptability indicator C is:

$$\boldsymbol{H_3} = \boldsymbol{W_3} \times \boldsymbol{Q_3} = \{0.0000, 0.7000, 0.3000, 0.0000\}$$

Based on the maximum membership principle, the evaluation grade for the adaptability indicator C is "Medium".

Through calculation, the comprehensive performance evaluation grade of the unmanned boat autonomous berthing system is determined to be "Excellent", the evaluation grade for the safety indicator A is "Excellent", the evaluation grade for the intelligence indicator B is "Medium", and the evaluation grade for the adaptability indicator C is "Medium". Further verification of the evaluation results has been carried out by consulting experts, and the evaluation conclusions have received unanimous approval from the experts. This demonstrates that the autonomous berthing system for unmanned boat performance evaluation indicator system and evaluation model constructed in this paper, based on the sequential relationship analysis method and FCE, are feasible.

V. CONCLUSION

This paper focuses on the autonomous berthing of unmanned boats and proposes a performance evaluation method of autonomous berthing system for unmanned boat. Firstly, a performance evaluation indicator system for the autonomous berthing system is established from three aspects: safety, intelligence, and adaptability, as well as the secondary indicators are quantified using both quantitative and qualitative methods. Then, based on the G1 method and FCE, a performance evaluation model for the autonomous berthing system is constructed, and the weights of each indicator at different levels are determined through expert scoring. Finally, the proposed method is validated through a case study using real test data from an unmanned boat's berthing test. The experimental results show that the proposed method can effectively and accurately reflect the performance differences of different unmanned boat autonomous berthing systems, and to some extent, fills the gap in the current research on performance evaluation of unmanned boat autonomous berthing systems.

ACKNOWLEDGMENT

The authors would like to thank the editors and anonymous reviewers for their valuable comments and suggestions to improve the presentation of this paper. This work was supported by the National Natural Science Foundation of China [Grant No.52231014] and Team Project of Dalian Maritime University [Grant No.3132023511].

REFERENCES

[1] Fan Y S, Su H, Wang G F. Testing technology and its application of autonomous navigation ability for unmanned surface vehicle[J]. Journal of Dalian Maritime University, 2020,46(03): 38-49.

[2] Su H. Research on Test and Evaluation Technology of Autonomous Navigation Ability of Unmanned Ship [D]. Dalian Maritime University, 2021.

[3] Zhang X Y, Zhang Y J. Research on the Methods Evaluating the Performance of the Autonomous Berthing System of Intelligent Ships[C]. 2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2022: 164-169.

[4] Fan Q Q. A Test and Evaluation Method for Ship Autonomous Collision Avoidance Performance in Complex Scenarios[D]. Dalian Maritime University, 2023.

[5] Li C X, Zhao J, Ding G G, et al. The Study of Risk Assessment Method for Ship Berthing Based on the "Human-Ship-Environment" Synergy[J]. Journal of Marine Science and Engineering, 2024, 12(11): 2022.

[6] Zhang H Z, Zhang Y J, Lu H R, et al. Research on the Decision-Making and Control System Architecture for Autonomous Berthing of MASS[J]. Journal of Marine Science and Engineering, 2024, 12(12): 2293.

Test Method for INS/GNSS Receivers

Anatoly Frolov
FSUE «VNIIFTRI»
Moscow region, Mendeleevo, Russia
frolov_aa@vniiftri.ru

Vladimir Pudlovskiy
FSUE «VNIIFTRI»
Moscow region, Mendeleevo, Russia
FSFEI HE «NRU «MPEI»,
Moscow,Russia
pudlovskiy@vniiftri.ru

Igor Silvestrov
FSUE «VNIIFTRI»
Moscow region, Mendeleevo, Russia
igsilv@vniiftri.ru

Abstract — The purpose of the work is to study a new test method for INS/GNSS receivers. The method is based on the reproduction of the reference motion trajectory by a three-axis rotary table with the provision for the tested INS/GNSS receivers to receive real GNSS signals. Estimates of errors in reproducing the parameters of the reference trajectory formed by the table (location coordinates, INS/GNSS-based Attitude Determination) are specified

Keywords — INS/GNSS receiver, tests, three-axis rotary table, radome

I. INTRODUCTION

Inertial navigation systems (INS) have been widely used for more than half a century to calculate the parameters of motion trajectories and GNSS-based attitude determination [1, 2]. To increase the accuracy of calculating motion trajectory parameters, INS are integrated with GNSS receivers [3, 4].

Tests of GNSS receivers in order to confirm its instrumental errors are usually carried out using GNSS-simulators [5].

As a rule, testing laboratories on wheeled vehicles are used for testing INS/GNSS receivers. The most common method is direct comparison with RTK GNSS receiver, used as a reference for full-scale tests with non-repeatable trajectories. The main disadvantage of this method is that the tested RTK GNSS receiver and INS/GNSS receiver are equally accurate measuring instruments. Various mechanical displacement tables (one-, two- and more axis, centrifuges, etc.) are widely used for testing sensitive elements of INS and INS as a complete technical product [7, 8].

The purpose of the work is to study the new test method for INS/GNSS receivers, using a three-axis rotary table (further table), located in the navigation field of GNSS signals.

II. PROPOSED TEST METHOD FOR INS/GNSS RECEIVERS

A. Idea of proposed method

The proposed test method can be considered an analogue of the test method using mobile laboratories; however, it has the necessary metrological margin of accuracy for estimating navigation parameters (location coordinates, INS/GNSS - based Attitude Determination, etc.). The method allows testing of INS/GNSS receivers with monoblock constructions when elementwise verification of the components of the INS/GNSS receivers (INS, GNSS, computing section) is not available.

To implement the proposed method, a three-axis rotary table is placed in a radome to protect the table's construction elements from the adverse effects of climatic factors (primarily precipitation), (Figure 1). The radome additionally attenuates GNSS signals in the L frequency band by no more than 0.5 dB.

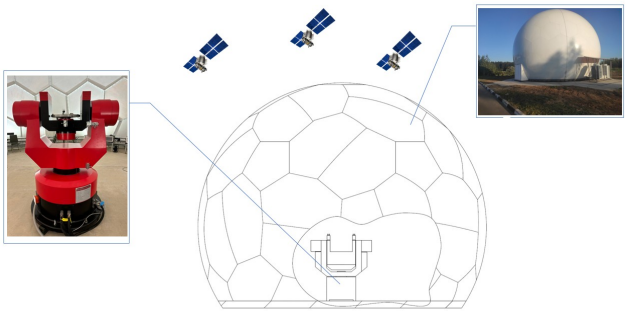


Fig. 1. Three-axis rotary table is placed in the navigation field of GNSS signals

The motion of the INS/GNSS receiver (reproduction of the reference trajectory) in space is carried out using angular movements of the axes of the three-axis table. At the same time, to increase the linear speed, an antenna of INS/GNSS receiver, as well as a block of sensitive inertial elements, is proposed to be placed on special technological mounting kit that ensures the removal of elements of the tested GNSS antenna from the intersection point of the table rotation axes, as shown in Figure 2.

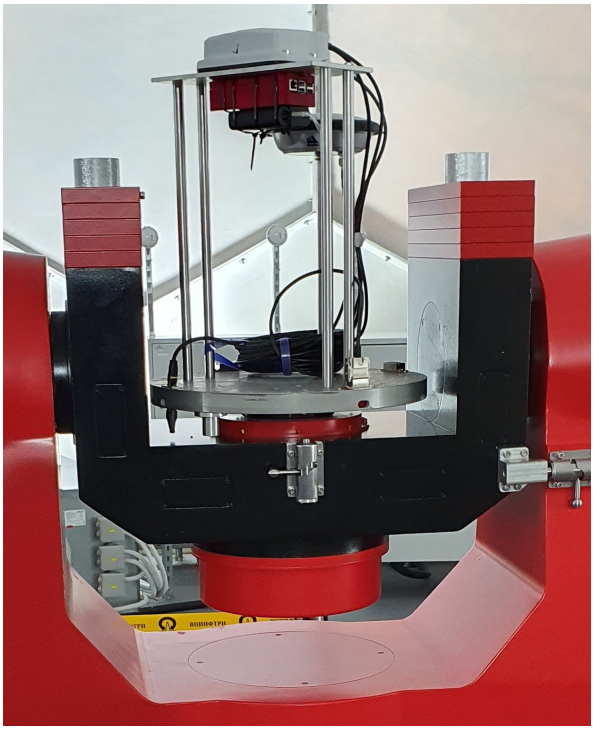


Fig. 2. INS/GNSS receiver structural elements placed on special technological mounting kit, installed on table top

Another variant of the technological mounting kit design may be a console (beam) fixed in the plane of the table top to ensure the displacement of the tested INS/GNSS receiver elements outside the table top, as shown in Figure 3.

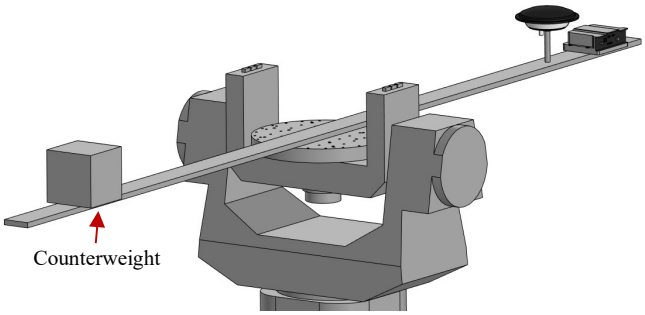


Fig. 3. Variant of technological mounting kit design for placing INS/GNSS receiver

However, the second variant of the technological mounting kit design does not allow the use of table top rotation, which somewhat limits the possibility of forming reference trajectories in space.

B. Using the table for INS/GNSS receiver test

To assess the errors of INS/GNSS receiver, using a table it is necessary:

- install the INS/GNSS receiver on the technological mounting kit rigidly connected to the table top;

- connect the INS/GNSS receiver to the power connectors and interface cable;

- determine the initial conditions for calculating the actual values of INS/GNSS receiver motion trajectory parameters;

- generate and reproduce single commands (or program) for angular movements of the table axes;

- calculate the differences between the measured INS/GNSS receiver navigation data and the values of the reference trajectory parameters.

The initial conditions for calculating the current values of the parameters of the reference trajectory for INS/GNSS receiver are:

- location coordinates of a given point of INS/GNSS receiver construction in the Earth-centered Earth Fixed (ECEF), for which reference values of the INS/GNSS receiver motion parameters will be calculated;

- the initial values of the INS/GNSS receiver attitude determination (course, roll, pitch). The initial values of INS/GNSS receiver attitude determination should be determined by measuring instruments according to the measurement procedure [9].

For example, the phase center of the GNSS antenna or the origin of the instrument coordinate system set by the accelerometer sensitivity axes (determined by the developer of the INS/GNSS receiver) can be selected as a given point of INS/GNSS receiver construction.

After reproducing the commands of the angular displacements of the table axes, the increments of the location coordinates of a given point of the INS/GNSS receiver construction in the topocentric coordinate system (TCS) ENU are calculated relative to the origin of the local coordinate system 0xyz (LCS), connected to the center of the table top, as shown in Figure 4.

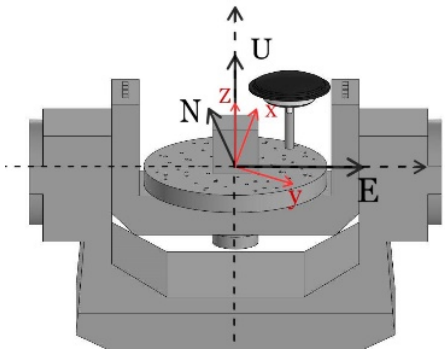


Fig. 4. INS/GNSS receiver location in the local coordinate system of table

Taking into account the known orientation of the LCS relative to the axes of the TCS, the coordinates of the location and the INS/GNSS receiver attitude determination in the LCS can be recalculated in the ECEF.

Thus, using the data on the current values of the rotation angles of the three axes of the table, it is possible to calculate the coordinates of the location in the ECEF and the orientation angles in the INS/GNSS receiver space, as well as derivatives of the specified navigation parameters.

To calculate the parameters of the reference trajectory reproduced by the table, it is necessary to know the current value of the attitude determination angles of the LCS relative to the local TCS ENU. By analogy with the known expressions for the orientation matrix of the platform trihedron of the table relative to the base trihedron [10], we can write:

$$\mathbf{R}_M = \mathbf{L}_3 \cdot \mathbf{L}_2 \cdot \mathbf{L}_1, \quad (1)$$

$$\mathbf{L}_1 = \begin{bmatrix} \cos\varphi_1 & -\sin\varphi_1 & 0 \\ \sin\varphi_1 & \cos\varphi_1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{L}_2 = \begin{bmatrix} \cos\varphi_2 & 0 & -\sin\varphi_2 \\ 0 & 1 & 0 \\ \sin\varphi_2 & 0 & \cos\varphi_2 \end{bmatrix},$$

$$\mathbf{L}_3 = \begin{bmatrix} \cos\varphi_3 & -\sin\varphi_3 & 0 \\ \sin\varphi_3 & \cos\varphi_3 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where $\mathbf{L}_k$ - rotation matrix of the platform trihedron of the table relative to the k-axis of the table at the appropriate angle φ_k.

Then the current position of the support construction vector in the LCS $\mathbf{D}_0$ is converted to the TCS as a vector $\mathbf{D}_{ENU}$:

$$\mathbf{D}_{ENU} = \mathbf{R}_M \cdot \mathbf{D}_0. \quad (2)$$

The angles of attitude determination (α – course angle; β – roll angle; γ – pitch angle) can now be calculated taking into account (1) as [11]:

$$\alpha = arctg\left(\frac{r_{21}}{r_{11}}\right),\ \gamma = arctg\left(\frac{-r_{31}}{\sqrt{r_{32}^2 + r_{33}^2}}\right),\ \beta = arctg\left(\frac{r_{32}}{r_{33}}\right), \quad (3)$$

where r_{ij} - the corresponding elements of the matrix $\mathbf{R}_M$.

The current values of coordinates and angles of INS/GNSS receiver attitude determination obtained this way in the TCS can be recalculated in the ECEF using well-known algorithms [12].

C. *Experimental verification of algorithms for calculating the actual values of INS orientation angles*

Verification of the correctness of the parameter formation of the reference trajectory in the proposed method of testing INS/GNSS receiver was carried out experimentally. In the experiments, INS/GNSS receiver of type NovAtel PwrPak7 [13] and a model of INS on the basis of micromechanical meters were used. The placement of these navigation devices on the technological mounting kit of the three–axis table is shown in Figure 2 (INS/GNSS receiver is on top of the platform, INS model is on the bottom). The initial values of the INS/GNSS receiver attitude determination angles were determined by the measurement procedure [9]. INS/GNSS receiver measurements were used to compare the orientation angles calculated from the readings of the axes of the three-axis table.

Angular movements along the axes of the table with speeds up to 5°/c were set in the interface software of the table. Diagrams of angle changes of table rotation in time are shown in Figure 5.

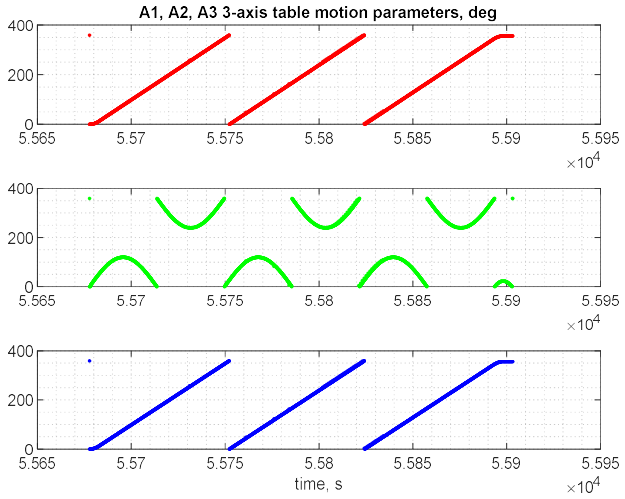


Fig. 5. Angle changes of table axis rotation in time

As a result of simultaneous rotations of the table axes, a reference trajectory was formed for the technological mounting kit. The view of the reference trajectory in the form of an increment of rectangular coordinates in the GCS is shown in Figure 6.

The current values of the angles of INS/GNSS receiver attitude determination and calculated from the readings of the axes of the three-axis table are shown in Figure 7.

The difference between the values of INS/GNSS receiver attitude determination and from the table did not exceed 1°, which generally corresponds to the error limits of this type of INS/GNSS receiver [13].

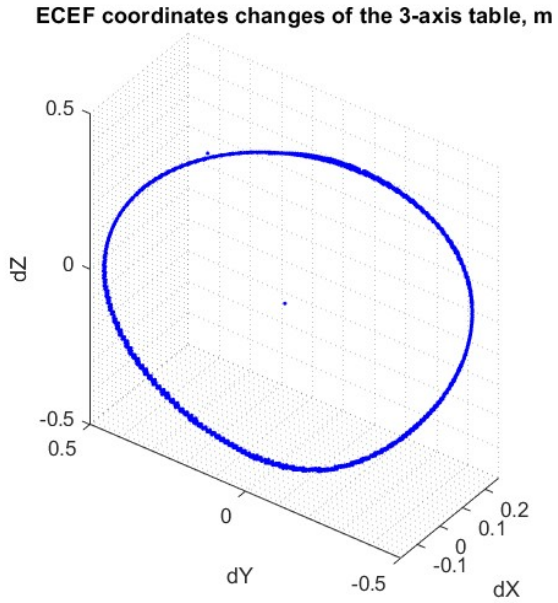


Fig. 6. Type of reference trajectory in space

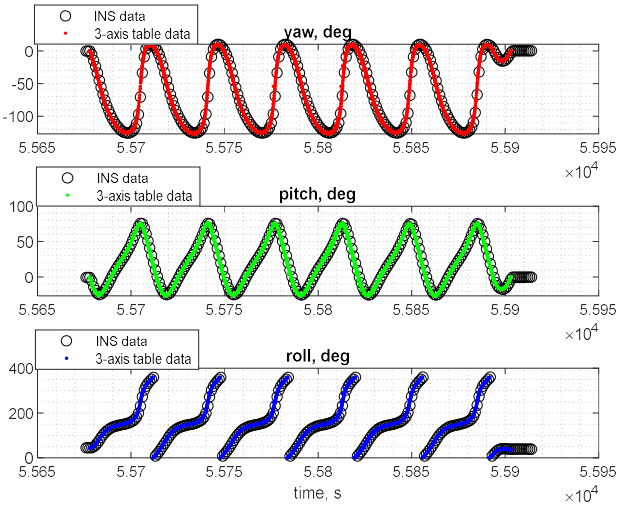


Fig. 7. Values of INS/GNSS receiver attitude determination and obtained from the angles of rotation of the table axes

III. ESTIMATION OF ERRORS IN REPRODUCING THE PARAMETERS OF THE REFERENCE TRAJECTORY

The error limits for the formation of the location coordinates of the reference trajectory in the ECEF and the angles of the INS/GNSS receiver attitude determination in space within a radius of 1 m from the center of the faceplate were obtained using the Monte-Carlo method in accordance with GOST 34100.3.2-2017.

Initial data for calculations:

- limits of the permissible error in reproducing angular displacement along each axis are ± 5″ [7];

- limits of the permissible error in determining the coordinates of a given point of INS/GNSS receiver construction in ECEF 0.01 m;
- limits of the permissible error in determining the initial values of the INS/GNSS receiver attitude determination angles ±0.5'.
- law of distribution of the current values of angles rotation errors of the table axes is uniform.

As a result of processing the simulated data array (100 reference trajectories from 14400 points each), the following values of errors in the formation of navigation parameters were obtained: for coordinates in the ECEF ± 0.03 m; for attitude determination angles ± 0.6'.

CONCLUSION

A new semi-natural method for testing INS/GNSS receiver is proposed. The method is based on the use of a three-axis rotary table located in the GNSS navigation field.

The method makes it possible to reproduce in the GNSS navigation field the repeating trajectories of linear-angular movements of the technological mounting kit support platform with a fixed INS/GNSS receiver with a total weight of up to 15 kg.

To implement the method, a mathematical apparatus for calculating the parameters of the reference (standard) trajectory has been developed.

Estimates of the error limits of the reference trajectory parameters reproduced by the table have been obtained to determine the measurement errors of the trajectory parameters obtained from the tested INS/GNSS receiver.

The method can be used to study promising methods for closely integrating measurements from inertial sensors and tracking circuits for GNSS signal parameters for the following reasons: the natural synchronization of measurements of heterogeneous sensors is ensured and the error in reproducing motion trajectory parameters in the geocentric coordinate system at the level of units of centimeters is guaranteed.

REFERENCES

[1] Lawrence A. Modern Inertial Technology. Navigation, Guidance and Control / SpringerVerlag, 1993. – 259 p.

[2] Titterton D. Strapdown Inertial Navigation Technology / D. Titterton, J. Weston. - 2 nd Edition. – Institution of Electrical Engineers, UK, 2004. – 558 p.

[3] GLONASS. Modernization and development prospects; edited by A.E. Perova. M.: Radiotechnics, 2020. p. 1072

[4] Kalchenko A.O. The task of calibrating a cardless inertial navigation system in flight using information from a satellite navigation system: dis. ... Candidate of Physical and Mathematical Sciences. M., 2016. p. 105

[5] Denisenko O.V., Silvestrov I.S., Mazurkevich A.V., Golub D.A., Pecheritsa D.S., Frolov A.A. Changes in the updated State verification schedule for coordinate-temporal measuring instruments // Almanac of Modern Metrology. 2019. No. 3 (19). pp. 25-30.

[6] A.S. Maltsev, S.Yu. Burtsev, D.S. Pecheritsa Investigation of the influence of base station parameters on the error in determining relative location coordinates // Materials of the XII Scientific and Practical Conference of Young Scientists, postgraduates and specialists. 2024. pp. 54-58.

[7] Rotary table 3T-30. User manual.

[8] Inchagov Yu.M., Shatilov A.Yu., Nagin I.A. Procedure for estimating errors of inertial satellite navigation systems // News of TulGU. Technical sciences. 2015. Issue 11. part 2. pp. 197-208.

[9] The procedure of measuring the angles of attitude determination with electronic tacheometer (project).

[10] Vasineva I.A. Calibration of a cardless inertial navigation system assembled on precision tables: dis. ... Candidate of Physical and Mathematical Sciences, Moscow, 2017. 99 p.

[11] Robot Dynamics Lecture Notes [Electronic resource]. Режим доступа: https://ethz.ch/content/dam/ethz/special-interest/mavt/robotics-n-intelligent-systems/rsl-dam/documents/RobotDynamics2017/RD_HS2017script.pdf (Accessed: 02.03.2025).

[12] Parameters of the Earth 1990 (PZ-90.11): A specialized reference book – Moscow: VTU GSH VSR. 2020. p. 64.

[13] PwrPak7 Family Installation and Operation User Manual [Electronic resource]. Access mode: https://docs.novatel.com/OEM7/Content/PDFs/PwrPak7_Installation_Operation_Manual.pdf (Accessed: 03/02/2025).

General Approach to Methods of Flight Tests for Airborne Flight and Navigation Equipment

E. G. Kharin

JSC "Gromov Flight Research Institute"
Zhukovsky, Russia
e-mail nio9@lii.ru

V. A. Kopelovich

JSC "Gromov Flight Research Institute"
Zhukovsky, Russia
e-mail nio9@lii.ru

A. V. Yasenok

JSC "Gromov Flight Research Institute"
Zhukovsky, Russia
e-mail nio9@lii.ru

E. B. Gorskiy

JSC "Gromov Flight Research Institute"
Zhukovsky, Russia
e-mail nio9@lii.ru

***Abstract*—This paper presents modern methods and tools for flight testing of aircraft flight and navigation equipment (FNE). The role of JSC "Gromov Flight Research Institute" in the development of FNE, as well as in the creation, development, and implementation of flight research and testing technologies, is highlighted. A developed general approach to FNE flight testing methods based on the use of onboard trajectory measurement systems (OTMS) and satellite navigation systems is discussed. Examples of applying an integrated flight test support system for FNE are provided.**

Keywords— aircraft, flight research and testing, flight and navigation equipment, satellite navigation systems.

I. INTRODUCTION

Flight and navigation equipment (FNE) of aircraft is continuously being improved to meet the growing demands for flight regularity, safety, and cost-efficiency. The scope of tasks assigned to FNE and the requirements for their execution are constantly increasing.

The FNE of an aircraft includes a large number of different systems:

- air data systems for measuring flight parameters;
- inertial navigation systems (INS);
- radio navigation, landing, and air traffic control equipment;
- satellite navigation systems (SNS);
- automatic and automated flight control systems;
- display and alerting systems.

The layout of airborne FNE has evolved from a fragmented structure to a highly integrated one. The cost of developing aviation equipment and its software accounts for 60÷80 % of the aircraft's total cost.

Flight testing plays a crucial role in the development of new aircraft. Its significance lies in being the final stage of aviation development, where shortcomings from all previous stages – from design decisions during planning to production and manufacturing – are identified.

II. HISTORY OF FNE FLIGHT TESTING

Let us briefly outline some stages of FNE development and the role of the "Gromov Flight Research Institute" (hereinafter referred to as LII) in the creation, development, and implementation of flight testing and research technologies for FNE [1].

In 1959–1970, the first domestic systems "Polet", "Polet-1", and "Polet-1I" were developed and tested. These are flight and navigation complexes (FNC) for airliners, military-transport aircraft, and maneuverable aircraft (Il-62, An-22). LII took an active part in the development of the concept and principles of construction, flight research and testing of the FNC on testbed and prototype aircraft.

In 1970–1980, unified baseline complexes – BNK-1P, BPK-1P (implemented on the Yak-42 aircraft), BNK-2P, BPK-2P (implemented on the Il-86 aircraft) were developed. Under LII's lead, a series of research projects were completed, defining the concept, design, and structure of FNE, unifying FNE systems, and conducting flight research and testing.

In 1980–1990, a large number of different types of aircraft and the "Buran" spaceplane were tested and commissioned. The flight testing technology, methods, and tools developed by LII were successfully applied in the testing of airborne FNC and aircraft systems.

In 1990–2000, the Il-96-300, Tu-204, and others were tested and certified. Certification testing was supported by LII using developed methods and tools. Work was carried out to create an onboard trajectory measurement system (OTMS) and implement it in testing advanced civil and military aircraft. Significant efforts were made to integrate satellite technologies into FNE complexes and flight testing practices. During this period, state and qualification tests were conducted on LII's testbed and prototype aircraft for various satellite navigation user equipment, inertial systems, and other FNE systems.

From 2000 to the present, extensive work has been done to implement these technologies in aviation flight testing:

- a major research cycle was conducted to improve OTMS and software - for processing and analyzing flight test data (over fifty software packages were developed);
- the accuracy of trajectory measurements was enhanced through differential SNS modes, including phase-based methods;
- flight tests were performed on LII's testbed aircraft for domestic satellite navigation equipment (SNS-2, SNS-3, SN-3301, "Inter-A", "Leader", A-737, A-744,

"Avia-P", PPA-S/V, PNAP-PLA, etc.) and inertial navigation systems (BINS-85, BIMS-T, LINS-100RS, etc.);

- certification testing was ensured for aircraft such as the Il-96-300, Be-200, Tu-214, RRJ-95, MS-21-300, Il-114-300, and others.

III. Modern Flight Testing Technology

Improving the efficiency of flight testing is a complex organizational and technical challenge, the solution of which is vital for reducing aircraft deployment timelines and associated costs. Hundreds of test flights are conducted during flight testing, extending the overall aircraft development cycle. The temporal and financial aspects of preparing and conducting flight tests are determined, among other factors, by the expenses of creating and using necessary methodological and technical support, including: measurement tools, test programs and methodologies, and software for processing and analyzing data obtained during test flights.

Flight testing is preceded by thorough preparation, which includes:

- analysis of regulatory and technical documentation (RTD) – technical specifications (TS), FNE structure, composition, characteristics, algorithms, airworthiness standards (AWS), flight manuals (FM), etc.;
- analysis of FNE performance requirements – correct functioning, accuracy, ergonomic indicators, fault tolerance, operational characteristics, etc.;
- selection of testing methods for evaluating FNE and its systems;
- selection of onboard and trajectory measurement tools, defining the list of recorded parameters and evaluation methods;
- selection of trajectory measurement methods.

Based on this analysis and the selection of methods and tools, ground and flight test methodologies and programs for FNE are developed.

For civil aircraft, methodologies may be based on previously developed conformity assessment methods and standard methodologies used for FNE testing. For state aviation aircraft, they are based on aviation testing guidelines and standard methodologies. The trajectory measurement system and software for data processing are adapted to the specific aircraft and FNE.

The full deployment of SNS by the mid-1990s and the emergence of high-precision equipment created the prerequisites for satellite-based trajectory measurements. Drawing on years of FNE evaluation experience, LII has established a unified testing cycle that records FNE data, performs trajectory measurements, and processes flight data [1]. LII has developed and implemented an integrated flight test support system for all types of modern, upgraded, and future aircraft and onboard equipment (Fig. 1). This system is designed to evaluate and analyze the characteristics of flight and navigation and other onboard equipment during flight testing, as well as determine flight performance.

The integrated flight test support system includes:

- an onboard trajectory measurement system (OTMS), comprising an onboard unit, base control stations (BCS), including a mobile station (MBCS), and other devices;
- testbed aircraft;
- test benches;
- methodological support;
- a computing network, automated workstations with software for processing and analyzing flight test data.

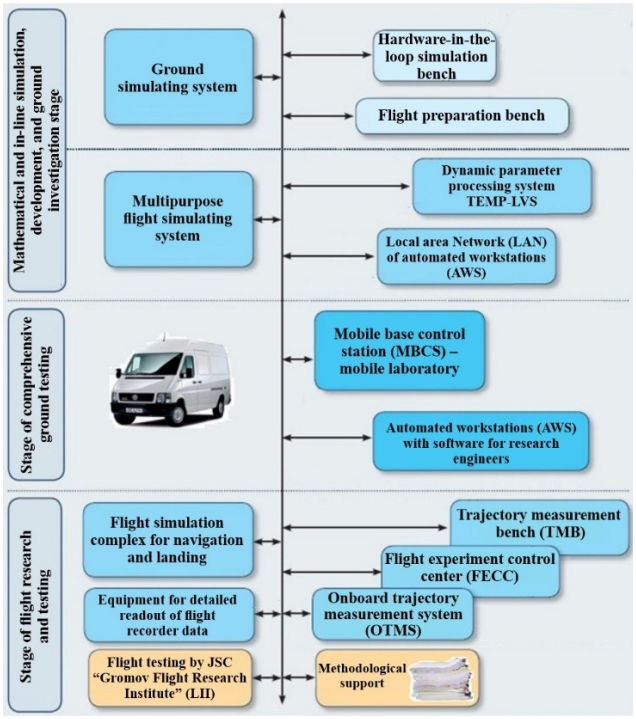


Fig. 1. Integrated aircraft flight test support system for aircraft

Methodological support includes:

- regulatory documents;
- aviation testing guidelines;
- ground and flight test methodologies and programs;
- conformity assessment methods.

The OTMS (Fig. 2) was created to evaluate and analyze FNE characteristics in real time during flight and in post-flight processing.

The OTMS onboard unit with an SNS antenna is installed onboard an aircraft, interfacing with the onboard equipment being tested via serial and multiplex lines. For real-time data analysis, a control and monitoring unit is installed for the senior engineer.

As the OTMS evolved, high-precision methods for processing satellite data were implemented. A stationary or mobile base control station is deployed in the flight zone to ensure differential mode operation. High-precision trajectory parameters are obtained post-flight by jointly processing onboard and ground data.

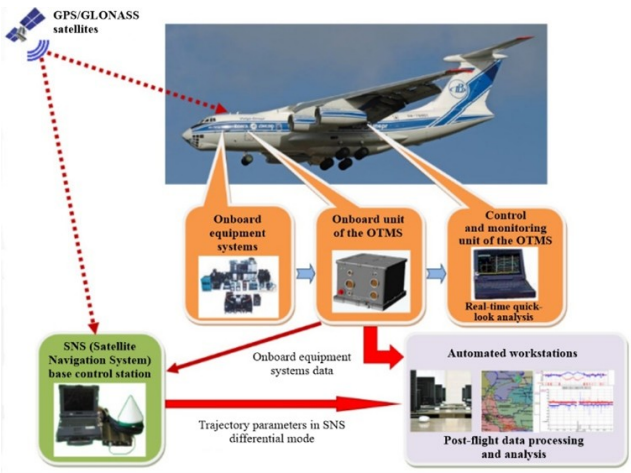


Fig. 2. OTMS application diagram in flight testing

The technological cycle for testing and evaluating aircraft and FNE systems performance using OTMS is shown in Fig. 3.

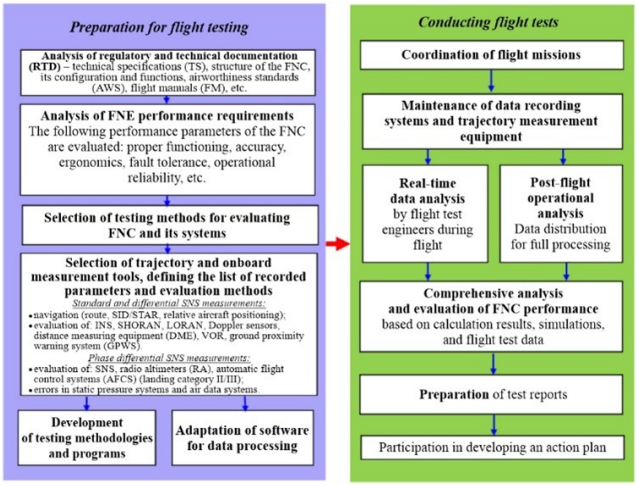


Fig. 3. Technological cycle for testing and evaluating FNE systems performance using OTMS

The implementation of OTMS has enabled a qualitative leap in evaluating various FNE systems:

- inertial navigation systems [2];
- radio navigation and landing equipment (RNLE) [3];
- satellite navigation systems [4];
- flight management system and others.

In addition to flight and navigation system testing, LII has developed and implemented the following flight testing technologies:

- analysis of automatic flight control system functioning and performance, including Category IIIA automatic landing;
- determination of relative positions of moving objects (used for testing aircraft carrier landings and aerial refueling) [5];
- evaluation of takeoff and landing performance;
- determination of aerodynamic errors in air data sensors, among others.

Below are examples of flight tests using the integrated flight test support system for FNE.

IV. Evaluation of Area Navigation and Approach Procedures

The inclusion of specialized navigation computers in FNE has significantly improved navigation accuracy, expanding the range of tasks that can be performed. Modern requirements for area navigation (RNAV) methods involve transitioning from sensor-based navigation to performance-based navigation.

The application of RNAV methods, based on radio technical (SNS, DME, VOR) and onboard (INS, platformless INS) systems, has enhanced the operational efficiency of existing airspace. European methods for evaluating approach procedures use precision radar, so initial certification flights had to be conducted in Finland. The use of OTMS has enabled the evaluation of total navigation error, separate navigation error, and piloting error at nearly all flight stages:

- en-route flight;
- execution of SID and STAR procedures (an example is shown in Fig. 4);
- RNP Approach procedures.

LII specialists participated in testing the Il-96-300, Tu 214, and RRJ-95 to evaluate RNAV performance. The resulting materials include system error assessments, piloting errors, deviations from prescribed procedures, and flight parameters during procedure execution.

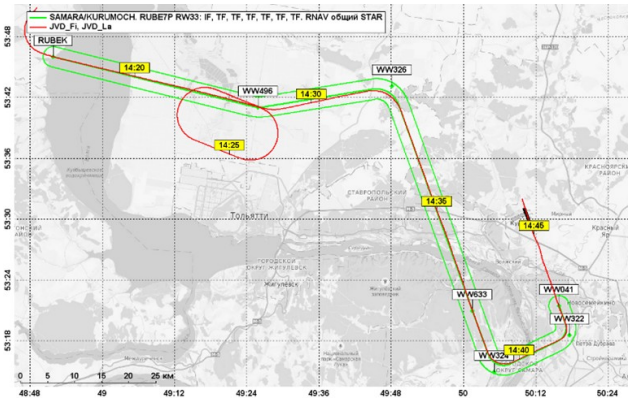


Fig. 4. Example of evaluating a STAR approach procedure

V. Evaluation of Inertial Navigation Systems

One of the primary systems ensuring aircraft navigation and piloting tasks is the inertial navigation system (INS). LII specialists have developed the “Etalon” software package, which performs:

- assessment of INS navigation and piloting parameter errors in autonomous mode;
- evaluation of instrumental errors in INS sensitive elements;

- data processing with integration of redundant navigation information using a Kalman filter for forward and backward smoothing.

The flight study and testing methods, tools, and methodologies developed at LII, based on satellite and computer technologies, have enabled qualification and certification testing of inertial and inertial -satellite navigation systems from various developers. Tests were conducted for systems such as BINS-SP-2, BINS-SP-2M, BIMS-T, INS-2000-4, NSI-2000, BINS-1DT-01, SNV-1A, BINS-RT-1, and LINS-100RS on numerous aircraft types. An example of INS performance analysis is shown in Fig. 5 and 6.

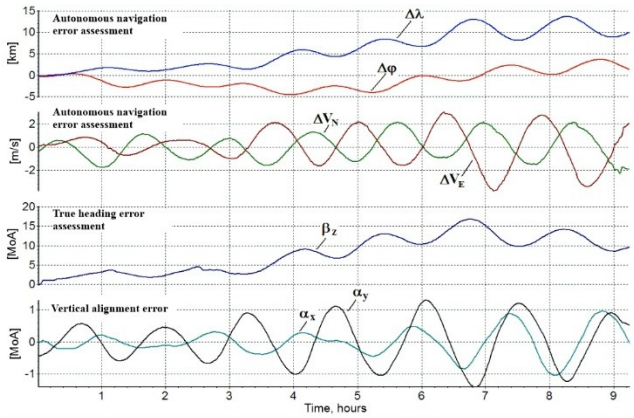


Fig. 5. Results of INS performance analysis based on data from one flight

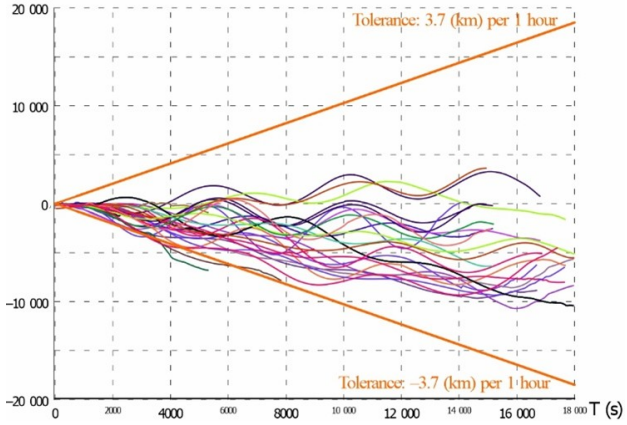


Fig. 6. Results of INS performance analysis based on data from multiple flights

VI. Evaluation of Radio Navigation and Landing Equipment

Modern aircraft FNE includes various types of onboard radio navigation and landing equipment:

- short-range radiotechnical navigation system (RSBN, VOR/DME system);
- instrument landing systems (ILS);
- satellite navigation and landing equipment;
- automatic direction finders (ADF);
- radio altimeters (RA) for low and high altitudes;
- Doppler speed and drift-angle sensors;
- long-range radio navigation system (LORAN).

RNLE performance is evaluated based on precise quantitative indicators, which are functions of aircraft motion parameters: spatial coordinates, velocity vector components, and angular position. Therefore, reference values for RNLE parameters are primarily determined using trajectory measurement data.

An example of radio navigation system performance analysis is shown in Fig. 7 and 8.

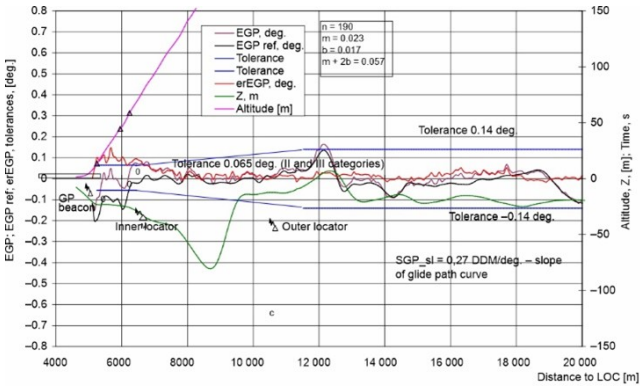


Fig. 7. Evaluation of ILS glide path deviation signals

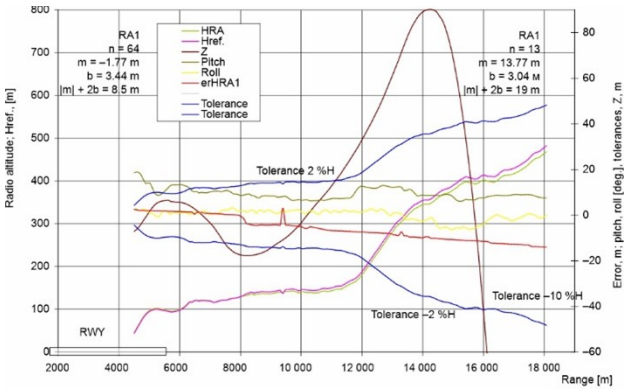


Fig. 8. Evaluation of radio altimeter measurements

Acknowledgment

Modern testing methods and tools, based on the methodological and software support developed at JSC “Gromov Flight Research Institute” enable prompt and objective analysis of FNE flight research and testing data.

As a result of ongoing improvement of flight testing methodology and technology proposed by LII, a consistent approach to FNE flight testing methods has been developed, and an integrated flight test support system has been created for advanced aircraft, based on the OTMS onboard trajectory measurement system.

The developed technologies have been successfully applied in testing maneuverable, passenger and transport aircraft including the An-70, An-140, An-148, An-225, Il-76 family, Il-96, Il-38, Il-114-300, Tu-204, Tu-214, Tu-334, Be-200, RRJ-95, Il 96 300, Il-114-300, SSJ-NEW, and MS-21-300.

References

[1] Kharin E. G., Kopylov I. A. Technology of Flight Testing for Aircraft Onboard Equipment Using the Onboard Trajectory Measurement System. M.: MAI-PRINT, 2012. 360 p.

[2] Kharin E. G., Kopelovich V. A., Kopylov I. A., Yasenok A. V., Klabukov E. V., Makarova A. Yu., Ilyin V. B., Yakushev V. A. Methods and Tools for Flight Testing of Inertial and Inertial Satellite Navigation Systems // Proceedings of MIEA. Navigation and Aircraft Control. 2015. Issue 14. Pp. 26–44.

[3] Gorsky E. B., Kopylov I. A., Kharin E. G., Kopelovich V. A., Yasenok A. V. Trajectory Measurements in Monitoring and Testing Ground and Onboard Landing Radio Equipment for Aircraft // XXVI St. Petersburg International Conference on Integrated Navigation Systems. 2019. Pp. 91–94.

[4] Ilyin V. B., Kharin E. G., Kopelovich V. A., Kopylov I. A., Yakushev A. F., Zhabin P. Yu., Tsarev V. M. Flight Research on the Performance of Onboard SNS Equipment When Operating with Various Global Satellite Navigation Systems // Navigation News. 2022. No. 1. Pp. 21–35.

[5] Yasenok A. V., Kharin E. G., Kopylov I. A., Kopelovich V. A., Yakushev A. F. Trajectory Measurements Based on Inertial Satellite Navigation Systems in Experimental Research on Aerial Refueling from the Il-96-300T Tanker // XXIII St. Petersburg International Conference on Integrated Navigation Systems. 2016. Pp. 234–240.

Adaptive Method for Estimation and Navigation of Unmanned Vehicles Using Intelligent Technologies

O.S. Amosov

Laboratory of Intellectual Control Systems and Modelling
V.A. Trapeznikov Institute of Control Sciences of RAS
Moscow, Russia
osa18@yandex.ru

S.G. Amosova

Laboratory of Cyber-Physical Systems
V.A. Trapeznikov Institute of Control Sciences of RAS
Moscow, Russia
amosovasg@yandex.ru

***Abstract*—An approach to solving the problem of adaptive estimation based on intelligent technologies is considered. The problem statement is formulated under structural and parametric uncertainty. Synthetic algorithms for its solution are discussed. Two illustrative examples are given for the adaptive estimation problem.**

Keywords—adaptive estimation, navigation, unmanned vehicle, neural network, fuzzy system, machine learning, intelligent technology, integration

I. INTRODUCTION

One of the actual current challenges in the navigation of unmanned vehicles today is to ensure their autonomy, stealth, and immunity to interference. This largely depends on the specifics of the onboard equipment. Unmanned vehicles are usually equipped with a standard set of devices and measuring instruments, so they are subject to typical navigation problems. For example, global navigation satellite system (SNS) signals may be distorted (multipath distortion) or unavailable. Regular calibration of inertial and other sensors is also required to maintain navigation accuracy by inertial navigation system (INS). Therefore, navigation methods and algorithms are needed that will help effectively combine data from various sensors (inertial measurement units, lidars, video cameras, etc.) to improve accuracy, taking into account limited computing power on board and energy resources, loss of communication, interference, and noise. Along with traditional methods and algorithms, neural network methods and machine learning are also used in the navigation of moving objects. Their main advantages are high speed and the ability to adapt to changing environmental properties. Fuzzy logic systems are effective for structural and parametric tuning of the systems.

Therefore, the aim of this paper is to present the adaptive method for the estimation and navigation of unmanned vehicles using intelligent technologies based on neural networks (NN), machine learning and fuzzy inference systems for decision making in the integration of measurements and system adaptation.

II. ESTIMATION PROBLEM STATEMENT AND SOLUTION

A. Types of Uncertainties

When processing navigation information, one must deal with the presence of a priori uncertainty, which means the incompleteness or unreliability of information, the influence of randomness or counteraction. In this case, both parametric and structural uncertainty are considered [1–4].

In the case of parametric uncertainty, the mathematical model describing the processes used is specified with an accuracy of up to unknown parameters and is conveyed by introducing a vector of accompanying, unknown parameters $\boldsymbol{\theta}(t) = (\theta_1(t), \ldots, \theta_l(t))^T$. In the case of structural uncertainty, both the mathematical model $\mathbf{M}(t)$ of the process itself and the parameters $\boldsymbol{\theta}(t)$ specifying it are unknown [1–4].

Adaptive estimation algorithms are algorithms for solving problems in the presence of structural and parametric uncertainty. We will consider adaptive estimation problems in discrete time.

B. The Problem Statement of Adaptive Estimation

It is necessary to estimate a $(n+l)$ -dimensional random sequence $\mathbf{z}_i = (\mathbf{x}_i^T, \boldsymbol{\theta}_i^T)^T$, $\mathbf{x}_i = (x_{1i}, \ldots, x_{ni})^T$, $\boldsymbol{\theta}_i = (\theta_{1i}, \ldots, \theta_{li})^T$ of a dynamical system by the set $\mathbf{Y}_i = [\mathbf{y}_1^T \ \ldots \ \mathbf{y}_{i-1}^T \ \mathbf{y}_i^T]^T$ of m-dimensional measurement vectors $\mathbf{y}_i = [y_{1i} \ \ldots \ y_{mi}]^T$, starting from the condition for minimization of the given criterion. Here $\mathbf{x}_i$ is the n-dimensional state vector and $\boldsymbol{\theta}_i$ is the l-dimensional vector of unknown accompanying parameters.

It is assumed that in the presence of structural uncertainty, the system is represented by a finite number of models M^j corresponding to a given set of hypotheses h^j :

$$M^j \to h^j, j = \overline{1,J}.$$

Let us consider the case when the equations of system dynamics and measurements are involved in the presence of parametric and structural uncertainties of their models. For the case of additive noises $\mathbf{n}_i$ and $\mathbf{v}_i$, the equations of dynamics and measurements can generally be written as follows [2]:

$$\mathbf{x}_i = \boldsymbol{\Phi}_i(\boldsymbol{\theta}_{i-1})\mathbf{x}_{i-1} + \boldsymbol{\Gamma}_i(\boldsymbol{\theta}_{i-1})\mathbf{n}_i; \quad (1)$$

$$\mathbf{y}_i = \mathbf{H}_i(\boldsymbol{\theta}_i)\mathbf{x}_i + \boldsymbol{\Psi}_i(\boldsymbol{\theta}_{i-1})\mathbf{v}_i; \quad (2)$$

$$\boldsymbol{\theta}_i = \boldsymbol{\theta}_{i-1} + \mathbf{n}_{\boldsymbol{\theta}i}. \quad (3)$$

For each hypothesis j , $j = \overline{1,J}$, in equations (1)–(3) $\boldsymbol{\Phi}_i$, $\boldsymbol{\Gamma}_i$, $\mathbf{H}_i$, and $\boldsymbol{\Psi}_i$ are known matrices; $\mathbf{n}_i$, $\mathbf{v}_i$, and

The research was supported by the Russian Science Foundation, project no. 24-29-00671, https://rscf.en/project/24-29-00671/.

$\mathbf{n}_{\boldsymbol{\theta}i}$ are white-noise centered Gaussian sequences of the corresponding dimensions.

C. The Solution of Adaptive Estimation Problem

There are two approaches to solving the problem. In the first group of adaptation algorithms, it is necessary to determine the number j of the hypothesis that maximizes a posteriori probability $P\{M^j / \mathbf{Y}_i\} \to h^j, j = \overline{1,J}$, and to find the optimal Bayesian estimates of the vectors $\mathbf{x}_i^j$, $\boldsymbol{\theta}_i^j$ corresponding to this number [2].

In another group of algorithms, a posteriori probabilities of all hypotheses are found, and the optimal Bayesian estimates of the vectors $\mathbf{x}_i^j$, $\boldsymbol{\theta}_i^j$ are found as weighted estimates of these vectors for all hypotheses [3–5].

To implement estimation algorithms, the most commonly used multi-alternative method is based on a Kalman filter (KF) bank [2–4]. This method uses a set of filters, each of which is configured for a specific hypothesis about the model describing the behavior of the process. Filter residuals are used to calculate the a posteriori probabilities of hypotheses about different states.

D. Solution of Adaptive Estimation Problem Based on Intelligent Technologies

An alternative approach uses a bank of so-called synthetic filters (algorithms) built on the basis of machine learning and intelligent technologies. These are neural network filters [6–8], filters based on fuzzy systems [9, 10], and wavelets [9, 10]. A synthetic algorithm is also used to calculate a posteriori probabilities of hypotheses about different states, which takes into account the filter residuals.

The difference between the approach using intelligent technologies and the well-known one based on the Kalman filter bank is that each of the synthetic filters configured using supervised learning can be nonlinear. A characteristic feature of this approach is the presence of a training set.

To solve the problem of estimating navigation parameters, it is proposed to use synthetic algorithms that provide the same accuracy of estimates as traditional algorithms in Bayesian, non-Bayesian approaches, and the least squares method [6–10]. Within the framework of the Bayesian approach, if the construction of synthetic algorithms is based on the same assumptions about a priori information about models of the object's dynamics, interference, and noise as for traditional algorithms, then synthetic algorithms are identical to traditional ones [7]. However, they may provide computational advantages. In addition, if there is a need to adapt to changes in dynamic models or the environment, it is possible to customize existing synthetic algorithms based on reinforcement learning within the framework of the least-squares method [8, 10].

When integrating measurements, it is possible to use the adaptive properties of synthetic algorithms. In these tasks, in addition to current measurements, there is a training sample for constructing synthetic algorithms. Such examples do exist for navigation applications. In particular, they are the problems of measurement processing in integrated inertial/satellite systems [7]. In these systems, with the availability of satellite measurements, they can be used as reference data for learning a synthetic which is intended to predict the behavior of parameters being estimated (navigation parameters themselves and sensor errors) in the case that the satellite data are missing [7]. To solve the estimation problems, non-recursive and recursive synthetic algorithms are analyzed. To improve the performance, it is proposed to use algorithm decomposition [9, 10].

To select the integration scheme and structural-parametric adjustment of the estimation system, it is proposed to use a fuzzy logic system for decision making. The solution to this problem requires the creation of a bank of estimation models and integration schemes depending on the availability of different measurement systems: onboard inertial navigation systems, sensors of the Earth's geophysical fields, beacons, satellite data, video systems, and other measurement tools.

To form a rule base for a fuzzy system, it is necessary to first determine the input and output linguistic variables [11]. Navigation systems record diagnostic parameters that allow the technical condition of the system to be predicted. Based on the values of the input diagnostic parameters $\mathbf{a} = a_1 \times ... \times a_p$, it is possible to assess the correct operation of the object or a disruption in its functioning process, i.e. the state of the navigation system. Based on the technical condition, a scheme for integration and structural-parametric adjustment of the estimation system is selected. The corresponding rule base consists of a set of fuzzy "*If-then*" rules in the following form:

$$R_k: \ If\ \mathbf{a}\ is\ \mathbf{A}_k, then\ \ \mathbf{b}\ is\ \mathbf{B}_k\ ;\ k = \overline{1.\lambda}, \qquad (4)$$

where $\mathbf{a}$ is input vector for fuzzy system, $\mathbf{A}_k = A_{k1} \times ... \times A_{kp}$ are fuzzy sets defined on the Cartesian product $\mathbf{A}$ of universal sets of input linguistic variables and having membership functions $\mu_{A_{ki}}(a_i)$, $k = \overline{1.\lambda}$, $i = \overline{1.p}$; $\mathbf{B}_k = B_{k1} \times ... \times B_{kq}$ are fuzzy sets defined on the Cartesian product $\mathbf{B}$ of universal sets of output linguistic variables and having membership functions $\mu_{B_{ki}}(b_i)$, $k = \overline{1.\lambda}$, $i = \overline{1.q}$; λ is total number of fuzzy "*If-then*" rules in the rule base.

To select options measured using fuzzy parameters, various methods are used: ranking of alternatives on a set of linguistic assessments; multi-criteria selection based on a fuzzy preference relation; fuzzy inference on a set of production rules. Note that the fuzzy decision making system can be used in an alternative approach for hypothesis selection or weighting of hypotheses when choosing mathematical models (1)–(3).

III. Examples Of Solving The Adaptive Estimation Problem

A. The task of filtering when integrating SNS and INS readings based on an invariant scheme

The simplest version of the problem is to refine the coordinates of the INS using the SNS data. An example is the problem of complex processing to refine the altitude of the aircraft using measurements from the SNS.

Following [12] we assume that there are measurements

$$y_i^I = X_i + \Delta y_i^I,\ y_i^{II} = X_i + \Delta y_i^{II}, \quad (5)$$

where $X_i, y_i^I, y_i^{II}, \Delta y_i^I, \Delta y_i^{II}$ are scalar sequences describing the behavior of the true values of the coordinate, the readings of the INS (*I=INS*), the satellite navigation system (*II=SNS*) and their errors. We will assume that the errors of the INS $\Delta y_i^I = \varepsilon_i^I + v_i^I$ are determined by the sum of the drift described by the Wiener process, the constant, and white-noise components, i.e.

$$x_i = x_{i-1} + \sigma_w^X w_i,\ x_0 = d_i = d_{i-1} = d, \varepsilon_i = d_i + x_i + v_i^I, \quad (6)$$

where x_i is the coordinate X_i drift; σ_w^X is the coefficient determining the root mean square (r.m.s.) forcing drift noise; d is the constant component of the error, v_i^I is the white noise of the measurements.

The errors of the SNS contain only the white noise component $\Delta y_i^{II} = v_i^{II}$. The task is to find an estimate of the drift x_i for coordinate X_i against a background of white noise. Following the invariant scheme, we can formulate the desired estimate as [12]

$$\tilde{X}_i = y_i^I - \tilde{\varepsilon}_i^I,\ \tilde{\varepsilon}_i^I = H_i^I \tilde{x}_i, \quad (7)$$

which we find using the linear Kalman filter.

We will assume that measurements from the SNS are received from 1 to 100 readings, from 101 to 200 readings become unavailable, and appear from 201 readings. An example of the implementation of measurements and the results of the solution of the estimation algorithm are shown in Fig. 1.

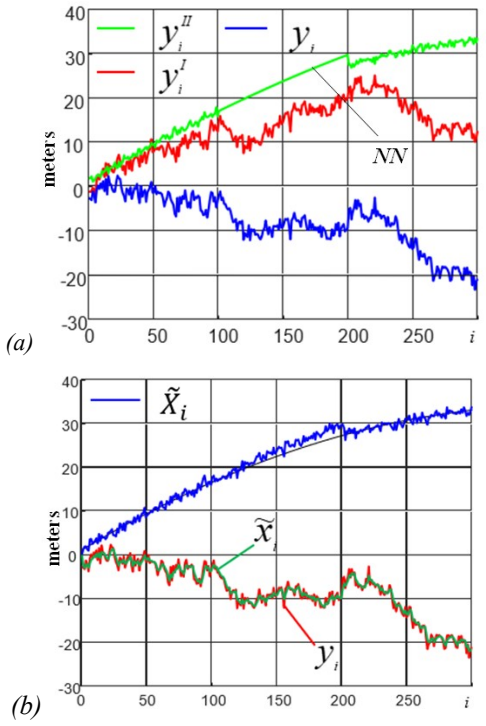


Fig. 1. An example of the implementation of (*a*) measurements and (*b*) the results of the solution of the estimation of coordinate and its drift.

The studies were conducted with the following parameters: sampling interval $T = t_i - t_{i-1} = 0.09\ s$; initial coordinate $X_0 = 1\ m$; initial velocity of the UV $V_0^X = 2\ m/s$; acceleration $a^X = -0.06\ m/s/s$; the constant component of the INS error is unknown, we assume that its mathematical expectation $m_X = 0\ m$ and r.m.s. deviation $\sigma^d = 2\ m$; the r.m.s. of discrete white noise of the forcing drift noise $\sigma_w^X = 0.5\ m$; the r.m.s. of discrete white noise of INS measurements $\sigma_v^I = 1\ m$; the r.m.s. of the discrete white noise of SNS measurements $\sigma_v^{II} = 0.5\ m$.

To predict the behavior of the UV based on the SNS data when the data from the SNS becomes unavailable, a two-layer neural network with one hidden layer with two neurons is used. The hidden layer of the neural network uses a sigmoid activation function, and the output layer uses a linear one. For complex models of UV behavior, it is planned to use a deep LSTM (Long Short-Term Memory) neural network to solve the problem of predicting time sequences [13].

B. The Problem of Filtering when Combining Gyroscope and Accelerometer Measurements Based on a Non-Invariant Scheme

To illustrate the application of the adaptive algorithm, an experiment was conducted with rotation around the axes of the unmanned vehicle at three angles: roll (*X* axis), pitch (*Y* axis) and yaw (*Z* axis). The measurement system consists of an inertial measurement unit based on the MPU-6050 integrated circuit, containing the 3-axis gyroscope and the 3-axis accelerometer. It is necessary to obtain estimates of the Euler angles for an unmanned vehicle based on measurements of the components of the angular velocity $[\omega_X\ \omega_Y\ \omega_Z]^T$ coming from the gyroscope and accelerations $[a_X\ a_Y\ a_Z]^T$ from the accelerometer. The estimated state vector for each of the angles contains the angle φ and angular velocity ω: $\mathbf{x}_i = [x_{1i}\ x_{2i}]^T = [\varphi_i\ \omega_i]^T$. Unknown accompanying parameter $\boldsymbol{\theta}_i = \theta_{1i} = [\Delta\omega_i]$ represents the gyroscope zero bias. $\mathbf{z}_i = [x_{1i}\ x_{2i}\ \theta_{1i}]^T = [\varphi_i\ \omega_i\ \Delta\omega_i]^T$.

Dynamics model, where $T = t_i - t_{i-1} = const$:

$$\varphi_i = \varphi_{i-1} + \omega_{i-1} \cdot T - \Delta\omega_{i-1} \cdot T + w_i^{\dot{\omega}} T^2 / 2; \quad (8)$$

$$\omega_i = \omega_{i-1} + \Delta\omega_{i-1} + w_i^{\dot{\omega}} T;\ \Delta\omega_i = \Delta\omega_{i-1} + w_i^{\Delta\omega}, \quad (9)$$

Measurement model:

$$\mathbf{y}_i = \begin{bmatrix} y_{1i} \\ y_{2i} \end{bmatrix} = \begin{bmatrix} \varphi_i \\ \omega_i \end{bmatrix} + \begin{bmatrix} v_{1i} \\ v_{2i} \end{bmatrix}. \quad (10)$$

To obtain a linear model of angle measurements φ, the primary measurements of the acceleration components from the accelerometer can be transformed and presented in the form: $y_{1i} = \text{atan2}(a_i^X, a_i^Z)$ for yaw and

$y_{1i} = \text{atan2}(a_i^X, a_i^Z)$ for pitch. The linear problem posed can be solved using either a Kalman filter or a neural network. In [6] it is shown that an algorithm using a recurrent linear neural network, with an appropriate choice of learning criterion, provides estimates that are close in their properties to the estimates produced by a linear Kalman filter. The a priori information in constructing the KF is provided by equations (1)–(3) and the joint probability density function (p.d.f.) $f(\mathbf{z}_0, \mathbf{z}_{i,}, \mathbf{y}_i)$. When constructing a neural network algorithm, equations (1)–(3) are also considered given, and instead of the p.d.f. $f(\mathbf{z}_0, \mathbf{z}_{i,}, \mathbf{y}_i)$, sets of realizations $\{\mathbf{z}_0^{(j)}, \mathbf{y}_1^{(j)}, \mathbf{z}_1^{(j)}, ..., \mathbf{y}_i^{(j)}, \mathbf{x}_i^{(j)}\}$, $j = \overline{1.N}$ are used, which are necessary for training the NN.

To solve the estimation problem, a linear single-layer NN with three neurons (according to the number of estimated parameters $x_{1i}, x_{2i}, \theta_{1i}$), with the identity activation function, three outputs for estimates $\tilde{x}_{1i}, \tilde{x}_{2i}, \tilde{\theta}_{1i}$, and five inputs $[\tilde{x}_{1i-1}, \tilde{x}_{2i-1}, \tilde{\theta}_{1i-1}, y_{1i}, y_{2i}]^T$ is used.

The studies were carried out at the following parameters: sampling interval $T = 0.09$ c; r.m.s. deviation of discrete white noise angular measurements $\sigma^{\varphi} = \sqrt{0.03}\ ^\circ$; r.m.s. of discrete white noise angular velocity measurements $\sigma^{\omega} = \sqrt{0.05}\ ^\circ/s$; r.m.s. deviation of the drift noise $\sigma^{\Delta\omega} = \sqrt{0.003}\ ^\circ/s$ degree/s, r.m.s. deviation forcing noise for the dynamics model $\sigma^{\dot{\omega}} = 0.3\ ^\circ/s/s$. Figure 2 show the results of solving the problem of estimating the angular position (roll) and error of the gyroscope.

Note that for the problem considered in the examples, the training sample can be obtained experimentally [14]. In addition, the training sample can be obtained by using computer models with real characteristics of sensors that are known.

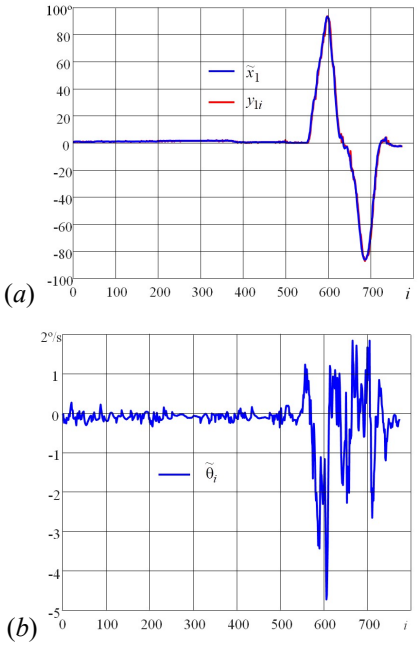


Fig. 2. Estimation (roll): (*a*) of angular position; (*b*) of error gyroscope.

IV. Conclusion

An approach to solving the problem of adaptive estimation based on intelligent technologies is presented. The problem statement is formulated under structural and parametric uncertainty. To solve it, synthetic algorithms based on machine learning and intelligent technologies are discussed. These are neural network filters, filters based on fuzzy systems and wavelets. To integrate measurements, a decision-making system based on a fuzzy logic system is proposed. Illustrative examples are considered for an unmanned vehicle. The task of adaptation in refining the coordinates of the INS using the SNS data is considered when satellite data are lost. A neural network is used to predict the behavior of a moving object. The task of adaptive estimation of Euler angles and errors of gyroscopes using a recurrent neural network algorithm is solved.

References

[1] V.I. Tikhonov, V.N. Kharisov, Statistical Analysis and Synthesis of Radio Engineering Devices and Systems. Moscow: Radio and Communications, 1991 (in Russian).

[2] O.A. Stepanov, A.V. Motorin, “Methods of Adaptive Estimation in Problems of Navigation Information Processing,” XIII All-Russian Meeting on management problems (VSPU-2019), Moscow, June 17-20, 2019. pp. 1359–1366 (in Russian).

[3] S.Z. Kuzmin, Digital Radar. Introduction to Theory. Kyiv: Computer and Publishing Information Center, 2000 (in Russian).

[4] Y. Bar-Shalom, X.-Rong Li, and Kirubarajan Thiagalingam, Estimation with Applications to Tracking and Navigation, New York: John Wiley & Sons, 2001.

[5] O.S. Amosov, S.G. Amosova, “Adaptive Estimation of the Processes Having Disorders in Navigation Applications Using Machine Learning,” 28th Saint Petersburg International Conference on Integrated Navigation Systems, 2021, vol. 28, pp. 9470814.

[6] O.A. Stepanov, O.S. Amosov, “Optimal Linear Filtering Using a Neural Network,” Giroskopiya i Navigatsiya, 2004, no. 3 (46), pp. 14–29 (in Russian).

[7] O.A. Stepanov, O.S. Amosov, “Application of Neural Networks to Nonlinear Navigation Problems,” 13th Saint Petersburg International Conference on Integrated Navigation Systems, 2006, vol. 13, pp. 207–210.

[8] O.S. Amosov, S.G. Amosova, “Machine Learning with Reinforcement for Optimal and Adaptive Estimation Problems in Navigation Applications,” 29th Saint Petersburg International Conference on Integrated Navigation Systems, 30 May – 1 June 2022.

[9] O.S. Amosov, S.G. Baena, “Decomposition Synthetic Approach for Optimum Nonlinear Estimation,” IFAC-PapersOnLine, 2015, T. 48, no. 11, pp. 819–824.

[10] O.S. Amosov, S.G. Amosova, “Intelligent Technologies for Joint Navigation and Functioning of Mobile Objects in Different Physical Environments,” Informacionnye Tehnologii, 2025, no. 1. vol. 31, pp. 24–34 (in Russian).

[11] O.S. Amosov, S.G. Baena, Y.S. Ivanov, S. Htike, “Roadway Gate Automatic Control System with the Use of Fuzzy Inference and Computer Vision Technologies,” 12th IEEE Conference on Industrial Electronics and Applications, 2018, vol. 12, pp. 706-711.

[12] O.A. Stepanov, Fundamentals of the Theory of Estimation with Applications to Navigation Data Processing, Vol. 1: Introduction to the Theory of Estimation. St. Peterburg: CSRI Elektropribor, 2017 (in Russian).

[13] O.S. Amosov, S.G. Amosova, “Predicting of a Person's Position in Trajectory Tracking from a Continuous Video Stream,” X International Annual Conference “Industrial Technologies and Engineering” (ICITE 2023), 2024, pp. 02022.

[14] S.O.H. Madgwick S.O.H. AHRS algorithms and calibration solutions to facilitate new applications using low-cost MEMS. University of Bristol, 2014.

About a Method of AUV Group Positioning using the Hydro Acoustical Navigation System with a Short Base Line

A.F. Scherbatyuk

IACP FEB RAS

Vladivostok, Russia
alex-scherba@yandex.ru

***Abstract*— The report considers the task of providing navigation support for a group of autonomous underwater vehicles (AUV) based on the use of a hydroacoustic navigation system (HANS) with a short base line (SBL). SBL HANS is proposed, in which the processing unit for generating time delay measurements of received signals on the AUV is complemented by a function for measuring the frequency of the received signals. The proposed estimation algorithm based on a particle filter is described, as well as some results of computer modeling of its operation.**

Keywords- hydroacoustic navigation system with a short base line, group navigation, autonomous underwater vehicle

I. INTRODUCTION

One of the ways to solve the problem of determining the coordinates for a group of autonomous underwater vehicles is based on the use of a hydroacoustic navigation system with a short base line. SBL HANS includes a separate beacon that periodically emits an acoustic signal, and several (at least two) receiving antennas on board each AUV.

Consider the case when there are two receiving antennas located along the AUV body. They record the moments (*t1, t2*) of signals receiving from the beacon in each cycle of its operation. The time interval $\Delta t12 = t1 - t2$ is proportional to the distance difference between the receiving antennas and the beacon. The distance difference is related to the bearing θ to the beacon and the value of the base distance a between the receiving antennas of SBL HANS on AUV by:

$$\Delta t_{12} = (a*\cos\theta)/c, \quad (1)$$

where c is the speed of propagation for the sound signal in the water, θ is the angle between the line directed along the hull of the AUV and the line directed to the beacon [1].

In [2], a method for estimating the coordinates of the AUV based on the SBL HANS using the AUV motion model is described. It assumes that the AUV has two receiving antennas of the SBL HANS, heading and depth sensors on board, and there are no expensive high-precision time system and an AUV absolute velocity sensor. Taking into account the fact that the vertical coordinate is measured with the required accuracy, a two-dimensional problem is considered when the AUV moves uniformly and rectilinearly in a horizontal plane on a certain section of the trajectory. An algorithm based on a particle filter is used to solve this problem.

Due to the lack for synchronization of the AUV and the beacon clock, it is not possible to directly determine the location of the AUV using the SBL HANS, since it is impossible to calculate the distance between the AUV and the beacon. Due to the lack of a velocity sensor, the area of uncertainty for the AUV location is large and the process of estimating the horizontal coordinates (x, y) and velocity v at the initial stage requires a significant number of iterations in order for the parameter estimates to achieve the true values. During this time, the location of the AUV remains uncertain. The need to evaluate three parameters – horizontal coordinates (x, y) and velocity v using the particle filter imposes on AUV significant limitations with the performance of the on-board computer.

In this paper, a version of the SBL HANS is proposed, in which the received signal processing unit for generating measurements is supplemented by the function of measuring the frequency of the received signals. The implementation of this function in the processing unit is associated with a minimal upgrade, since the main hardware modules necessary for this are already available in the device and it is only necessary to modify the input signal processing algorithm. The description of estimation algorithm based on a particle filter and some results for computer modeling of its operation are supplemented.

II. DESCRIPTION OF THE ALGORITHM

Consider the coordinates estimation of one AUV from a group. Taking into account the fact that the vertical coordinate (depth) is measured with the required accuracy, for simplicity of presentation, assume that the AUV moves in a horizontal plane rectilinearly and uniformly. Without generality limitation, it can be assumed that the beacon is located at a fixed point with coordinates (x_a, y_a, z_a). It is assumed that the AUV and beacon clocks are not synchronized, but the period τ of navigation signals radiation by the beacon is known.

The following measurements are available in the k-th cycle of the SBL HANS operation:

- the difference in the time of the signal arrival to the first and second receivers of the SBL HANS on the AUV:

$$\Delta t_{12k} = (d_{1k} - d_{2k})/c, \quad (2)$$

where d_{1k} and d_{2k} are the distances from the beacon to the first and second receiver of the SBL HANS on the AUV in the k-th cycle of the SBL HANS, respectively;

- the difference $\Delta t_{k,k-1}$ of the time for arrival of the signal in (k -1)-th and k-th cycles to the first receiver SBL HANS on the AUV:

$$t_k = t_{sk} + d_{1k}/c, \quad t_{k-1} = t_{sk-1} + d_{1k-1}/c, \quad t_{sk} - t_{sk-1} = \tau, \quad (3)$$

where t_{sk} and t_k are the times of emission and reception of the signal in the k-th cycle, t_{sk-1} and t_{k-1} are the times of emission and reception in the $(k-1)$-th cycle, from where by subtract we have:

$$\Delta t_{k,k-1} = (d_{1k} - d_{1k-1})/c + \tau. \quad (4)$$

- a frequency of the received signal f_{ok}.

According to the Doppler equation, we have:

$$f_{0k} = (1 \pm (v_k/c)\sin\theta) * f_s \quad (5)$$

where f_s is the known frequency of the source signal, v_k is the receiver speed and θ_k is the bearing to the signal source. In accordance with (1) the bearing θ_k to the beacon is related to the value of the base distance a between the SBL HANS receivers on the AUV in the k-th clock cycle by the ratio:

$$\theta_k = \arccos (c*\Delta t_{12k}/a). \quad (6)$$

Using expressions (5-6), we can determine the unknown velocity v_k.

It is required to obtain estimates of the location for each AUV from the group, provided that measurements of the propagation times and propagation velocity of sonar signals, as well as measurements of the motion heading, contain random independent errors.

An algorithm based on a particle filter is used to solve this problem. A cloud of N particles evenly distributed in 2-dimensional space (x, y) is used, where x and y are horizontal coordinates of AUV motion in the horizontal plane The weights of all particles at the initial moment of time are assumed to be the same and equal to $\omega = 1/N$.

To calculate the current coordinates on board the AUV, information about the speed of its motion is generated in each cycle of the SBL HANS operation. At the stage of extrapolation of the particle filter operation, the predicted state vector (x^e_k, y^e_k) is calculated. To do this, the estimates obtained in the previous step (x^p_{k-1}, y^p_{k-1}) are used, as well as the estimated motion of the AUV:

$$x^e_k = x^p_{k-1} + \tau * v_k * \cos(\varphi_k),\ y^e_k = y^p_{k-1} + \tau * v_k * \sin(\varphi_k),$$

where v_k is the calculated speed and φ_k is the measured heading of the AUV.

At the stage of correction for the particle filter operation at the current k-th clock cycle, the estimate of the location (x^p_k, y^p_k) of each AUV from the group is calculated based on the mismatch of the measured difference Δt_{k12} of the arrival time for the signals to the first and second AUV receivers of the SBL HANS, the difference $\Delta t_{k,k-1}$ of the arrival time for the signal in $(k-1)$-th and k-th cycles on the first AUV receiver for the SBL HANS and calculated the differences in the arrival time of the signals for all particles, taking into account their current estimated coordinates.

III. SOME RESULTS OF MODELING FOR AUV POSITIONING BASED ON SBL HANS

The assessment of the accuracy and operability of the navigation algorithm for determining the location of the AUV was carried out within the framework of numerical simulation. The model experiments were performed in the IDLE environment (Python). During the simulation, the period of operation for SBL HANS was 10 seconds. The speed of the acoustical signal propagation was assumed to be equal to 1500 m/s.

The simulation was performed using the Monte-Carlo method, while each experiment was performed in a cycle 100 times with different random errors. Based on the results obtained, the average value and variance of errors in estimating the location and velocity of the AUV were calculated. The initial area of uncertainty for the AUV location was a square with a side of 20 m.

It was assumed that the measurements for the course and arrival times of the signals contained random errors. To compare the results of the proposed method with those described in [2], a model trajectory of the AUV motion was selected, similar to that described in this work. The AUV started from a point with coordinates (-3000, -2000) m. The beacon was located at a point with coordinates (0, 0, 0) m. Parameters of the AUV motion in the horizontal plane: the speed is 1 m/s, the course is 45 degrees. The distance traveled was about 500 meters and the travel time was about 8.3 minutes (fig. 1).

In a situation where, during modeling, random errors in measuring the AUV course were in the range of (-0.5, 0.5) degrees, speeds were in the range of (-0.01, 0.01) m/sec, and random errors in measuring the arrival times of signals were in the range of $(-10^{-3}, 10^{-3})$ seconds, the average value of the error in determining the location of the AUV was 0.79 m, the standard deviation was 0.90 m.

The result indicates a reduction for more than an order of magnitude in error compared to the result given in [2] under similar modeling conditions (fig. 2). At the same time, according to estimates obtained during the modeling process, the time required to implement one clock cycle of the proposed algorithm on board the AUV is reduced by more than 30 times compared to the implementation of a single clock cycle of the algorithm proposed in [2].

CONCLUSION

The simulation results of the proposed method have shown that it allows to estimate the location of the AUV on board in real time with the required accuracy in the absence of an expensive accurate time system on board the AUV. In this situation, due to the out-of-sync for the AUV and the beacon clock, it is not possible to directly calculate the location of the AUV using only the SBL HANS.

ACKNOWLEDGEMENTS

The research was carried out at the expense of the state assignment of the IACP FEB RAS FWFW-2025-0002 (design of modernized algorithm for AUV group navigation based on SBL HANS) and a grant from the Russian Science Foundation No. 23-61-10024 (simulation of the algorithm based on a particle filter).

REFERENCES

[1] Borodin V., Smirnov G., Tolstyakova N. et al. Hydroacoustic navigation devices. L: Sudostorenie, 1983, 263 p.

[2] Dubrovin F., Rodionov A., Scherbatyuk A. On improving the accuracy of the acoustical positioning system with a short base line for AUV group positioning. Proceedings of the XXXIth International Conference on Integrated Navigation Systems, Saint Petersburg, 27 - 29 May 2024.

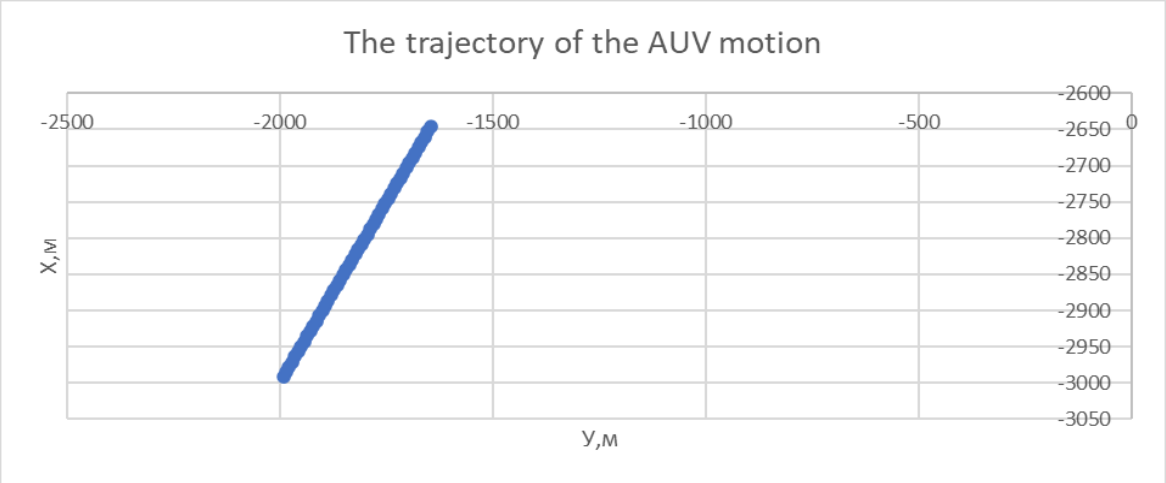


Fig. 1. The trajectory of the AUV (the beacon is located at a point with horizontal coordinates 0.0, 0.0).

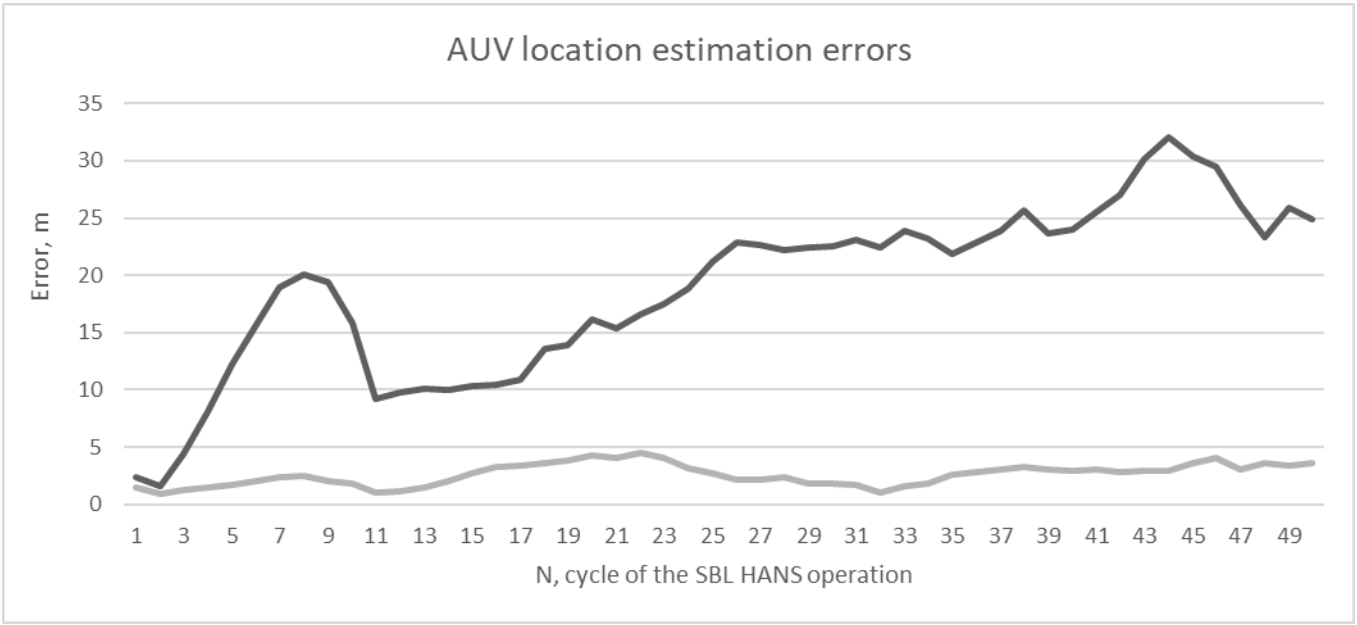


Fig. 2. Errors in estimating the location of the AUV (black curve – without measuring the frequency of the incoming signal, gray curve – with measuring the frequency of the incoming signal).

Method of AUV Coordinates Correction during Observation in the Northern Latitudes

L.A. Martynova
Concern CSRI Elektropribor, JSC
St. Petersburg, Russia
martynowa999@bk.ru

Abstract—The paper considers AUV observation in the northern latitudes by the signals from global navigation satellite systems (GNSS). In the areas where icebergs appear due to global warming and seismic activity the signals from some navigation satellites can be unavailable. The influence of the number of navigation satellites and pseudorange errors for the satellites positioned low above the horizon on HDOP and positioning error is estimated using the developed mathematical model further used in numerical experiments. The results have confirmed that relative positions of satellites do affect AUV positioning error. The simulation results can be used in predicting the place and time of AUV observation with the most favorable navigation conditions for AUV positioning with a minimum error.

Keywords—autonomous underwater vehicle (AUV), observation, global navigation satellite systems (GNSS), iceberg, dilution of precision (DOP).

I. INTRODUCTION

When using autonomous underwater vehicles (AUV) [1, 2] for measuring the marine environmental parameters, ice [3] or bottom [4] conditions in the Arctic region, periodical surface observations by GNSS signals should be made. It is believed that during observation in the northern latitudes, the only problem for the AUV control system is to find an ice hole of appropriate size [5], and there will be no problems with positioning using the GNSS signals [6].

However, due to the global warming and earthquakes in the Russian Arctic [7], large massifs of ice split off and move on the sea surface as icebergs in unpredictable directions at unknown speed. The height of such icebergs can reach several meters, and if the AUV emerges close to it, the angle of navigation satellite observation may be more than 40 deg. This may increase the time of the AUV observation using GNSS signals and degrade the positioning accuracy, because in the northern latitudes the satellites are located no higher than 70 deg above the horizon and mostly close to the horizon [8]. When the full satellite constellation is visible, this configuration does not significantly affect the AUV positioning and even improves the horizontal dilution of precision (HDOP).

If the icebergs shield the signals of low-positioned satellites, the quantity of visible satellites decreases, and HDOP grows, which may prolong the observation process and increase AUV positioning error. Positioning error is not critical when the AUV moves along a route. However, AUV performs some tasks during the mission such as searching for some objects, bottom stations for refueling [9], or measuring the environmental parameters. In this case the AUV needs a more accurate autonomous coordinate fix to identify the detection or measurement point.

AUV long surfacing for obtaining more accurate coordinates from GNSS signals may degrade AUV condition: low temperatures near the water surface discharge the battery faster; AUV hull may be damaged by ice, ships, etc.; AUV can be demasked and observed from above from reconnaissance spacecraft, UAV, scout aircraft, etc.

Due to these features of observation in northern latitudes in iceberg regions it is important to determine how shielding the signals of some satellites influences the observation results. No research of this kind has been conducted before, and no publications on the subject are available.

II. PROBLEM STATEMENT

Let the following be known:

- observation period determined by the accumulation of AUV positioning error in submerged position using onboard inertial navigation system;
- coordinates of the observation area in the vicinity of a route point.

It is needed to determine the AUV positioning error by GNSS signals at these points.

Only GLONASS signals are used as the most reliable of available systems (GPS, Beidou, Galileo) to reduce the risk of determining the false position, particularly during the period of countries' competition for the natural resources.

III. FEATURES OF SATELLITE CONSTELLATION IN NORTHERN LATITUDES

AUV positioning accuracy by GNSS signals is mostly influenced by direct visibility of the satellites, their mutual positions and GNSS signal ionospheric delays [9]. The positions of constantly drifting icebergs cannot be predicted, so AUV may emerge near the iceberg (Fig. 1), which will shield the GNSS signals. This may reduce the number of visible satellites to the critical number of four, when the positioning is still possible, and degrade HDOP. If the satellites are arranged on one line (Fig. 2), this will also degrade the AUV positioning error.



Fig. 1. Icebergs in northern latitudes.

HDOP, which shows how many times the measurement accuracy is degraded, serves as a quantitative measure of the positioning error. Pseudorange measurement error due to the ionospheric signal delays may reach 50 m [10, p. 93].

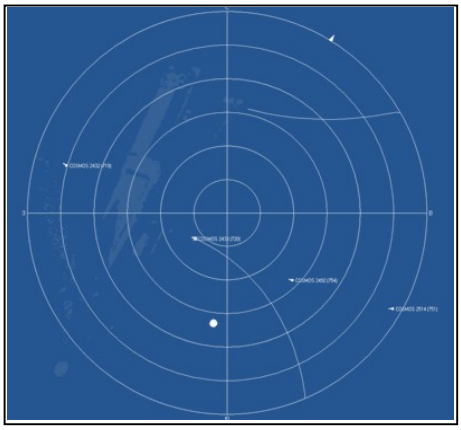


Fig. 2. Changes in satellite number with time [11].

IV. Mathematical Model for the Research

To estimate the effect of satellite position on the results of AUV observation under GNSS signals shielded by the icebergs, we have developed a mathematical model including the mathematical model of satellite orbital motion; navigation algorithm; PDOP determination algorithm; and AUV positioning error determination algorithm.

A. Mathematical Model of Satellite Orbital Motion

The satellite orbital motion is modeled based on non-disturbed (Keplerian) satellite motion subject only to the Earth's gravity force (one attraction center) [12]. Using the Newton's second law and the law of universal gravitation, and taking the time derivative as a total derivative in inertial frame, the equation of motion of the satellite center of mass in inertial frame has the form:

$$m \cdot (d^2r / dt^2) = \mathbf{F},$$

where m is the satellite mass, $\mathbf{F}$ is the gravity force vector, t is the time; r is the radius vector determined by the expression

$$r = \sqrt{(x_0^2 + y_0^2 + z_0^2)},$$

where x_0, y_0, z_0 are the satellite current coordinates (projection of radius vector r on the coordinate axes).

Using the model of satellite orbital motion allows determining the coordinates of all satellites and pseudoranges from the actual AUV position to each satellite.

Based on the results, the group of satellites visible from the AUV actual position and the group of satellites with elevation angle over 40 deg (assuming that the satellites with elevation angles below 40 deg are shielded by the icebergs) are determined. The navigation solution was obtained using the group of non-shielded satellites.

B. Navigation Algorithm

AUV coordinates are determined using the signals from all satellites within the AUV visibility zone [12, p. 523]. To solve the navigation problem by the least squares method we introduce the AUV state vector $\mathbf{x}=|x\ y\ z\ Д|^{\mathrm{T}}$ and the vector of observations of estimated pseudoranges to the k-th satellite, $k=1,\ldots,N$:

$$\mathbf{y}_{\tau,k} = |\tau_{1,k}\ \tau_{1,k}\ \ldots\ \tau_{1,N}\ |^{\mathrm{T}}.$$

where (x,y,z) are the coordinates, Д is the range, τ is the pseudodelays of GNSS signal, N is the quantity of all visible satellites [12, p. 194].

These data and the formulas from [12, p. 523] are used to compute the estimates of the AUV state vector. Based on the navigation solution the estimated AUV coordinates are determined.

C. Determining PDOP

At the step of secondary processing of navigation data the errors in the AUV state vector $x=|x\ y\ z\ Д|^{\mathrm{T}}$ are determined [12, p. 300]:

$$\Delta x = (\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H})^{-1} \mathbf{H}^{\mathrm{T}} \cdot \Delta y_{Д});$$

where $\mathbf{H}$ is the directional cosine matrix, $\Delta y_{Д}$ is pseudorange vector error.

Positional DOP $K_{ПП}$ (PDOP) is determined as

$$K_{ПП} = \sqrt{(\sigma_x^2 + \sigma_y^2 + \sigma_z^2)},$$

where σ_x, σ_y, σ_z are the pseudorange variances. The positional DOP consists of two components: vertical DOP – VDOP and horizontal DOP – HDOP:

$$\mathrm{PDOP}^2 = \mathrm{HDOP}^2 + \mathrm{VDOP}^2,$$

$$\mathrm{HDOP}^2 = \sqrt{(\sigma_x 2 + \sigma_y^2)},$$

$$\mathrm{VDOP}^2 = \sigma_z^2.$$

D. Estimating the Positioning Error Using the Satellite Visible Constellation

The error ΔL of the calculated AUV position (x_0,y_0,z_0) compared to the actual position (x,y,z) was determined using the formula

$$\Delta L = \sqrt{((x-x_0)^2 + (x-x_0)^2 + (x-x_0)^2)}.$$

V. Numerical Experiment

The following initial data were set: observation time; actual AUV coordinates unknown to it; coordinates of satellites visible from the actual AUV position; errors in pseudoranges due to ionospheric delays of the signal from the satellite to AUV.

The following data were determined during the research: satellite coordinates; quantity of satellites visible above the AUV actual position at the current time; satellites with elevation (observation) angle above 40 deg; calculated AUV coordinates using the visible satellites assuming that the satellites positioned low above the horizon are shielded; PDOP and its horizontal and vertical components; AUV positioning errors.

The following parameters were determined by the results of numerical experiments and statistical processing: percentage ratio of the periods with minimum required satellite configuration (4 satellites) to the total modeling period; effect of the pseudorange ionospheric errors on AUV positioning errors; effect of the quantity of satellites and

HDOP on AUV positioning error; degradation of navigation solutions if some signals from low-positioned satellites are shielded.

VI. Results from Numerical Experiments

We selected the point with coordinates 77° N and 90° E, where the hydrocarbon deposits are positioned, as the test point [13]. The numerical experiment was conducted continuously for 6 hours.

The changes in the number of satellites with time are shown in Fig. 3. Upper line shows the quantity of all visible satellites, the lower line shows the number under shielding and observation angle over 40 deg.

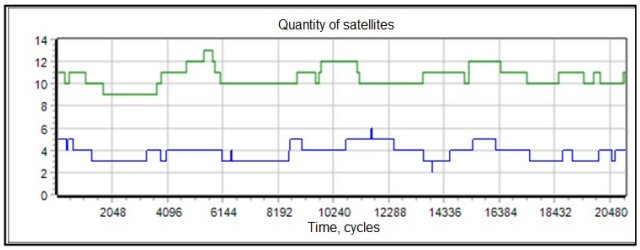


Fig. 3. Changes in the satellite number with time.

As seen from Fig. 3, there were from 9 to 13 satellites with no shielding, however, at some intervals with the shielded signals from the satellites below 40 deg, only three satellites were visible, and at short intervals, two satellites. Three satellites were observed within the half of observation interval, which actually did not allow solving the navigation problem. In other numerical experiment, three satellites were observed continuously within 50 s. This suggests that periods with critically low number of satellites do exist. And even if there were periods with minimal required four satellites, the experience shows that observation by the minimal required constellation can take dozens of minutes.

With the full constellation of visible satellites, PDOP does not exceed 2.5, and HDOP oscillates about 1, because the visible satellites are mostly positioned low above horizon up to 40 deg. However, when shielding the low positioned satellites, the observation results reveal short 40 s HDOP outliers (Fig. 4), which, however, did not affect the positioning error because of their short duration.

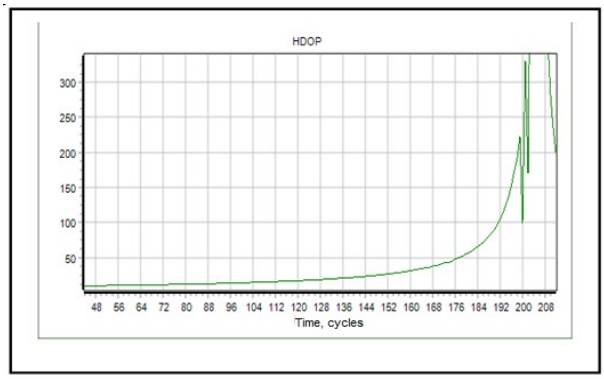


Fig. 4. Changes in the satellite number with time.

At the same time, longer HDOP increases up to 200 were observed when the positions of two satellites coincided. This is explained by the fact that with shielded satellites located low above the horizon only high positioned satellites participate in coordinate determination.

When signals from the whole visible satellite constellation were used, the positioning error was within 1 m. When some satellites were shielded, the error sometimes increased up to 395 m and even more for short several seconds periods.

To study the pseudorange effect on AUV positioning error, a 25 m error was added to pseudoranges of two random visible satellites with elevation above 40°.

Simulation results have shown that AUV positioning error was from 410 m to 1.0–4.5 km, with DOP factors taking the values HDOP – 29, VDOP – 30, PDOP – 32; the number of visible satellites at this time changed from 4 to 3. Short 1 min kilometer AUV positioning errors were observed in the computations.

VII. Conclusions

Numerical experiments have shown that changing the constellation of visible satellites and their elevation angles affect HDOP and AUV positioning error. Some intervals from 1 to 50 min were detected with less than four required satellites, which makes it impossible to solve the navigation problem, and with four satellites HDOP could increase, and AUV positioning error could grow, though shortly, from the fractions of meter to kilometer.

Therefore it is proposed to calculate the AUV orbital constellation when forming the AUV mission, because iceberg positions cannot be predicted.

The proposed AUV positioning error reduction method consists in selecting the observation site and time with the most favorable satellite constellation assuming possible appearance of icebergs shielding the signals from some satellites. This approach selecting the favorable satellite constellation is used in land navigation surveys [14]. The initial data for the prediction are the coordinates of the potentially determined observation point on the route path within the scheduled observation period and the current satellite ephemeride data. For each date of scheduled observation, the plot of visible satellite number and PDOP (GDOP) values depending on the time of the day are predicted. The obtained plots and tables are used to select the periods optimal for satellite observation at the observation site.

References

[1] Ageev, M.D., Kiselev, L.V., Matvienko, Yu.V. et al., Avtonomnye podvodnye roboty: sistemy i tekhnologii (Autonomous Underwater Robots: Systems and Technologies), Ageev, M.D., Ed., 2005.

[2] Inzartsev. A.V., Kiselev L.V., Kostenko, V.V., Matvienko, Yu.V., Pavin, A.M., Shcherbatyuk, A.F., Podvodnye robototekhnicheskie kompleksy: sistemy, tekhnologii, primenenie (Underwater Robotic Complexes: Systems, Technologies, Applications), Kiselev, L.V., Ed., Insitute of Marine Technologies of RAS, Vladivostok, 2018.

[3] Bozhenov, Yu.A., Use of autonomous underwater vehicles for Arctic and Antarctic research, Fundamentalnaya i prikladnaya gidrofizika, 2011, vol. 4, no. 1, pp. 47-68.

[4] Martynova, L. A.Method of coordinated behavior of the emitter and autonomous underwater vehicles for effective seismic exploration, Informatsionno-upravlyayuschie sistemy, 2017, no. 1 (86), pp. 83-92.

[5] Martynova, L.A., Mashoshin, A.I., Pashkevich, I.V., and Sokolov, A.I., Algorithms implemented by the AUV integrated control system, Izvestiya YuFU. Tekhnicheskie nauki, 2015, no. 1 (162), pp. 50-58.

[6] Laptev, K.Z. and Bagnitsky, A.V., An ice hole search problem for a long-range AUV in the Arctic regions, Izvestiya YuFU. Tekhnicheskie nauki, 2019, no. 1 (203), pp. C.-28-38, DOI 10.23683/2311-3103-2019.

[7] Why more icerbergs appear in the arctic, https://rg.ru/2024/05/13/reg-szfo/plavuchie-tanki.html.

[8] Challenges for Positioning and Navigation in the Arctic. https://mycoordinates.org/challenges-for-positioning-and-navigation-in-the-arctic/ (date of access 20.01.2025).

[9] Kolesnikov, M.P., Martynova, L.A., Pashkevich, I.V., and Shelest, P. S., Method of positioning an autonomous underwater vehicle in the process of bringing it to the mooring device, Izvestiya Tulskogo gosudarstvennogo universiteta. Tekhnicheskie nauki, 2015, no. 11-2, pp. 38-49.

[10] Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J., Global Positioning System. Theory and Practice, Springer-Verlag Wien New York, 1994.

[11] Orbitron - Satellite Tracking System. https://www.stoff.pl/

[12] GLONASS. Printsipy postroeniya i funktsionirovaniya (GLONASS. Principles of Design and Functioning), Perov, A.I., Kharisov, V.N., Eds., fourth edition, Moscow: Radiotekhnika, 2010.

[13] Arctic zone of the Russian Federation: risks and development prospects. URL: https://yakovpartners.ru/publications/arkticheskaya-zona-rf-riski-i-perspektivy/

[14] Instructions for the development of survey justification and survey of the situation and terrain using global navigation satellite systems GLONASS and GPS, Moscow: TsNIIGAiK, 2002.

Navigation System Neural Network Architecture Parameters Preformance Influence Investigation

Ivan Kuznetsov

Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
kuznetsovim@mai.ru

Konstantin Veremeenko

Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
nio3@mai.ru

Maksim Zharkov

Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
zharkovmv@mai.ru

Andrei Pronkin

Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
pronkinan@mai.ru

Roman Zimin

Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
ziminry@mai.ru

***Abstract*—When addressing the challenges of Unmanned Aerial Vehicles (UAVs) autonomizing, developers typically seek ways to improve the performance of their navigation systems. One of the relatively new and potentially promising approaches to constructing algorithms for UAV navigation systems (NS) is the use of artificial intelligence techniques. The advantages of artificial neural networks (NN) presented in the literature (including, among others, maintaining the operational capability of the complex with an acceptable level of accuracy for a certain period during the absence of signals from the primary corrector – GNSS) are unachievable without appropriately provided preliminary training. The selection and design of the NN architecture, as well as the selection of training datasets and UAV flight trajectories, constitutes a separate scientific task requiring detailed elaboration. Often, as can be observed from the results of other studies, the best solution lies in empirical techniques and the selection of an appropriate combination of all mentioned factors. This article presents a study on the influence of neural network hyperparameters on accuracy, convergence time, training time, and the overall operational efficiency of the integrated system. The investigated parameters are neural network architecture, hidden layers size, batch size, initial learning rate, and other parameters. To identify patterns, a comprehensive study of the influence of such neural network hyperparameters on the UAV NSs performance is provided. The research is carried out using realistic application scenarios for such NSs – employing a developed post-processing suite, simulating both the loss of signals from the primary corrector (GNSS) and considering scenarios where the signal significantly degrades due to intentional interference. Comparative results are presented, reflecting various implementation scenarios of the navigation complex depending on the values of the considered parameters, demonstrating achievable accuracy levels of the navigation system.**

Keywords— Navigation system, artificial intelligence, machine learning, extended Kalman filter, integrated information processing, hyperparameters

I. INTRODUCTION

The task of improving navigation accuracy and the level of autonomy of Unmanned Aerial Vehicles (UAVs) remains one of the most challenging in modern avionics. Constantly growing requirements for the UAV navigation systems (NSs) performance stimulate developers to search for new approaches and technologies [1 – 4]. The integration of UAVs into general airspace, as noted in strategic documents such as the Federal Project for Unmanned Aviation Systems (FP UAS) [6], requires not only high accuracy but also reliability of the navigation solution, including challenging environments, for example, during GNSS signal jamming or spoofing [5].

Traditional integration techniques, based on optimal filtering, typically using various forms of the Kalman filter [1, 7], have proven themselves well but have limitations, especially when dealing with nonlinear models or in conditions of uncertainty. In recent years, there has been significant interest in applying artificial intelligence (AI) techniques, and particularly artificial neural networks (NNs), to solve navigation problems [8-11]. Their ability to approximate complex nonlinear dependencies, adaptability, and the lack of need for a priori error models open up new possibilities for constructing UAV NSs.

In previous work by the authors, presented at the ICINS-2024 conference [15], an architecture for a UAV NS with an integrated AI module (UAV AI NS) based on an LSTM network was proposed. Preliminary results of field tests have been presented, demonstrating the potential of the approach: the system maintained operational capability with an acceptable level of accuracy during relatively long periods of GNSS signal absence, while predicting navigation parameters. However, shortcomings were also identified such as the significant time required for training the neural network module in the proposed architecture, as well as the potential dependence of prediction accuracy on the nature of the UAV's flight trajectory used during training.

This work is a continuation of the previous one [15] and is dedicated to studying the influence of key parameters of the neural network module – so-called hyperparameters – on the overall performance of the UAV AI NS. The choice of NN architecture, its configuration (size of layers, training batch size, learning rate, etc.), as well as the selection of training datasets, are nontrivial tasks that are often solved through empirical methods. The aim of this work is a systematic investigation of the influence of these factors on

The research was funded by the Russian Science Foundation (project No. 24-29-20304, https://rscf.ru/en/project/24-29-20304/

the accuracy of the navigation solution, the convergence time of the training algorithm, and the overall efficiency of the NS, especially in realistic operational scenarios, including conditions of GNSS signal degradation or loss.

II. Proposed Approach

This section describes the main mathematical relationships underlying the functioning of the integrated navigation system (INS), the structure of the UAV INS with an integrated artificial intelligence (AI) module, and details the approach to investigating the influence of this module's parameters on the overall system performance.

A. Basic INS Equations

The object is an integrated inertial-satellite navigation system, including a tactical-grade inertial measurement unit (IMU) and a GNSS receiver. Joint position-velocity and attitude correction of inertial navigation system (INS) errors using an optimal Kalman filter (OKF) is considered.

The state equation of the system, describing the error dynamics, in general form is:

$$\dot{\bar{x}}(t) = F(t)\cdot\bar{x}(t) + G(t)\cdot\bar{w}(t), \quad (1)$$

where $\bar{x}(t)$ – is the system state vector, $F(t)$ – is the dynamics matrix, $\bar{w}(t)$ – is the system noise vector, $G(t)$ – is its intensity matrix.

The state vector is estimated from measurements $\bar{z}(t_k)$, available at discrete time instants t_k:

$$\bar{z}(t_k) = H(t_k)\cdot\bar{x}(t_k) + \bar{v}(t_k). \quad (2)$$

where $H(t_k)$ – is the measurement matrix, $\bar{v}(t_k)$ – is the measurement noise vector.

The state vector $\bar{x}$ includes INS errors in determining coordinates $\delta\bar{P}$, velocity projections $\delta\bar{V}$ and attitude angles $\delta\bar{A}$, as well as parameters of the sensors error model (gyroscopes and accelerometers) $\delta\bar{S}$ – constant components of drifts and zero biases:

$$\bar{x} = \left[\delta\bar{P}^T, \delta\bar{V}^T, \delta\bar{A}^T, \delta\bar{S}^T\right]^T. \quad (3)$$

If necessary, the sensor error model can be expanded, which will lead to a corresponding increase in the dimensionality of the state vector. The dynamics matrix $F(t)$ is designed based on the INS error equations and the sensor error model. The system noise vector $\bar{w}(t)$ includes random components of INS sensor errors, and the measurement noise vector $\bar{v}(t_k)$ – includes random components of the corrector error.

The measurement vector $\bar{z}$ is formed as the difference between the INS and the corrector (GNSS) coordinates $\Delta\bar{P}$, velocity projections $\Delta\bar{V}$ and, if corresponding information is available, attitude angles $\Delta\bar{A}$:

$$\bar{z} = \left[\Delta\bar{P}^T, \Delta\bar{V}^T, \Delta\bar{A}^T\right]^T. \quad (4)$$

To estimate the state vector $\bar{x}$ an Extended Kalman Filter (EKF) in discrete Joseph form [7], is used, as in the previous work [15].

B. Structure of UAV INS with AI Module

The overall structure of the INS remains the same: the INS is the core of the system, providing continuous output of navigation parameters, and the GNSS receiver(s) is the main corrector, information from which is used by the OKF to estimate and compensate for accumulating INS errors.

The key element is the module based on an artificial neural network (NN) integrated into the NS structure. As shown in [15], such a module can be used to predict navigation parameters (coordinates, velocities, attitude angles) during periods of absence or degradation of GNSS signal quality, by approximating the output data of the OKF or directly the INS dead-reckoning algorithms (see Fig. 1 and Fig. 2 from [15]). The neural network is trained on data obtained during normal INS operation with reliable GNSS signals.

C. Investigation of the Influence of Neural Network Hyperparameters

Previous results [15] demonstrated the fundamental operational capability of the UAV AI NS using an LSTM (Long Short-Term Memory) architecture, but also revealed the need to optimize the training process and improve the robustness of the neural network module. The current research main focus is the systematic study of the influence of neural network hyperparameters on the NS overall performance. Hyperparameters are parameters that are set before the training process begins and define both the network structure and the training process itself.

Within this work, the influence of the following key hyperparameters is investigated:

1) Neural network architecture: Although LSTM remains the base, its modifications (e.g., bidirectional LSTMs, GRU – Gated Recurrent Units) and other types of recurrent networks adapted for time series are considered, aiming to find an optimal balance between the ability to remember long-term dependencies and computational complexity (an example of an architecture is shown in Fig. 1).

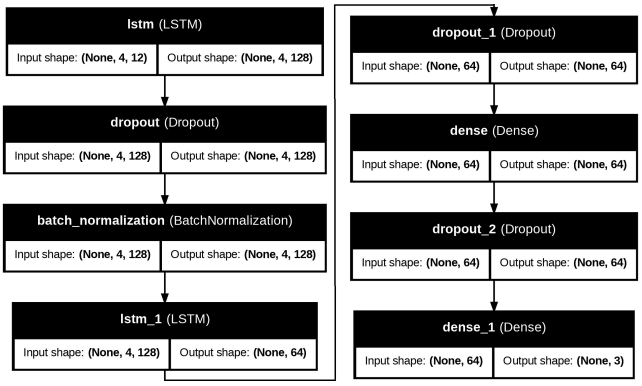


Fig. 1. Example of one of the neural network architectures considered in the study.

2) Hidden layers size: The number of neurons in the hidden layers of LSTM or other recurrent cells directly affects the capacity of the model – its ability to remember and approximate complex dependencies. Too small a size can lead to underfitting, while too large can lead to overfitting and increased computation time.

3) Number of layers: The depth of the network (number of sequential hidden layers) also affects its expressive power

but increases the risk of vanishing gradients and complicates training.

4) Batch Size: The number of training examples used for one step of updating the network weights. This parameter can affect its generalization ability.

5) Initial Learning Rate: Determines the step size when updating network weights during gradient descent. Too high a rate can lead to training instability and jumping the optimum, while too low a rate can lead to excessively long training. Adaptive learning rate strategies (e.g., Adam, RMSprop) are also investigated.

6) Activation function: The type of non-linear function applied in neurons (e.g., ReLU, sigmoid, tanh) affects the training dynamics and the network's ability to approximate non-linearities.

7) Regularization methods: The application of techniques aimed at preventing overfitting, such as Dropout (randomly "turning off" some neurons during a training iteration) or L1/L2 weight regularization, is investigated.

The aim of this research is not just to select optimal values for a specific task, but to identify general patterns and relationships between hyperparameters and the final characteristics of the UAV AI NS: the accuracy of predicting navigation parameters in autonomous mode, the time required to train the neural network to a given accuracy, and the stability of operation in various scenarios, including those that differ from the training trajectories. Understanding these dependencies will allow for a more relevant approach to designing neural network modules for navigation complexes.

III. Experimental Results

The aim of the experimental part of the work is to investigate the influence of the hyperparameters of the neural network module on the accuracy and operational characteristics of the UAV navigation system with AI modules (UAV AI NS), presented in [15]. The experiments are carried out using data obtained during field tests with the MAI Navigation System Car Testbed [12, 13], and a post-processing software developed by the authors (an example of a test trajectory (34 minutes) with the Testbed driving through Moscow is shown in Figure 2).

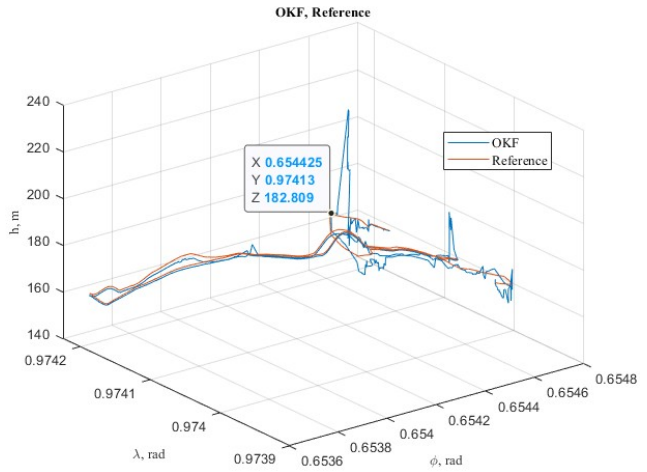


Fig. 2. Trajectory of the MAI Navigation System Car Testbed (in the figure, OKF – Optimal Kalman Filter solution, Reference – data obtained from post-processing).

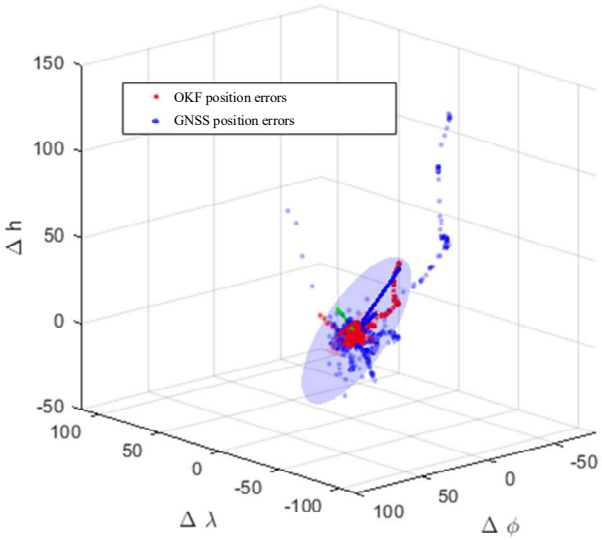


Fig. 3. Position errors on the test trajectory of MAI Navigation System Car Testbed (in the figure, error ellipsoids – blue – GNSS, red – OKF).

A. Experimental Research Methodology

To systematically study the influence of the hyperparameters described in section II.C, the following approach is applied:

1) *Formation of training and test datasets:* Data recorded during drives of a vehicle equipped with a tactical-grade IMU (characteristics similar to Table 1 from [15]) and GNSS receivers are used. The data include raw IMU measurements (angular velocities and linear accelerations) and GNSS measurements (code and carrier phase pseudoranges, Doppler shifts). Particular attention is paid to including data recorded in poor GNSS signal reception conditions (e.g., dense urban development in Moscow, driving under bridges, in tunnels) in the datasets to assess the robustness of the algorithms. The output data of the reference trajectory and INS parameters (coordinates, velocities, attitude angles), obtained using the OKF during periods of good GNSS signal, are used to train the neural network modules. Separate, non-overlapping datasets are formed for training, validation (hyperparameter tuning), and testing (final evaluation).

2) *Training of neural network modules:* For each investigated set of hyperparameters (architecture, number of layers/neurons, batch size, learning rate, etc.), the corresponding neural network model is trained. INS parameters and/or raw IMU measurements are used as input data for training, and GNSS/OKF output parameters are used as target output data. Training is carried out until convergence or a specified number of epochs.

3) *Evaluation of neural network performance:* Trained models are evaluated on test datasets. The main evaluation scenario includes simulating GNSS signal loss of varying durations. During these periods, the neural network module operates in prediction mode, using its previous output data and current IMU measurements to generate a navigation solution. The following characteristics are evaluated:

a) Prediction accuracy: Predicting coordinates, velocities, and attitude angles are calculated by comparing the output of the neural network module with the reference solution (obtained using MAI software and Novatel Inertial Explorer [14] when processing with continuous GNSS data). Metrics used: root mean square error (RMSE), mean absolute error (MAE), mathematical expectation (ME) of error.

b) Training time: The time required to train each neural network configuration to a specified error threshold on the validation dataset is recorded.

c) Stability under various conditions: The degradation of accuracy when operating on trajectories and in conditions (motion dynamics, level of GNSS interference) different from those used during training is assessed.

B. Obtained Results

Based on the results of the comprehensive study, quantitative assessments of the influence of various hyperparameters on the characteristics of the UAV AI NS were obtained. The main results are presented below.

- **Comparative error graphs:** Similar to Fig. 6 from [15], within this work, graphs of errors in determining coordinates, velocities, and attitude angles over time in prediction mode (Figure 5) were plotted for various configurations of the neural network module, examples of which are given in Table 1. This allows assessing the influence of hyperparameters on the dynamics of error accumulation.

For quantitative assessment, as well as for greater clarity of the influence of each hyperparameter on the target metric (e.g., test_mae), a machine learning model based on RandomForestRegressor was trained. Its features were the hyperparameters of the models, and the target variable was the corresponding metric (test_mae). Based on the trained machine learning model, the importance of each hyperparameter (feature_importances) was calculated. The results are visualized as a bar chart showing the relative contribution of each hyperparameter to the prediction of the target metric (Figure 4).

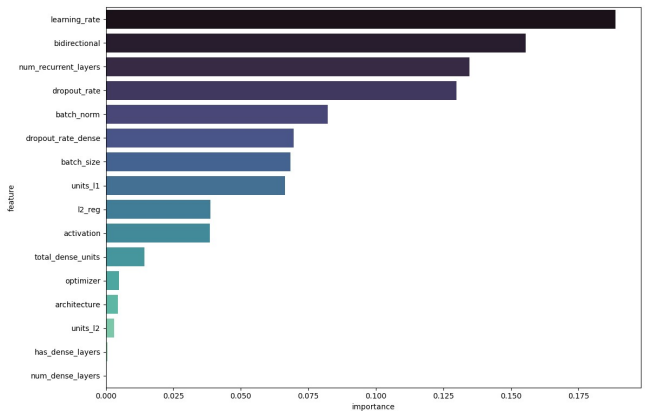
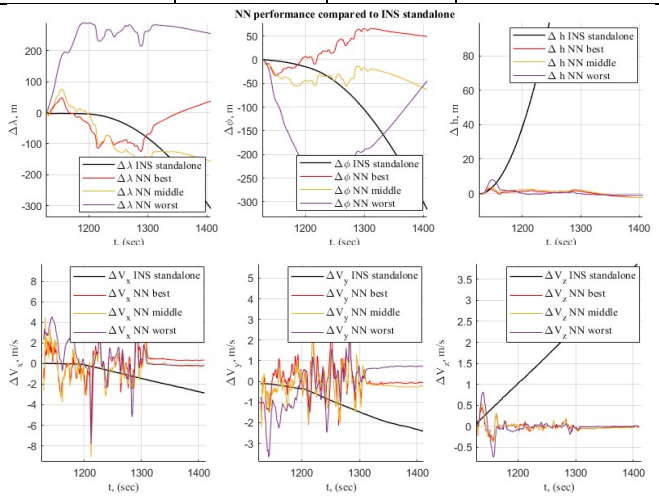


Fig. 4. . Example of calculating the importance of neural network hyperparameters for predicting the Test MAE target metric.

TABLE I. EXAMPLES OF CONFIGURATIONS OF THE UAV AI NS NEURAL NETWORK MODULE (THE BEST MODEL ACCORDING TO THE OBTAINED VALUES OF KEY TRAINING METRICS IS HIGHLIGHTED IN ITALICS)

model_ name	**test_mae**	**test_mse**	**training_ time_ sec**
kt_best_ trial_1	0,017	0,002	51,087
kt_best_ trial_2	*0,016*	*0,002*	*52,428*
kt_best_ trial_3	0,022	0,003	45,747
kt_best_ trial_4	0,039	0,005	56,523
kt_best_ trial_5	0,024	0,003	276,634
epochs_ to_ best_val_ loss	**architecture**	**units_l1**	**num_ recurrent_ layers**
58	GRU	0	1
44	*GRU*	*0*	*1*
42	GRU	0	1
30	GRU	0	1
35	LSTM	0	1
units_l2	**batch_size**	**learning_ rate**	**optimizer**
0	64	0,005	adam
0	*16*	*0,001*	*rmsprop*
0	32	0,001	rmsprop
0	32	0,01	adam
0	16	0,0005	rmsprop
dropout_ rate	**l2_reg**	**activation**	**batch_ norm**
0,1	0,001	tanh	0
0,4	*0,0001*	*relu*	*1*
0,3	0	tanh	0
0	1,00E-06	tanh	1
0,2	0,001	relu	1
bidirectional	**num_ dense_ layers**	**total_ dense_ units**	**has_ dense_ layers**
1	0	0	0
0	*0*	*0*	*0*
0	0	0	0
1	0	0	0
1	0	0	0



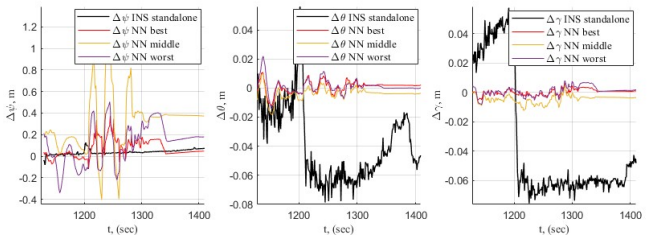


Fig. 5. Performance of the UAV AI NS compared to a standalone INS (the graphs show the performance of the UAV AI NS algorithms (for the best/average/worst configurations of the neural network selected based on research results, in terms of obtained training metrics (indicated in red/yellow/purple colors respectively)), black color – standalone INS, 1st row (3 graphs): errors in determining longitude, latitude, and altitude, 2nd row (3 graphs): errors in determining linear velocity projections along the respective axes, 3rd row (3 graphs): errors in determining heading, pitch, and roll respectively).

- **Tables with final metrics:** The obtained results were summarized in Table 2 for clarity, reflecting the ME, RMSE, and Mean Absolute Error (MAE) of errors in determining basic navigation parameters for various sets of hyperparameters when predicting during periods of absence/distortion of GNSS signals (Figure 4 shows such a time period of about 300 seconds).

TABLE II. TABLE 2 – OBTAINED QUANTITATIVE PERFORMANCE CHARACTERISTICS OF UAV AI NS COMPARED TO STANDALONE INS

Error	Source of information	ME	RMSE	MAE
Coordinates (longitude, latitude, and altitude), [m]	Standalone INS	-85,53/ -100,6/ 185,73	94,5/ 97,44/ 164,2	-
	Best NN	-32,4/ 28,4/ 0,19	50,25/ 31,9/ 1,32	22,1/ 16,3/ 0,74
	Average NN	-85,77/ -37/0.8	71,23/ 13,5/ 1,59	45,46/ 29,4/ 1.2
	Worst NN	245,42/ -155,5/ 0,44	58,1/ 63,9/ 1,92	194,12/ 115,3/ 2,46
Linear velocity projections (ΔVx, ΔVy, ΔVz), [m/s]	Standalone INS	-1,12/ -1,15/ 1,96	0,94/ 0,76/ 1,11	-
	Best NN	0,06/ 0,08/ -0,008	1,34/ 0,7/ 0,09	1,9/ 0,97/ 0,06
	Average NN	-0,26/ -0,1/ -0,009	1,54/ 0,75/ 0,11	2,06/ 1,11/ 0,06
	Worst NN	0,42/ -0,06/ 0,003	1,68/ 1,05/ 0,17	2,27/ 1,81/ 0,07
Heading, pitch, and roll, [deg]	Standalone INS	0,036/ -0,039/ -0,034	0,02/ 0,026/ 0,026	-
	Best NN	0,07/ 0,009/ 0,0025	0,24/ 0,008/ 0,006	0,18/ 0,005/ 0,004
	Average NN	0,31/ -0,008/ -0,004	0,47/ 0,008/ 0,006	0,46/ 0,006/ 0,006
	Worst NN	0,11/ 0,0094/ 0,0094	0,28/ 0,008/ 0,007	0,31/ 0,005/ 0,005

- **Dependency graphs:** For the convenience of analyzing the cross-correlation, as well as the direct effect of hyperparameter values on the values of key metrics, heat maps and boxed statistical diagrams (Figure 6) were used to illustrate the dependence of key metrics on the value of a particular hyperparameter. During the research, this greatly simplified the task of identifying optimal ranges of hyperparameter values, as well as possible trade-offs when choosing a specific model.

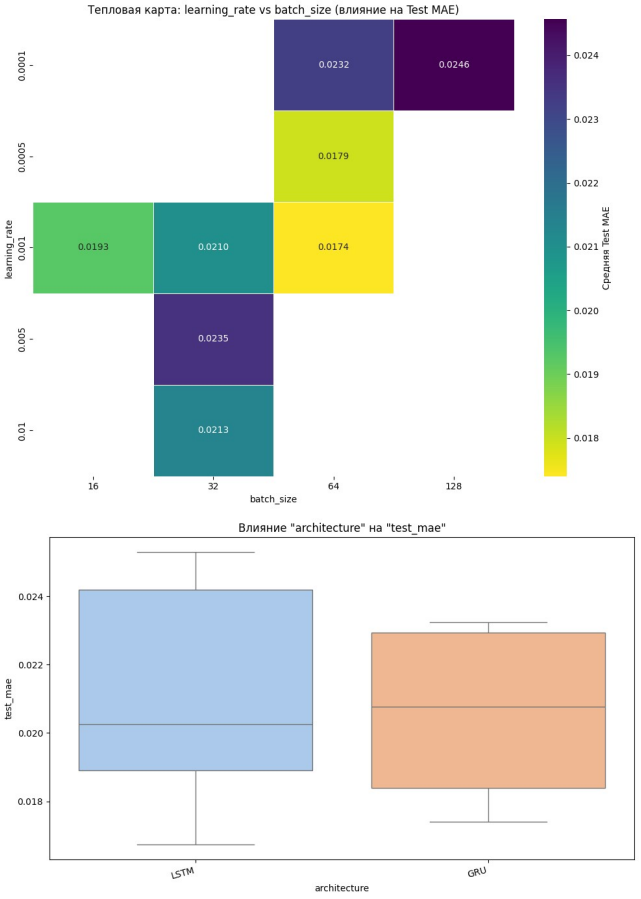


Fig. 6. Examples of auxiliary analysis tools – heat maps illustrating the dependence of key metrics on the value of a specific hyperparameter, as well as box plots.

- **Analysis of the impact of difficult conditions:** Accuracy estimates were obtained for scenarios with intentional/accidental GNSS interference in dense urban environments for test drives using the MAI Navigation System Car Testbed. The obtained results demonstrated the influence of hyperparameter selection on system robustness.

The research allowed identifying combinations of hyperparameters that provide the best compromise between prediction accuracy, training time, and operational stability of the neural network module within the UAV AI NS. For example, based on the research results, it can be seen how increasing model complexity (more layers/neurons) affects accuracy, but at the same time increases both training time and the risk of overfitting, and how regularization methods or selecting an optimal batch size can mitigate these effects. It was also found that the choice of a specific architecture (such as LSTM or GRU considered in the work) has less

influence on the ability to model long-term dependencies when predicting navigation parameters.

IV. Conclusion

This paper presents the development of research dedicated to the application of artificial intelligence methods for improving the characteristics of unmanned aerial vehicle navigation systems (UAV AI NSs) in conditions of absence or degradation of GNSS signals. Based on the results of previous work [15], where the principal possibility of using neural network modules (using the LSTM architecture as an example) for predicting navigation parameters was shown, this study focused on solving identified problems related to training duration and the potential dependence of accuracy on the specifics of training trajectories.

A key aspect of the presented approach is the systematic investigation of the influence of neural network hyperparameters (such as architecture, layers size, batch size, learning rate, regularization methods) on the final performance of the UAV AI NS. A methodology for experimental studies using field test data and a post-processing software was described, allowing not only the assessment of prediction accuracy in autonomous mode but also training time, as well as system robustness in realistic operational scenarios, including urban environments with unstable GNSS signal reception.

The research results provided quantitative estimates and identified patterns of influence of various hyperparameters. This will allow developers to approach the selection of neural network module configuration more soundly, finding an optimal balance between accuracy, computational costs for training and operation, and stability under various flight conditions. Understanding these dependencies is a critically important step for transitioning from conceptual models to creating effective and reliable UAV AI NSs capable of meeting the stringent requirements for autonomous navigation.

Further research will be aimed at validating the obtained conclusions on a wider range of flight data and UAV scenarios. It is also planned to study the possibility of applying adaptive hyperparameter tuning methods and online learning techniques to enhance the adaptability of the neural network module to changing operating conditions. Furthermore, a promising direction is the investigation of the influence of the diversity and representativeness of training datasets on the generalization ability of the neural network and the development of methods to reduce the dependence of prediction accuracy on a specific flight trajectory. Successful solution of these tasks will open new doors for creating highly autonomous next-generation navigation systems for UAVs.

The research results demonstrate the possibility of using the proposed approach for UAV navigation and attitude determination, with the integrated system functioning with an artificial intelligence module in its structure under conditions of GNSS sufficiently long signal loss. Subsequently, it is planned to investigate the influence of the trajectory on the performance and achievable accuracy levels of the neural network, structurally a part of the UAV NS.

References

[1] O. Stepanov, A. Motorin «Models and methods in estimation tasks in the processing of navigation information» // Materials of the XI conference in memory of the outstanding designer of gyroscopic devices N. N. Ostryakov, St. Petersburg, pp. 285-288, October 2020

[2] K. Veremeenko, K. Sypalo, D. Kozorez «Modeling of the processes of functioning of integrated aircraft systems based on the methodology of object-oriented analysis and design» // Moscow, p. 93, 2010.

[3] Tang, Y.; Jiang, J.; Liu, J.; Yan, P.; Tao, Y.; Liu, J. A GRU and AKF-Based Hybrid Algorithm for Improving INS/GNSS Navigation Accuracy during GNSS Outage. Remote Sens. 2022, 14, 752. https://doi.org/10.3390/rs14030752

[4] M. Zharkov, K. Veremeenko, D. Antonov, I. Kuznetsov «Attitude determination using ambiguous GNSS phase measurements and absolute angular rate measurements» // Gyroscopy and navigation, Vol. 26, № 3(102), pp. 54-68, 2018.

[5] M. Zharkov, K. Veremeenko, I. Kuznetsov, A. Pronkin, «Global Navigation Satellite System Spoofing Detection in Inertial Satellite Navigation Systems» // Inventions p.8, 2023. https://doi.org/10.3390/inventions8060158

[6] Decree of the Government of the Russian Federation dated June 21, 2023 No. 1630-r "On Approval of the Strategy for the development of unmanned Aircraft of the Russian Federation for the period up to 2030 and for the future up to 2035"

[7] Veremeenko, K.K., Zharkov, M.V., Kuznetsov, I.M. et al. «Strapdown Inertial Navigation System Transfer Alignment: Algorithmic Features and Simulation Performance Analysis» // Russ. Aeronaut. 63, 618–626 (2020).

[8] W. Fang, J. Jiang, S. Lu, Y. Gong, Y. Tao, Y. Tang, P. Yan, H. Luo, J. Liu «A LSTM Algorithm Estimating Pseudo Measurements for Aiding INS during GNSS Signal Outages» // MDPI Remote Sens, Switzerland, p.24, 2020.

[9] Sheng, Y.; Liu, H.; Li, J.; Han, Q. UAV Autonomous Navigation Based on Deep Reinforcement Learning in Highly Dynamic and High-Density Environments. Drones 2024, 8, 516. https://doi.org/10.3390/drones8090516

[10] Fang, W.; Jiang, J.; Lu, S.; Gong, Y.; Tao, Y.; Tang, Y.; Yan, P.; Luo, H.; Liu, J. A LSTM Algorithm Estimating Pseudo Measurements for Aiding INS during GNSS Signal Outages. Remote Sens. 2020, 12, 256. https://doi.org/10.3390/rs12020256

[11] H. Dai, H. Bian, R. Wang, H. Ma, An INS/GNSS integrated navigation in GNSS denied environment using recurrent neural network, Defence Technology, Volume 16, Issue 2, 2020, Pages 334-340, ISSN 2214-9147, https://doi.org/10.1016/j.dt.2019.08.011.

[12] M. Zharkov, K. Veremeenko, I. Kuznetsov, A. Pronkin «Experimental Results of Attitude Determination Functional Algorithms Implementation in Strapdown Inertial Navigation System» // Sensors 2022, 22, 1849

[13] MAI Navigation and Attitude Determination Systems Research and Development Laboratory. Available online: https://mai.ru/science/studies/equipment/index.php?ELEMENT_ID=167952 (accessed on 20 April 2025)

[14] NovAtel Inertial Explorer. Онлайн доступ: https://novatel.com/products/waypoint-post-processing-software/inertialexplorer (accessed on 20 April 2025)

[15] K. K. Veremeenko, M. V. Zharkov, R. Yu. Zimin [et al.] // 31st Saint Petersburg International Conference on Integrated Navigation Systems : Collection of plenary and poster paper, Saint-Petersburg, 27–29 of May 2024. – Saint-Petersburg: State Research Center of the Russian Federation Concern CSRI Elektropribor, JSC, 2024. – P. 121-125.

Fast Block LD Kalman Filter for Visual-Inertial Navigation Tasks

N. Tsiopliakis

Dept. of Information and Measuring Technology
South Ural State University
Chelyabinsk, Russia
ntsiopliakis@mail.ru

***Abstract* — This report presents the Fast Block LD Kalman Filter (FBKF-LD) for visual-inertial navigation. The filter estimates the full state vector — including visual feature states — with *O*(*N*) computational complexity. Multi-core processor implementation demonstrates real-time performance while tracking hundreds of features.**

Keywords— visual-inertial navigation, Kalman filtering.

I. INTRODUCTION

Vision-aided navigation systems require joint processing of data from multiple simultaneously tracked visual features. To ensure real-time operation, a fast fusion filter is essential. However, traditional Kalman-based filters such as the Extended Kalman Filter (EKF) exhibit computational complexity that scales as $O(N^3)$ with the number of features N. For visual-inertial navigation, fast filters with $O(N)$ complexity (e.g., the Multi-State Constraint Kalman Filter, MSCKF [1-2]) exist, but they exclude feature states from both the measurement equations and filter state vector. Consequently, feature parameter estimation occurs outside the Kalman filter framework.

In contrast, the FBKF – extending its formulation in [3] – provides an $O(N)$-complexity approximation of the EKF for visual-inertial navigation, while retaining full state vector estimation.

II. SYSTEM MODEL

The system model used in the filter is suitable for describing a linearized error model of a visual-inertial navigation system with features parameterized in both the body-fixed and world-fixed frames.

The residual of the i-th feature's measurement equation is modeled as a linear combination of the computer vision system's measurement noise $\mathbf{Y}_i$, errors in the strapdown inertial navigation system (SINS) parameters $\mathbf{X}_s$, and errors in the i-th visual feature's location parameters $\mathbf{X}_i$:

$$\mathbf{Y}_i = \mathbf{v}_i - \mathbf{H}_{is}\mathbf{X}_s - \mathbf{H}_i\mathbf{X}_i, \quad i = 1...N\,. \tag{1}$$

The measurement noises for different features are uncorrelated. The dynamics model consists of the dynamics equation of the SINS errors and the dynamics equations of the visual feature errors:

$$\begin{aligned} \breve{\mathbf{X}}_s &= \mathbf{F}_s\mathbf{X}_s + \mathbf{w}_s\,, \\ \breve{\mathbf{X}}_i &= \mathbf{F}_i\mathbf{X}_i + \mathbf{F}_{is}\mathbf{X}_s + \mathbf{w}_i + \mathbf{M}_{i\,ws}\mathbf{w}_s, \quad i = 1...N, \end{aligned} \tag{2}$$

where $\breve{\mathbf{X}}_s$, $\breve{\mathbf{X}}_i$ are the state vectors of the error model for the predicted estimates; $\mathbf{w}_s$, $\mathbf{w}_i$, $i = 1...N$ are the dynamic noise vectors (uncorrelated with each other).

III. FBKF APPROACH

Immediately after initializing N new visual features, the i-th feature's parameter errors $\mathbf{X}_i^0$ can be expressed as a linear combination of errors in the data used for its initial location estimation:

$$\mathbf{X}_i^0 = \mathbf{M}_{i\,s}^0 \cdot \mathbf{X}_s^0 + \boldsymbol{\gamma}_i^0, \quad i = 1...N\,, \tag{3}$$

where $\mathbf{X}_s^0$ is the SINS error model state vector at initialization; $\boldsymbol{\gamma}_i^0 = \mathbf{M}_{i\,v}^0\mathbf{v}_i^0$ represents the contribution of noise of the measurements used for the i-th feature initialization ($\boldsymbol{\gamma}_i^0$ is uncorrelated with $\mathbf{X}_s$ and $\boldsymbol{\gamma}_j^0$, $j \neq i$).

It can be shown that under the measurement model (1) and initial structure of $\mathbf{X}_i^0$ (3), the full state vector's Kalman-optimal a posteriori estimate $\left[\bar{\mathbf{X}}_s^T \quad \bar{\mathbf{X}}_{1..N}^T\right]^T$ can be obtained with $O(N)$ complexity and preserves the original structure:

$$\bar{\mathbf{X}}_i = \bar{\mathbf{M}}_{i\,s} \cdot \bar{\mathbf{X}}_s + \bar{\boldsymbol{\gamma}}_i, \quad i = 1...N\,, \tag{4}$$

where $\bar{\boldsymbol{\gamma}}_i$ is residual uncorrelated with $\bar{\mathbf{X}}_s$ and $\bar{\boldsymbol{\gamma}}_j$, $j \neq i$.

Performing the update step with $O(N)$ complexity at an arbitrary iteration of the filter requires maintaining the decomposition of $\mathbf{X}_i$ in the form similar to (3):

$$\mathbf{X}_i = \mathbf{M}_i \cdot \mathbf{X}_c + \boldsymbol{\gamma}_i, \quad i = 1...N\,, \qquad \mathbf{X}_c = \left[\mathbf{X}_s^T, \mathbf{X}_e^T\right]^T, \tag{5}$$

where $\mathbf{X}_c$ is the shared state vector (representing a common state shared across blocks of the filter), $\mathbf{X}_e$ is the shared state extension vector (additional state dimensions beyond the initial shared state $\mathbf{X}_c^0 = \mathbf{X}_s^0$), and $\boldsymbol{\gamma}_i$ is residual uncorrelated with $\mathbf{X}_c$ and $\boldsymbol{\gamma}_j$, $j \neq i$.

It can be shown that preserving (5) as exact decomposition under the dynamic model (2) requires increasing $\mathbf{X}_e$'s dimensionality by $n_q = \dim(\mathbf{w}_s)$ at each prediction step. To limit the growth of computational cost, the filter uses an approximation of $\mathbf{X}_e$ with fixed dimensionality n_e (n_e is a tunable parameter). At each prediction step, the first n_e right singular vectors of the matrix $\mathbf{A}$ of projections of the normalized feature state

This work was supported by the Ministry of Science and Higher Education of Russia (State Assignment No. FENU-2024-0004, dated 17.01.24).

vectors onto the subspace $H_{\mathbf{d}} = H_{\mathbf{X}_c} \backslash H_{\breve{\mathbf{X}}_s}$ ($\mathbf{d}$ is an orthonormal basis of $H_{\mathbf{d}}$) are taken as $\mathbf{X}_e$:

$$\mathbf{A}\cdot\mathbf{d}=\mathcal{P}_{H_{\mathbf{d}}}\left(\left[\tilde{\mathbf{X}}_1^T \ldots \tilde{\mathbf{X}}_N^T\right]^T\right)=\left[\mathbf{U}_e,\mathbf{U}_{\breve{\mathbf{d}}}\right]\begin{bmatrix}\Sigma_e & \\ & \Sigma_{\breve{\mathbf{d}}}\end{bmatrix}\left[\mathbf{V}_e,\mathbf{V}_{\breve{\mathbf{d}}}\right]^T, \quad (6)$$

$$\breve{\mathbf{X}}_e = \mathbf{V}_e\mathbf{d}, \quad \breve{\mathbf{d}} = \mathbf{V}_{\breve{\mathbf{d}}}\mathbf{d},$$

where $H_{(.)}$ denotes a subspace in the Hilbert space generated by a random variable (.), $\mathcal{P}_{(.)}$ is the projection operator onto the subspace (.).

The projections of $\mathbf{X}_i$ onto $H_{\breve{\mathbf{d}}}$ are included in $\boldsymbol{\gamma}_i$. When n_e is chosen large enough so that $\mathrm{trace}\left(\Sigma_{\breve{\mathbf{d}}}\right) << \mathrm{trace}\left(\Sigma_e\right)$, (5) will be approximately satisfied (principal component analysis method). The computational complexity of the prediction step is $O\left(N\cdot[n_q+n_e]^2\right)$; the complexity of the update step is $O\left(N\cdot[n_s+n_e]^2\right)$, where $n_s = \dim\left(\mathbf{X}_s\right)$.

IV. Implementation and Perfomance

An LD-factorized version of the FBKF (FBKF-LD) was implemented in C++ using the Eigen 3.4.0 and OpenBLAS 3.29 libraries. During each iteration, the FBKF-LD updates block-specific covariance matrix factors $\mathbf{L}_c$, $\mathbf{D}_c$, $\mathbf{L}_{\boldsymbol{\gamma}_i}$, $\mathbf{D}_{\boldsymbol{\gamma}_i}$, $\mathbf{M}_i$, calculates correction terms for $\mathbf{X}_s$ and $\mathbf{X}_i$, $i = 1...N$. For orthogonalization, the filter uses MWGS-LD (lower-triangular modified weighted Gram-Schmidt method) [4].

The trade-off between FBKF speed and accuracy is adjustable via n_e selection. With n_e = 12 the FBKF demonstrated a relative divergence from the standard KF

$$c_x \triangleq \left\|\hat{x}_{FBKF} - \hat{x}_{KF}\right\|_2 / \left\|\hat{x}_{KF} - x\right\|_2$$

of less than 0.05% during a 15-second numerical simulation of UAV descent, utilizing a MEMS-based vision-aided SINS with KLT feature tracking from onboard camera images. The average count of simultaneously tracked features was 205, while camera altitude decreased from 20 m to 2 m.

Figure 1 shows the computation time per filter step for N = 200 features with varying numbers of parallel threads. The standard KF-LD required 11.9×10^5 μs for the update step and 5.4×10^5 μs for the prediction step (without parallelization), making it two orders of magnitude slower than the FBKF-LD. The dimensions of the system vectors were: n_s = 15, n_q = 12, n_e = 12, n_{X_i} = 3, n_{Y_i} = 2.

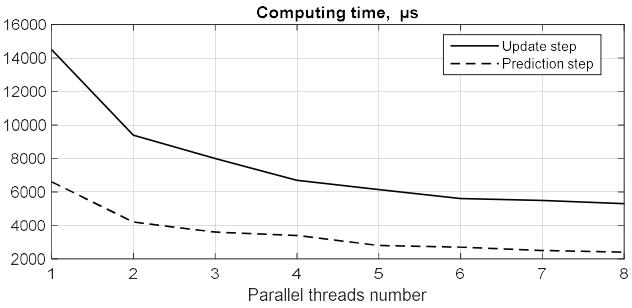


Fig. 1. Computing time for N = 200 of the FBKF-LD with n_e = 12

Conclusion

The report presents the FBKF-LD algorithm. Multi-core processor implementation demonstrated its suitability for real-time applications with hundreds of tracked features; for N = 200, the filter achieves estimation frequencies exceeding 100 Hz.

References

[1] A. I. Mourikis and S. I. Roumeliotis, “A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation,” in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), Rome, Italy, 2007, pp. 3565-3572. DOI: 10.1109/ROBOT.2007.364024.

[2] 2 G. Huang, “Visual-Inertial Navigation: A Concise Review,” in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), Montreal, QC, Canada, 2019, pp. 9572-9582. DOI: 10.1109/ICRA.2019.8793604.

[3] 3 N. I. Tsiopliakis, “Block Kalman filter with linear computational complexity for inertial navigation systems complexed with computer vision,” Vestnik of the South Ural State University. Series: Computer Technologies, Control, Radio Electronics, vol. 24, no. 4, pp. 43-56, 2024. DOI: 10.14529/ctcr240404.

[4] 4 Yu. V. Tsyganova and M. V. Kulikova, “On modern orthogonalized algorithms for optimal discrete filtering,” Vestnik of the South Ural State University. Series. Mathematical Modeling and Programming, vol. 11, no. 4, pp. 5-30, 2018. DOI: 10.14529/mmp180401.

Numerical Stability of Various Discrete Kalman Filter Implementations When Solving Ill-Conditioned Problem

M.V. Kurdeneva

Navigation Algorithms Department
PJSC "Perm Scientific-Industrial Instrument Making Company"
Perm, Russia
mariya-kurd@mail.ru

D.Kh. Mukhametzyanov

Navigation Algorithms Department
PJSC "Perm Scientific-Industrial Instrument Making Company"
Perm, Russia
damirmyhamet@gmail.com

***Abstract* — Different discrete Kalman filter implementations are considered: scalarized Kalman filter, Potter square root filter, Thornton-Bierman filter. Analysis of the accuracy and the numerical stability of these implementations are made in a solving ill-conditioned problem.**

Keywords — discrete Kalman filter, discrete Kalman filter implementations, numerical stability, roundoff errors, ill-conditioned problem

I. INTRODUCTION

The discrete Kalman filter (KF) is one of the main tools for obtaining an optimal RMS unbiased linear estimate. Due to its convenience and ease to use, this filter is widely applied in various fields of activity: from digital image processing to navigation of moving objects. However, the use of the discrete KF for solving applied problems can lead to divergence of the filter estimate because of its instability to computational roundoff errors. This disadvantage of the KF is especially critical when the problem is ill-conditioned. Therefore, in practice, more numerically stable implementations of the discrete KF should be used. In this paper, accuracy and numerical stability of various KF implementations in a solving ill-conditioned problem are studied.

II. STATEMENT OF THE PROBLEM

An n-dimensional discrete linear dynamic system is given

$$x_i = \Phi_i x_{i-1} + \Gamma_i w_i \tag{1}$$

and there are m-dimensional measurements

$$y_i = H_i x_i + v_i, \tag{2}$$

where x_i is the n-dimensional state vector (x_0 is the centered n-dimensional initial vector with a covariance matrix P_0); Φ_i is the state transition matrix; Γ_i is the process noise matrix; w_i is the p-dimensional process noise (considered centered discrete white noise with a covariance matrix Q_i); H_i is the observation matrix; v_i is the m-dimensional measurement white noise with a covariance matrix R_i.

The discrete KF [1] has been used to estimate the state vector x from measurements (2):

$$\begin{aligned}
\hat{x}_{i/i-1} &= \Phi_i \hat{x}_{i-1}, \\
P_{i/i-1} &= \Phi_i P_{i-1} \Phi_i^T + \Gamma_i Q_i \Gamma_i^T, \\
\hat{x}_i &= \hat{x}_{i/i-1} + K_i (y_i - H_i \hat{x}_{i/i-1}), \\
K_i &= P_{i/i-1} H_i^T (H_i P_{i/i-1} H_i^T + R_i)^{-1}, \\
P_i &= (E - K_i H_i) P_{i/i-1},
\end{aligned} \tag{3}$$

where $\hat{x}_{i/i-1}$ and $\hat{x}_i$ are the predicted state estimate and the actual state estimate, respectively; K_i is the Kalman gain matrix; $P_{i/i-1}$ and P_i are covariance matrices of the prediction errors and estimate errors, respectively; E is the identity matrix.

As noted in many papers, e.g. in [2, 3, 4], accuracy of the algorithm (3) reduces due to accumulation of computational roundoff errors.

III. IMPROVING NUMERICAL STABILITY

Computational roundoff errors are inherent to computer architecture and discrete representation of numbers and they cannot be eliminated at the hardware level. Therefore, the main way to improve the numerical stability of the discrete KF is to modify it mathematically.

A. Scalarized mesurment vector processing

One of such mathematical modifications is the scalarized processing of the measurement vector [4]. It provides sequential processing of each state vector element at each KF iteration.

This mathematical modification eliminates the computationally complex direct calculation of the inverse matrix from the algorithm of discrete KF.

B. Factorization of covariance matrices

Another mathematical modification that can be used for the discrete KF numerical stability improvement is the factorization of covariance matrices. Instead of covariance matrices, their factors are used in the discrete KF algorithm.

Note that there are different ways to factorize matrices. One of them is the square root decomposition, based on which James E. Potter proposed his own discrete KF implementation in 1963 [5]. Another variant of matrices factorization is the UD-decomposition (decomposition into upper triangular and diagonal matrices). The UD-decomposition is a modification of the square root decomposition but without the square root extraction that

saves computational efforts. Gerald J. Bierman was the first who proposed the UD-decomposition for the discrete KF [6].

Factorization of covariance matrices increases the calculation accuracy when the limited number of bits is used to represent numbers in the PC [7].

Moreover, this mathematical modification prevents the loss of symmetry and nonnegative definiteness by covariance matrices [7].

It is important that during the KF algorithm calculation covariance matrices remain nonnegative definite. As shown in [2], the loss of nonnegative definiteness by the covariance matrix of prediction errors can lead to Kalman gain matrix changing its sign and finally to divergence of the filter estimate. And even if in the course of further calculations the Kalman gain matrix takes the correct sign, the estimate will converge more slowly than before.

IV. Discrete Kalman Filter Implementations

Based on mathematical modifications mentioned above, various discrete KF implementations were developed. Descriptions of some of them are given in [2, 4, 8].

In this paper, the following discrete Kalman filter implementations were considered:

- KF – discrete KF with the direct matrix inversion;
- SKF – scalarized KF;
- PSRF – Potter square root filter using Cholesky decomposition algorithm for covariance matrices factorization;
- MPSRF – Potter square root filter using Gram-Schmidt orthogonalization process to factorize the covariance matrix of prediction errors and explicitly calculate the square root of the covariance matrix of estimate errors;
- TBF – Thornton-Bierman filter using Cholesky *UD*-decomposition algorithm for covariance matrices factorization;
- MTBF – Thornton-Bierman filter using weighted Gram-Schmidt orthogonalization process to factorize the covariance matrix of prediction errors and explicitly calculate the *UD*-factorization of the covariance matrix of estimate errors.

V. Ill-Conditioned Problem

Problem conditioning refers to the sensitivity of its solution to variations of input parameters [2]. The problem is considered ill-conditioned, if small variations in initial conditions influence strongly the final result, otherwise it is a well-conditioned problem. Problem conditioning depends on input parameters errors and applied solution procedure.

According to [2], main factors that affect conditioning of the optimal filtration problem with the help of the discrete KF are the following:

- uncertainty of the inverse matrix that occurs during calculation of the discrete KF algorithm;
- large KF parameter matrix;
- limited number of bits used to represent numbers in the PC;
- poor system observability.

To analyze the accuracy and the numerical stability of different discrete KF implementations the problem (1) and (2) with the following parameter matrices:

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1+\delta & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} \delta^2 & 0 & 0 \\ 0 & \delta^2 & 0 \\ 0 & 0 & \delta^2 \end{bmatrix},$$
$$\Phi = E_4,\ \Gamma = E_4,\ P_0 = E_4,\ Q = O_4, \tag{4}$$
$$x_0 = \left[1, 1, -1, 1\right]^{\mathrm{T}}, \quad y_1 = Hx_0 + \nu.$$

where E_4 and O_4 – identity and zero matrices of 4-th order; $\delta^2 < \varepsilon_{roundoff}$, $\delta > \varepsilon_{roundoff}$, $\varepsilon_{roundoff}$ – parameter of computational roundoff, i.e. $\varepsilon_{roundoff} + 1 \neq 1$, but $\varepsilon_{roundoff} / 2 + 1 \overset{roundoff}{=} 1$.

We will show that the conditioning of this problem depends on the value of the parameter δ, which is connected with the parameter of computational roundoff. To do that, let us find the measurement residual covariance matrix:

$$HP_{1/0}H^T + R = H(\Phi P_0 \Phi^T + \Gamma Q \Gamma^T)H^T + R =$$
$$= HH^T + R = \begin{bmatrix} \delta^2 + 4 & \delta + 4 & 4 \\ \delta + 4 & 2\delta^2 + 2\delta + 4 & \delta + 4 \\ 4 & \delta + 4 & \delta^2 + 4 \end{bmatrix}. \tag{5}$$

Determinant of this matrix equals to

$$\left| HP_{1/0}H^T + R \right| = 2\delta^4(\delta^2 + \delta + 9). \tag{6}$$

Note that the matrix inverse to the matrix (5) is included into the expression of the discrete KF gain, so it is important this matrix not to be degenerate. However, in conditions of computational roundoff, when $\delta \to \varepsilon_{roundoff}$, the determinant (6) tends to zero, therefore, the matrix (5) tends to a degenerate matrix that results in calculation errors in the KF algorithm.

The analytic expression for the covariance matrix of estimate errors after the first iteration of the discrete KF has the form:

$$P_1 = (E - K_1 H)P_{1/0} = \frac{1}{2(\delta^2 + \delta + 9)} \cdot$$
$$\cdot \begin{bmatrix} 2\delta^2 + 2\delta + 13 & -5 & -(\delta + 3) & -5 \\ -5 & 2\delta^2 + 2\delta + 13 & -(\delta + 3) & -5 \\ -(\delta + 3) & -(\delta + 3) & \delta^2 + 9 & -(\delta + 3) \\ -5 & -5 & -(\delta + 3) & 2\delta^2 + 2\delta + 13 \end{bmatrix}. \tag{7}$$

Note that this matrix does not tend to the degenerate matrix when $\delta \to \varepsilon_{roundoff}$, so it can be used as a reference for evaluation of accuracy of the considered discrete KF implementations.

VI. Estimation of the Accuracy and Numerical Stability of Kalman Filter Implementations

A. The criterion for evaluating the accuracy of Kalman filter implementations

To compare the accuracy of various discrete KF implementations the relative root mean square error (RMSE) of calculation of the covariance matrix of estimate errors after the first KF iteration was introduced:

$$\Delta P^a = \sqrt{\frac{\sum_{i=1}^{3}\sum_{j=1}^{3}\left(p_{ij}^a - p_{ij}\right)^2}{\sum_{i=1}^{3}\sum_{j=1}^{3}p_{ij}^2}}, \quad (8)$$

where p_{ij}^a – element of the covariance matrix of estimate errors P_1^a, that was received by one of the aforementioned discrete KF implementations; a – index denoting the KF implementation chosen for computing, $a \in \{KF, SKF, PRSF, MPRSF, TBF, MTBF\}$; p_{ij} – element of the analytically received covariance matrix of estimate errors P_1; i, j – matrix row and column numbers.

The value ΔP^a shows how accurate is the calculation of the covariance matrix of estimate errors with the help of implementation a of the discrete KF, i. e. the difference between the numerically and analytically obtained value of this matrix. For brevity, we will call this value "accuracy of the implementation a".

B. Results

As can be seen from the Fig. 1, the accuracy of the discrete KF with the direct matrix inversion (KF implementation) decreases as the problem conditioning worsens.

Potter square root filter (MPSRF implementation) and Thornton-Bierman filter (MTBF implementation) have shown the best results. Relative RMSE of calculation of covariance matrix of estimate errors after the first KF iteration in these implementations grow slower when $\delta \to \varepsilon_{roundoff}$ compared with other KF implementations. For example, at $\delta = 10^{-8}$, the accuracy of MPSRF and MTBF is in the order of 10^{-8}-10^{-9}, while other discrete KF implementations show the accuracy approximately 10^{-1}. Thus, they will lose their efficiency when $\delta \le 10^{-8}$. MPSRF and MTBF remain numerically stable to the value $\delta = \varepsilon_{roudoff}$.

These results are consistent with the results obtained for another ill-conditioned problem given in [2] with addition of analysis of behavior of SKF, MPSRF and TBF implementations.

VII. Conclusion

In the course of the study, the accuracy and the numerical stability of various discrete KF implementation were analyzed. The numerical simulation has shown that the discrete KF with the direct matrix inversion (KF implementation) is unstable to computational roundoff errors, therefore it is not recommended for solution of practical problems. Potter square root filter (MPSRF implementation) and Thornton-Bierman filter (MTBF implementation) demonstrated greater numerical stability compared to other considered implementations of the discrete KF.

It is of interest to study further the behavior of various discrete KF implementations in solving practical ill-conditioned problems and navigation problems with a poor observability.

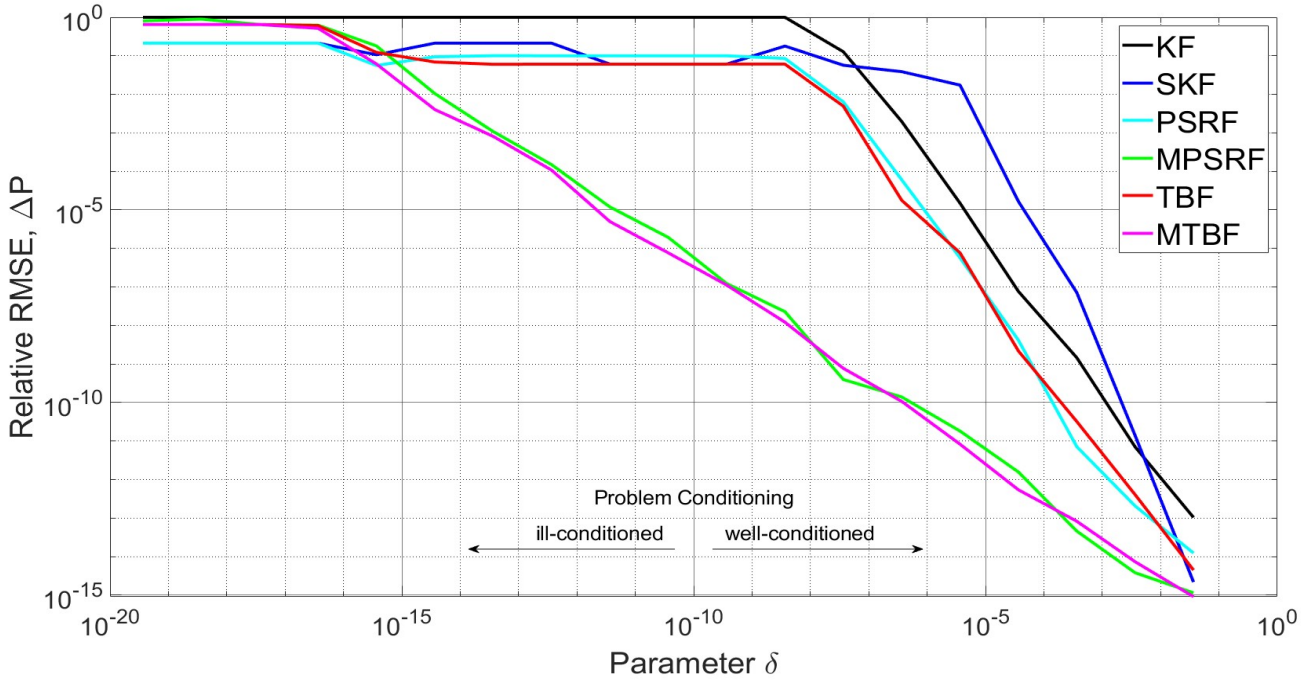


Fig. 1. Dependence of the relative RMSE from the problem conditioning in calculation of covariance matrix of estimate errors

REFERENCES

[1] O.V. Stepanov, Osnovy teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoi informatsii (Fundamentals of the Estimation Theory with Applications to the Problems of Navigation Information Processing), Part 1, Vvedenie v teoriyu otsenivaniya (Introduction to the Estimation Theory). St. Peterburg: Concern CSRI Electropribor, 2017. 509 p.

[2] M.S. Grewal, A.P. Andrews, Kalman Filtering: Theory and Practice. – New Jersey: Prentice-Hall, 2001. 410 p.

[3] M. Verhaegen, P. Van Dooren, Numerical Aspects of Different Kalman Filter Implementations. IEEE TRANSACTIONS ON AUTOMATIC CONTROL, vol. 31, №10, 1986, pp. 907-917.

[4] I.V. Semushin, Vychislitel'nye metody algebry i otsenivaniya (Computational Methods of Algebra and Estimation). Ulyanovsk: UlGTU, 2011. 366 p.

[5] J.E. Potter, R.G. Stern, Statistical Filtering of Space Navigation Mesurments. Proceedings of 1963 AIAA Guidance and Control Conference. New York: AIAA, 1963.

[6] G.J. Bierman, Factorization Methods for Discrete Sequential Estimation. New York: Dover Publications. 1977. P. 241.

[7] N.A. Parusnikov, A.A. Golovan, Matematicheskie modeli bortovykh algoritmov integrirovannykh inertsial'nykh navigatsionykh sistem (Mathematical Models of Onboard Algorithms of Integrated Inertial Navigation Systems). Moskow, 2005. 124 p.

[8] I.V. Semushin, Yu.V. Tsyganova, K.V. Zakharov, Ustoichivye algoritmy fil'tratsii – obzor i novye rezul'taty dlya sistem sudovozhdeniya. (Stable Filtering Algorithms – Overview and New Results for Navigation Systems) Informatsionnye tekhnologii i vychislitel'nye sistemy, №4, 2013, pp 90-112.

Application of Visual-Inertial Odometer for Autonomous Navigation of Unmanned Aerial Vehicle

Yevgeniy Zhirkov

development department
Kvantron Group LLC
Ryazan, Russia
jirckow.evgeny@yandex.ru

Yevgeniy Muratov

development department
Kvantron Group LLC
Ryazan, Russia
e.muratov@kvantron.com

***Abstract*—The work studies practical implementation of visual-inertial odometry for autonomous navigation of unmanned aerial vehicle based on airscape data and measurements of onboard inertial navigation subsystem. The results of algorithm operation based on the data obtained both in computer simulator and during flight tests at Protasovo airfield in Ryazan region are presented.**

Keywords— autonomous positioning, unmanned aerial vehicle, visual odometry, homography matrix, navigation, integrated navigation system.

I. INTRODUCTION

Nowadays the topic of designing autonomous navigation systems for unmanned aerial vehicles (UAVs) is more relevant than ever. One of the most efficient ways of autonomous navigation for low-altitude UAVs [1] is visual odometry, which implies the presence of optical sensor being the source of navigation information. The efficiency of visual odometry is manifested in the conditions when the effect of simulating and masking interference in radio wave range has negative effect on the reception of signals from global satellite navigation system used [2-4]. The basis for navigating with the help of visual odometer is the method that determines coordinate displacement and heading rotation of ψ carrier with nadir-oriented photo or video camera [5-7]. In this case, offset and angle of rotation are found by key points of two consecutive frames having intersecting fields of view with time interval Δt [8, 9]. The technique to determine carrier offset and angle of rotation over time interval Δt is based on estimating homography matrix **H** which correlates homogeneous pixel coordinates of key point pairs [8-11]. Flight altitude h and current orientation angles pitch θ and roll φ coming from flight controller should also be taken into account.

This work is devoted to the study of visual odometer error both on a mathematical model with high degree of abstraction, as well as on the basis of test results in flight simulator and video materials obtained by small aircraft—type UAV with nadir-oriented digital video camera.

II. RESEARCH OBJECTIVES

A. Main objective

The task to estimate the error of visual inertial odometer can be formulated as follows. When UAV is moving along a certain trajectory defined by a set of control points $\{\mathbf{A}_n\}$, its track is necessary to be restored a posteriori from the information received only from nadir-oriented camera and onboard inertial navigation subsystem, and the error (absolute and/or relative) of its own positioning to be estimated in global coordinate system. On-board inertial navigation subsystem of UAV can be considered as 4-dimensional sensor reporting vector readings (h, φ, θ, ψ). Height h of flight is reported by barometric altimeter, and angles φ, θ and ψ of current UAV are read by micro-electro-mechanical sensors of accuracy class corresponding to consumer one.

Geometric formulation of the problem to estimate coordinate displacement from two frames with intersecting fields of view using nadir-oriented camera is considered in [5].

B. Auxiliary objectives

Designing visual odometer is accompanied by solving the auxiliary tasks described below.

1) Calibration of microelectromechanical sensors [12].

2) Photogrammetric camera calibration [13].

3) Compensation of optical distortions from camera lens [13].

4) The choice of satisfactory detector and descriptor of singular points [8, 11].

5) Mutual calibration of flight controller and recording camera [14].

Separately, the problem to choose the algorithms for filtering and aggregating navigation data can be mentioned [15, 16]. Sensor synchronization of onboard inertial navigation subsystem, nadir-oriented camera and signals from global satellite navigation system are necessary to be taken into consideration.

III. BASIC ANALYTICAL EXPRESSIONS

Along with geometric formulation of problem [5], the main analytical expression is based on homography matrix **H** decomposition. Initial data for calculation is formed as a result of the following steps.

- To determine the normal to underlying surface **n** and parallel transfer vector **t**. The underlying surface is assumed to be strictly horizontal [5].
- To set matrices $\mathbf{R}_{\text{cam1}}$ and $\mathbf{R}_{\text{cam2}}$ of camera rotation relative to global coordinate system for the first and second frames, respectively.
- Knowing rotation matrices $\mathbf{R}_{\text{cam1}}$ and $\mathbf{R}_{\text{cam2}}$, to calculate matrix $\mathbf{R}_{\text{cam1}}{}^{\text{cam2}}$ of camera rotation (in camera associated coordinate system) per time Δt:

$$\mathbf{R}_{\text{cam1}}{}^{\text{cam2}} = \mathbf{R}_{\text{cam2}}\mathbf{R}_{\text{cam1}}{}^{-1}. \quad (1)$$

- Matrices $\mathbf{R}_{cam1}$ and $\mathbf{R}_{cam2}$ also allow obtaining the expressions for vector $\mathbf{n}_{cam1}$ of the normal to underlying surface in camera coordinate system before its rotation as well as for vector $\mathbf{t}_{cam2}$ of parallel transfer in camera coordinate system after its rotation:

$$\mathbf{n}_{cam1} = \mathbf{R}_{cam1}\mathbf{n}, \quad (2)$$

$$\mathbf{t}_{cam2} = \mathbf{R}_{cam2}\mathbf{t}. \quad (3)$$

- The matrix of internal camera **K** parameters as well as the matrix of camera mutual calibration and flight controller $\mathbf{R}_{INS->cam}$ follows from the solution of aforementioned auxiliary objectives list.

Thus, homography matrix **H** estimate decomposition takes the following form:

$$\mathbf{H} = \mathbf{K}\mathbf{R}_{INS->cam}\left[\mathbf{R}_{cam1}{}^{cam2} + h^{-1}\mathbf{t}_{cam2}(\mathbf{n}_{cam1})^{T}\right]\mathbf{R}_{INS->cam}{}^{-1}\mathbf{K}^{-1} \quad (4)$$

Expressions (1)–(4) are included in a generalized mathematical model with high degree of abstraction which assumes the definition of analytical expression for movement trajectory, the normal to underlying surface plane **n**, spatial coordinates of key points $\mathbf{M}_j$, a priori known internal parameters of camera (**K, R, t**), root-mean-square error of estimating pixel coordinates of singular points σ.

To estimate homography matrix by the analogy with [8, 11] the authors have chosen computationally efficient ORB method (***o****riented FAST and* ***r****otated* ***B****RIEF*): the combination of FAST feature detector (i.e., ***f****eatures from* ***a****ccelerated* ***s****egment* ***t****est*) and BRIEF descriptor (i.e., ***b****inary* ***r****obust* ***i****ndependent* ***e****lementary* ***f****eatures*) Matrix **H** in the work was estimated based on 50 best pairs of all matches found.

IV. Simulation Modelling

The authors tested successful operation of visual odometer on a generalized mathematical model with high degree of abstraction which involves specifying an analytical expression for movement trajectory [15], the normal to underlying surface plane, spatial coordinates of key points, a priori known internal and external camera parameters, and root-mean-square error in estimating pixel coordinates of key points.

The flight of multicopter aircraft was simulated in *Microsoft Airsim* simulator. The flight was simulated during daylight hours, in the absence of fog and haze. The flight trajectories were "vertical takeoff, 300 m straight flight, 180° turn, 300 m straight flight, vertical landing". Flight time was 2 min 10 sec. Frame processing time was Δt = 0.5 sec. Absolute error of visual-inertial odometer estimation was less than 25 m at the time of landing. Thus, relative positioning error (4...5%) is generally comparable with the results of trajectory reconstruction for ORB method given in study [8] (taking into account doubled value of root-mean-square error). The authors should note that in more than 95% of cases, stable correct matching for a simulator is true for all pairs of key points.

V. Results of Real Flight Data Processing

Real flights were performed in Ryazan region over Protasovo airfield territory in the conditions similar to a simulator: daylight and no fog or haze. The trajectory was seen as "8"-shaped three-dimensional curve (see Fig. 1).

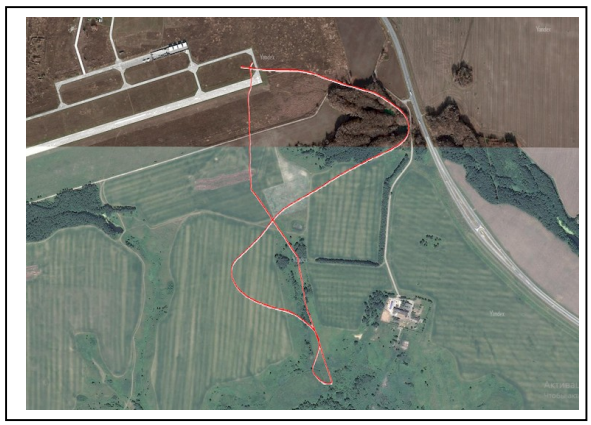


Fig. 1. Top view of UAV trajectory over Protasovo airfield, Ryazan region.

The flights were performed on small aircraft-type UAV with *iNav* flight controller. GoPro action video camera with resolution of 1920×1080 pixels was used as nadir-oriented camera. Average time of each flight was 6.5 min ± 20 s; average length of flight path was 6 km ± 0.5 km, angular velocity during turns was no more than 50 °/s. On real flight data, ORB singular point detector has absolute error more than an order of magnitude higher than in a simulator: its value is in the range from 430 to 900 m, it being explained, first of all, by much lower probability to correctly find pairs of matches, their non-uniform density, and the error in estimating instantaneous values of orientation angles based on the signals from inertial sensors in flight controller.

VI. Conclusion

According to the results obtained in a simulator and in real flight tests, it becomes obvious that classical methods [9, 10] of special points detecting do not give a relative error of visual-inertial odometer of the order of percent units. In this regard, the preferred option to develop the work is to justify detecting the pairs of special points that are robust to scale and direction of shooting based on machine learning.

Acknowledgments

The authors would like to express their gratitude to the team of Kvantron Group LLC for participating in the development of the project on autonomous navigation of unmanned aerial vehicle.

References

[1] Miroshnichenko A.P., Mukhin I.E. Methods of navigation using airscape data for small UAVs. Bulletin of the South-West State University. Series: Management, Computer Science, Informatics. Medical Instrumentation. 2023. Vol. 14. № 3. P. 144–156.

[2] Peshekhonov V.G. High-precision navigation without using information from global navigation satellite systems. Gyroscopy and navigation. 2022. Vol. 30. № 1. P. 3–11.

[3] Makarenko S.I. Counteraction to unmanned aerial vehicles. St. Petersburg: Science-intensive technologies, 2020. 204 p.

[4] Kashcheev A.A., Koshelev V.I. Evaluation of efficient suppression of satellite radio navigation systems signals by intentional interference. Journal of Radio Electronics. 2012. № 7. P. 1–12.

[5] Conte G., Doherty P. An integrated UAV navigation system based on aerial image matching. Proc. IEEE Aerospace Conference. 2008. P. 1–10.

[6] Lu Y., Xue Z., Xia G.-S., Zhang L. A survey on vision-based UAV navigation. Geo-spatial information science. 2018. Vol. 21. No. 1. P. 21–32.

[7] Yang T., Ren Q., Zhang F., Xie B., Ren H., Li J., Zhang Y. Hybrid camera array-based UAV auto-landing on moving UGV in GPS-denied environment. Remote Sensing. 2018. Vol. 10. P. 1829.

[8] Yang G., Meng W.-D., Dong H.-G., Feng N.-N. Using image keypoints and lines in real-time visual-inertial odometry. Gyroscopy and Navigation. 2023. Vol. 31. № 4. P. 96–117.

[9] Efimov A.I., Novikov A.I. Algorithm for step-by-step refinement of projective transformation for image alignment. Computer Optics. 2016. Vol. 40. № 2. P. 258–265.

[10] Szeliski R. Image alignment and stitching: a tutorial. Foundations and trends in computer graphics and vision. 2006. Vol. 2. No. 1. P. 1–104.

[11] Ali B., Sadekov R.N., Tsodokova V.V. Algorithms for navigation of unmanned aerial vehicles using technical vision systems. Gyroscopy and navigation. 2022. Vol. 30. № 4. P. 87–105.

[12] Tikhomirov V.V., Dzuev A.A., Golikov V.P., Trebukhov A.V. Calibration of SINS with a block of inertial sensors fixed on shock absorbers. Gyroscopy and navigation. 2019. Vol. 27. № 1. P. 33–46.

[13] Luhmann T., Robson S., Kyle S., Boehm J. Close-range photogrammetry and 3D imaging: 2nd edition. Berlin: De Gruyter, 2013. 684 p.

[14] Fleps M., Mair E., Ruepp O., Suppa M., Burschka D. Optimization based IMU camera calibration. Proc. IEEE/RSJ international conference on intelligent robots and systems. San Francisco, 2011. P. 3297–3304.

[15] Stepanov O.A. Integrated inertial-satellite navigation systems. Gyroscopy and navigation. 2002. № 1. P. 23–45.

[16] Al-Mansur M., Shuaib I., Jafar A., Potapov A.A. Analytical algorithm to estimate spatial position and heading of an object. Gyroscopy and Navigation. 2019. Vol. 27. № 1. P. 72–92.

On Motion Control in the Problem of Spacecraft Inspection

M. Klyushin

Department of Theoretical and Applied Mechanics
Saint Petersburg State University
Saint Petersburg, Russia
m.klyushin@spbu.ru

***Abstract*— Spacecraft formation flying in the gravitational and magnetic fields of the Earth is discussed in the report. Spacecraft formation is assumed to consist of uncharged target spacecraft and an electrically charged spacecraft-pursuer. The orbital motion control of a spacecraft-pursuer in the problem of a target spacecraft inspection is studied. Based on analysis of the system of differential equations of motion, a control algorithm for changing the distance between the spacecraft is proposed. Computer modelling of the considered problem is conducted, demonstrating efficiency of the proposed control method.**

Keywords—satellite formation, geomagnetic field, motion control, electric charge

I. Introduction

The rapid development of aerospace industry and the simultaneous desire to reduce the cost of space missions have led to the creation of small spacecraft that do not require significant time and financial costs for their production. The use of small spacecraft formation allows to significantly expand the range of problems to be solved, and also leads to the emergence of new research areas in mechanics and control theory. However, it is important to maintain the mutual position of spacecraft in order to ensure proper communication and interaction between them.

The size of small spacecraft imposes certain restrictions on systems of motion control. For instance, the use of jet engines is limited by the small reserves of fuel on board. Therefore, researchers have turned to developing new methods for controlling the orbital motion of a spacecraft based on interaction with natural sources of forces [1-3]. This report is devoted to orbital motion control of a spacecraft using the Lorentz force [4-7]. The control algorithm is applied to the problem of a target spacecraft inspection using a pursuer spacecraft.

II. Mathematical Modelling

A. The Problem Statement

The motion of a spacecraft formation, that consists of an uncharged target spacecraft and an electrically charged spacecraft-pursuer, in the gravitational and magnetic fields of the Earth is considered. The geomagnetic field is modelled as "direct magnetic dipole". It is assumed that the both spacecraft have a mass of m. The position of a target spacecraft's center of mass with respect to the coordinate system $Oxyz$, whose origin is at the center of the Earth, is given by the radius-vector $\boldsymbol{r}_C$. The position of a pursuer spacecraft's center of mass with respect to the same coordinate system is given by the radius-vector $\boldsymbol{r}_L$ (Fig. 1).

The work is supported by the Russian Science Foundation (grant No. -41-02031, https://rscf.ru/project/24-41-02031/)

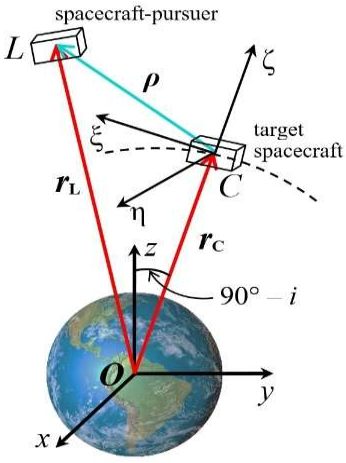


Fig. 1. Pursuer spacecraft and target spacecraft

The target spacecraft is assumed to move in a circular Keplerian orbit with radius r_C and with arbitrary inclination i, with angular velocity of the orbital motion ω_0. The right-hand orbital frame $C\xi\eta\zeta$ [8], with origin at the target spacecraft's center of mass, is considered. The $C\zeta$-axis is directed along the radius-vector $\boldsymbol{r}_C$, and the $C\eta$-axis is directed normal to the orbit plane. The position of a pursuer spacecraft's center of mass with respect to the orbital frame is given by the radius-vector $\boldsymbol{\rho}$. In this problem statement, the target spacecraft is only affected by the gravitational force. And the charged pursuer spacecraft is affected by the gravitational and the Lorentz forces.

B. Mathematical Model

Based on the application of the Newton's second law and the Transition theorem, one can obtain the following system of differential equations, describing the relative motion of an electrically charged pursuer spacecraft:

$$\begin{aligned}
\ddot{\xi} &= -\frac{\mu}{r_L^3}\xi - 2\omega_0\dot{\zeta} + \omega_0^2\xi + a_{L\xi},\\
\ddot{\eta} &= -\frac{\mu}{r_L^3}\eta + a_{L\eta},\\
\ddot{\zeta} &= -\frac{\mu}{r_L^3}\left(r_C + \zeta\right) + \frac{\mu}{r_C^2} + 2\omega_0\dot{\xi} + \omega_0^2\zeta + a_{L\zeta}.
\end{aligned} \tag{1}$$

Here $a_{L\xi}$, $a_{L\eta}$, $a_{L\zeta}$ are the components of the Lorentz accelaration, which is defined by the following formula: $\boldsymbol{a}_L = (q/m)(\boldsymbol{v}_r \times \boldsymbol{B})$. Here $\boldsymbol{v}_r$ is the pursuer spacecraft velocity with respect to the Earth and $\boldsymbol{B}$ is the magnetic induction vector of the geomagnetic field [7].

Let us introduce the dimensionless variables $\xi_1 = \xi/r_C$, $\eta_1 = \eta/r_C$, $\zeta_1 = \zeta/r_C$ and perform the linearization of system (1). As a result, one can obtain the system of Hill-Clohessy-Wiltshire-type equations [9, 10], generalizing these equations taking into account the Lorentz force:

$$\xi_1'' = -2\zeta_1' + a^1_{L\xi},$$
$$\eta_1'' = -\eta_1' + a^1_{L\eta}, \quad (2)$$
$$\zeta_1'' = 3\zeta_1 + 2\xi_1' + a^1_{L\zeta}.$$

Here $\xi_1' = d\xi_1/d\tau$, where $\tau = \omega_0 t$ is the dimensionless time. The same is for variables η_1 and ζ_1. Here $a^1_{L\xi}$, $a^1_{L\eta}$, $a^1_{L\zeta}$ are the components of the Lorentz accelaration, which are written in terms of the dimensionless variables and are linearized.

During the analysis of the linearized system (2), the following simple law for controlling the electric charge of the pursuer spacecraft is proposed. This control allows to change the relative distance to the target spacecraft:

$$q(t) = \pm q_0 \cdot \cos(\omega_0 \cdot t), \quad (3)$$

where q_0 is the amplitude of the charge change. Law (3) with "+" sign leads to a decrease in the relative distance between the spacecraft. Law (3) with "–" sign corresponds to the law with "+" sign with a phase shift of π and leads to an increase in the relative distance between the spacecraft.

III. Inspection of Target Spacecraft

Based on the charge change law (3), a feedforward control algorithm is proposed to change the distance between the spacecraft. The application of this algorithm is considered based on the problem of a target spacecraft inspection using a pursuer spacecraft. As a part of the considered problem, it is required to perform the spacecraft rendezvous maneuver. And, as a result, transition of the pursuer spacecraft to new relative orbits is also required. It is suggested to perform this maneuver in three stages:

1. Motion of a pursuer spacecraft without electrical charge along the initial relative orbit, which is an ellipse with a semi-major axis a. The ellipse is assumed to be in the plane of the target spacecraft's orbit. The motion occurs during k_1 rotations of the target spacecraft around the Earth.
2. "Switching on" the charge and it's changing according to law (3) at the end of k_1-th rotation. The motion occurs during $k_2 - k_1$ rotations of the target spacecraft around the Earth.
3. Neutralization of the charge at the end of k_2-th rotation, and the motion of the pursuer spacecraft along a new relative orbit.

As an illustration of the algorithm's operability, Fig. 2-3 show the analytically obtained trajectories for the parameters in Table 1.

Analysis of the plots in Fig.2-3 shows that in a short time (about 15 hours), it is possible to significantly change the pursuer spacecraft's relative orbit size. It should be noted that the proposed control does not require additional control systems.

TABLE I. Simulation parameters

Parameter	m, kg	r_C, km	ω_0, rad/s	i,°	q_0, C	a, km	k_1	k_2
Value	20	6900	0.0011015	40	0.001	0.5	5	15

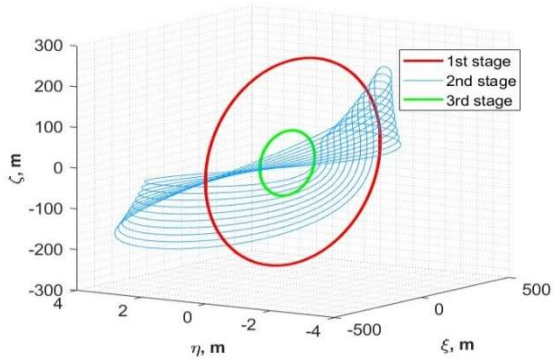


Fig. 2. Rendezvous maneuver

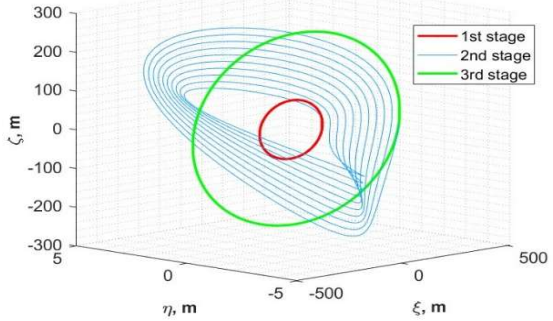


Fig. 3. Transfer of a pursuer spacecraft to a higher relative orbit

IV. Conclusion

As a result of this research, a feedforward control algorithm, based on changing the electrical charge according to the simple harmonic law, was proposed to change the distance between a pursuer and a target spacecraft. Analytical and numerical modelling demonstrated the operability of this control method. The control algorithm proposed can be used for problems of a target spacecraft inspection, that require changing the distance between a pursuer spacecraft and the target.

References

[1] D. Ivanov, U. Monakhova, A. Guerman, and M. Ovchinnikov, "Decentralized Control of Nanosatellite Tetrahedral Formation Flying Using Aerodynamic Forces," Aerospace, 2021. Vol. 8, p. 199.

[2] Y. Mashtakov, M. Ovchinnikov, T. Petrova, and S. Tkachev, "Two-satellite formation flying control by cell-structured solar sail," Acta Astronautica, 2020. Vol. 170, pp. 592–600.

[3] M. Shouman, M. Bando, and Sh. Hokamoto, "Controllability Analysis of Propellant-Free Satellite Formation Flight," Journal of Guidance, Control, and Dynamics, 2021. Vol. 44, pp. 1–11.

[4] G. Pollock, J. Gangestad, and J. Longuski, "Analytical solutions for the relative motion of spacecraft subject to Lorentz-force perturbations," Acta Astronautica, 2011. Vol. 68, pp. 204–217.

[5] X. Huang, Y. Yan, and Y. Zhou, "Optimal Lorentz-augmented spacecraft formation flying in elliptic orbits," Acta Astronautica, 2015. Vol. 111, pp. 37–47.

[6] M. Klyushin, A. Tikhonov, and D. Giri, "Orbital motion control of an electrically charged spacecraft," Acta Astronautica, 2025. Vol. 226, pp. 626–636.

[7] M. Klyushin and A. Tikhonov, "The Lorentz force-based control of a spacecraft formation flying using the harmonic charging law," Acta Astronautica, 2025. Vol 232, pp. 68–79.

[8] A. Tikhonov, "A method of semipassive attitude stabilization of a spacecraft in the geomagnetic field," Cosmic Research, 2003. Vol. 41, pp. 63–73.

[9] G. Hill, "Researches in the Lunar theory," American Journal of Mathematics, 1878. Vol. 1, pp. 5–26.

[10] W. Clohessy and R. Wiltshire, "Terminal guidance systems for satellite rendezvous," Journal of the Aerospace Sciences, 1960. Vol. 674, pp. 653–658

Estimating Torques for a Cubesat 6U Satellite with Deployed Solar Pannels

Polina Yakovleva

Inter University department of space research
Samara National Research University
Samara, Russia
polina.ya03@yandex.ru

Aleksey Kumarin

Inter University department of space research
Samara National Research University
Samara, Russia
alky_samara@mail.ru

Modeling angular motion of CubeSats nanosatellites usually considers gravitational and aerodynamic torques while the torque of sunlight pressure is one of the most underrated. Although it can have significant role in angular motion of a spacecraft, especially satellite with large deployable elements such as solar pannels. In this paper we develop models and software to estimate the torque of sunlight pressure and compare it with gravitational and aerodynamic torques for CubeSat 6U nanosatellite with deployed solar panels at low Earth orbits.

Keywords—CubeSat, nanosatellite, angular motion, sunlight pressure, passive orientation.

I. INTRODUCTION

According to statistics [1] there is a big increase of popularity of nanosatellites in last decade.

Nanosatellite formats like CubeSat 6U and CubeSat 3U are the most popular CubeSat configurations [1]. Having launch sizes not exceeding 20x10x30 centimeters and 10x10x30 centimeters respectively and relatively low cost they became an excellent base for science researches and commercial missions. CubeSat 6U is one of the best platforms for implementing scientific missions. Its payload capability to overall size ratio exceeds such values of popular CubeSat 3U platforms.

Despite all the interest there are not many angular motion models intended particularly for nanosatellites. Therefore, many assumptions are used for CubeSat motion modelling and only aerodynamical and gravitation torques impact is usually considered [2]. Thus, different inner and outer force's effect on nanosatellite motion is not accounted. However, they can affect its dynamics significantly.

Unaccounted torques include torques from moving masses inside the satellite; additional magnetic torques caused spacecraft magnetization and also sunlight pressure torque caused by reflection of solar particles from satellite's surface [3].

One of the most important, but underrated and usually unaccounted torque is sunlight pressure torque. Its value can be close or even greater than gravitation torque, as it will be further shown in this paper.

Accounting of sunlight pressure torque in the beginning of nanosatellite's mission development can help to gain a better sense about modes of CubeSat motion. It can be really useful for development of motion control system requirements.

This work was supported by the Russian Science Foundation, project no. 23-67-10007. https://rscf.ru/project/23-67-10007/

II. COORDINATE SYSTEMS

The math model of torques in this paper includes a number of coordinate systems. Orbital coordinate system OXYZ is shown in Fig. 1. In this model the origin point coincides with CubeSat's center of mass, axis OZ is directed along the satellite's radius vector, axis OX is pointed along the satellite's orbital motion vector and axis OY complements this system to the right Cartesian coordinate system.

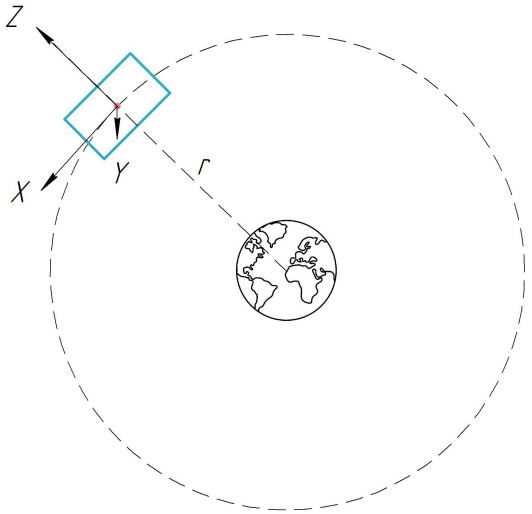


Fig. 1. Orbital coordinate system

In this paper all modelling is performed for CubeSat 6U nanosatellite with deployable solar panels perpendicular to the edge of the biggest side. The nanosatellite's model is shown in Fig. 2.

The relation between orbital coordinate frame and body-fixed coordinate frame Oxyz is determined using the Euler angles.

The first rotation is between orbital (OXYZ) and velocity ($OX_aY_aZ_a$) frames by angle ψ around OX axis.

The second rotation is between velocity ($OX_aY_aZ_a$) and intermediate ($OX_nY_nZ_n$) frames by angle α_n around OY_a axis.

The third rotation is between intermediate ($OX_nY_nZ_n$) and body-fixed (Oxyz) frames by angle ϕ around OX_n axis.

The relative rotations of the coordinate frames are shown in Fig. 2.

The transition matrix from orbital coordinate frame OXYZ to body-fixed coordinate frame Oxyz are written as follows [2]:

$$B=\begin{bmatrix} \cos\alpha_n & \sin\alpha_n\sin\psi & -\sin\alpha_n\cos\psi \\ \sin\alpha_n\sin\varphi & \cos\varphi\cos\psi-\cos\alpha_n\sin\varphi\sin\psi & \cos\varphi\sin\psi+\cos\alpha_n\sin\varphi\cos\psi \\ \sin\alpha_n\cos\varphi & -\sin\varphi\cos\psi-\cos\alpha_n\cos\varphi\sin\psi & -\sin\varphi\sin\psi+\cos\alpha_n\cos\varphi\cos\psi \end{bmatrix} \quad (1)$$

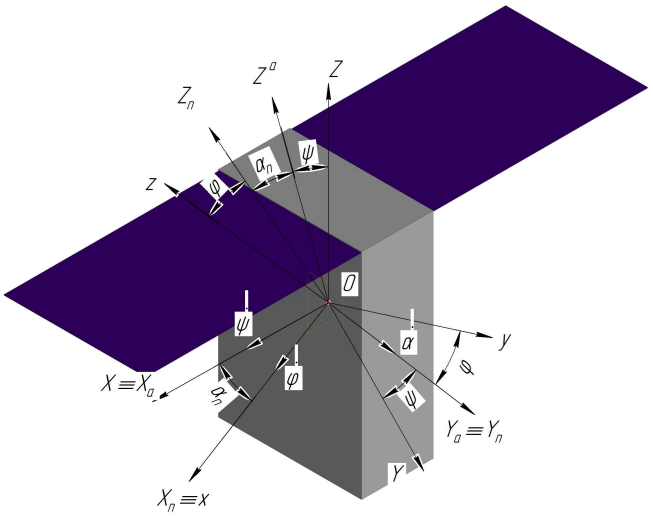


Fig. 2. The coordinate frames used

Sunlight direction in this paper is considered to be parallel to OY axis and directed oppositely to axis's direction. The velocity vector of the incoming atmospheric particles flow is parallel to OY axis and also directed oppositely to axis's direction.

III. TORQUES

A. Gravitational torque

The expressions for the projections of the gravitational moment on the body-fixed frame might be given as [2]:

$$\begin{aligned} M_{gx} &= \frac{3\mu}{r^3}\left(J_z - J_y\right)b_{23}b_{33}, \\ M_{gy} &= \frac{3\mu}{r^3}\left(J_x - J_z\right)b_{33}b_{13}, \\ M_{gz} &= \frac{3\mu}{r^3}\left(J_y - J_x\right)b_{13}b_{23}, \end{aligned} \quad (2)$$

where μ – the gravitational parameter of the Earth; r – spacecraft's radius-vector; J_x, J_y, J_z – the principal central moments of inertia; b_{13}, b_{23}, b_{33} – the elements of transition matrix B (formulae (1)).

B. Aerodynamic drag force torque

Aerodynamic drag force torque can be calculated as:

$$\mathbf{M}_a = \mathbf{r}_d \times \mathbf{Q}_{xv}, \quad (3)$$

where $\mathbf{r}_d$ – radius-vector of the center of pressure's position relatively to the geometric center of the nanosatellite; $\mathbf{Q}_{xv}$ – aerodynamic force.

Aerodynamic force is written as follows:

$$\mathbf{Q}_{xv} = -c_0 \mathbf{A} q, \quad (4)$$

where $c_0 = 2{,}2$;

$\mathbf{A}$ – the vector of the area of CubeSat geometry projection on a plane perpendicular to the velocity vector of the incoming flow;

$q = \frac{\rho(H)\left[v(H)\right]^2}{2}$ – the dynamic velocity pressure;

$v(H) = \sqrt{\frac{\mu}{R_з + H}}$ – the orbital speed;

H – the flight altitude;

$\rho(H)$ – the atmospheric density;

$R_з = 6371\ km$ – the radius of the Earth;

$\mu = 398600 \cdot 10^9\ m^3 s^2$ – is the gravitational parameter of the Earth [2], [4].

The vector of the area of CubeSat projection on a plane perpendicular to the velocity vector of the incoming flow is calculated by our software.

C. Sunlight pressure torque

For sunlight pressure on the surface without reflection calculations we used the following equation:

$$P = \frac{W}{c}, \quad (5)$$

where W – the Solar radiation flux density equals to 1368 W/m^2, c – the speed of the light in vacuum.

For the perfectly reflecting surface the equation (5) has to be multiplied by 2. This is caused by conservation of momentum law. Any intermediate case has this multiplier between 1 and 2. For top possible value estimation purposes we always consider it to be 2.

The force of sunlight pressure is given as:

$$\mathbf{F}_S = P \cdot \mathbf{A}, \quad (6)$$

where $\mathbf{A}$ – the vector of the area of CubeSat geometry projection on to a plane perpendicular to the velocity vector of the sunlight flow and is calculated by our software.

Sunlight pressure torque is written as follows:

$$\mathbf{M}_S = \mathbf{d} \times \mathbf{F}_S, \quad (7)$$

where $\mathbf{d}$ – radius-vector of the geometric center of the area of projection of the nanosatellite on the plane perpendicular to sunlight direction related to the center of mass projection on the same plane [5].

IV. RESULTS

For numeric simulation we developed our own special software in MATLAB. The software uses 3D models of the nanosatellite in STL format and rotates in all possible positions during two or three angle iterating.

Results of numeric simulation of aerodynamical and gravitational moments and also sunlight pressure for any possible nanosatellite's positions are demonstrated in figures 3-6.

Maximum values of all estimated torques are shown in table 1.

TABLE I. MAXIMUM TORQUES' VALUES

Sunlight pressure torque, N·m	*Gravitational torque, N·m*	*Aerodynamical torque, N·m*
$8.5454 \cdot 10^{-8}$	$6.4639 \cdot 10^{-8}$	$3.9734 \cdot 10^{-6}$

All values were calculated for the orbit altitude of 500 kilometers. However, torque with the prevailing influence on satellite motion can change with altitude. Therefore, we calculated the maximum values of all three moments for the altitudes in range from 300 to 1200 kilometers for comparison. Results are shown in figure 7.

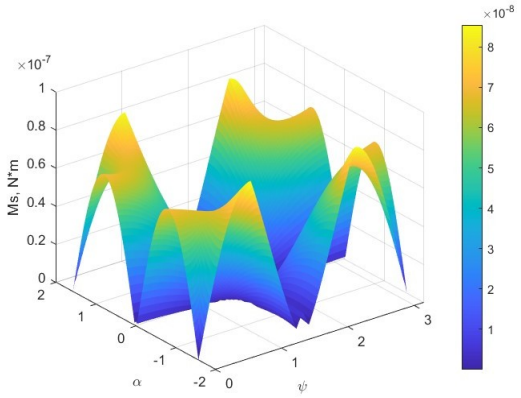


Fig. 3. Sunlight pressure torque values

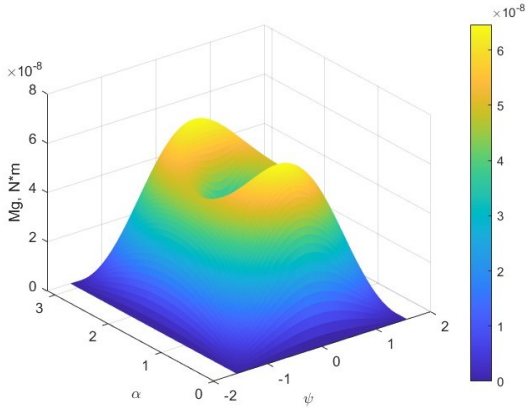


Fig. 4. Gravitational torque values for $\phi=\pi/4$

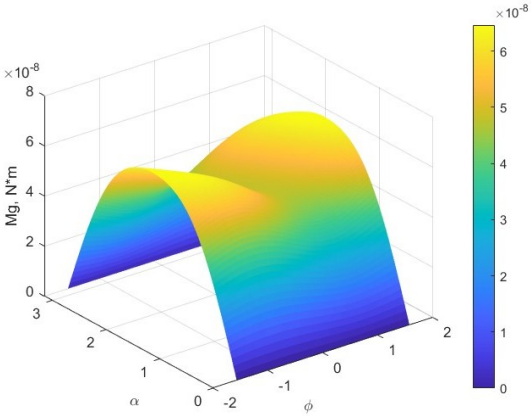


Fig. 5. Gravitational torque values for plane $\psi=0$

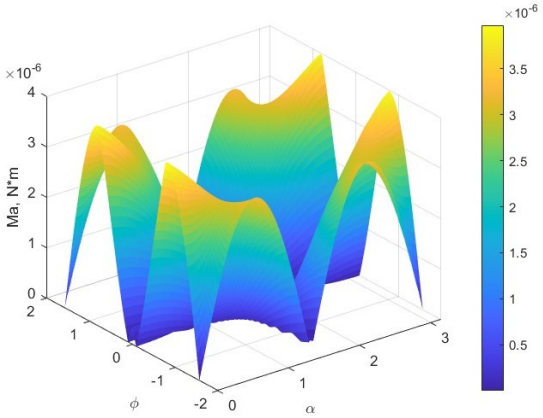


Fig. 6. Aerodynamical torque values

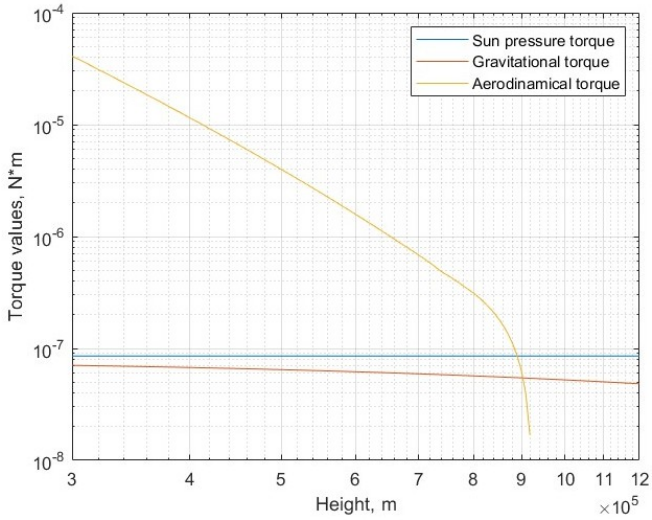


Fig. 7. Maximum torques at different altitudes

V. DISCUSSION

The results of numeric simulations of different torques show us all possible torque values. It can be seen that each torque has its maximum and minimum values. Analyzing all torques combined would allow us to determine the equilibrium positions. However, it should be noted that these charts show dependences with different angles, therefore they cannot be used directly.

Both aerodynamic and sunlight torques have angles which do not affect torque value: ϕ in case of sunlight torque and ψ in case of aerodynamical torque. Gravitational torque on the other hand depends on every angle.

From figures 3-7 we can conclude that the values of sunlight pressure torque for this particular nanosatellite are commensurable to gravitational torque values. At the altitude of 500 kilometers the values of sunlight pressure are nearly two times bigger, than gravitational torque's. With increasing altitude sunlight pressure torque's influence on satellites angular motion begins to predominate over other torques since it does not depend on satellite altitude. In addition, at altitudes greater than 900 kilometers sunlight pressure torque is becoming prevailing in the angular motion of nanosatellite with deployable solar panels.

These results can be used in different ways. The first is the simplest one – to find the maximum value of each type of torque. This would allow us to determine the prevailing torque if present and whether we need to consider sunlight pressure or it could be neglected.

The second is to determine the equilibrium points. Translating all the results to all possible combinations of all three angles and further analysis would allow us to find different equilibrium positions. Note that these results can no longer be represented as three-dimensional charts like shown in this paper.

The third way is to use torques values to simulate satellite dynamics. In that case results can be formed as a three-dimensional array and be called upon during simulation to quickly determine the required values. The main problem would be the resolution of such matrices. It can be improved using interpolation. However, it should be noted that sunlight pressure torque model is designed for estimation purposes only and is not suited for precise simulations.

VI. CONCLUSION

In this paper we designed math models and software to estimate aerodynamical, gravitational and sunlight pressure torques for a given CubeSat model. The model and the software are not limited to a specific mission and can be used for different satellites.

Using the designed software, we calculated values of the torques for a CubeSat-6U satellite with deployable solar panels. We concluded that sunlight pressure torque cannot be neglected and prevails at higher altitude orbits.

The torques affect the angular motion of the nanosatellite and can be used by satellite designers to estimate motion modes and dynamics of a nanosatellite with deployable elements. It can be useful for control system development. At the same time knowledge about the prevailing torque at particular altitude can be used for passive stabilization system development. The results of simulation allow to simulate angular motion dynamics to assess motion modes for different initial conditions and mission parameters.

REFERENCES

[1] E. Kulu, “CubeSats & Nanosatellites - 2024 Statistics , Forecast and Reliability,” Int. Astronaut. Congr. (IAC 2024), vol. c, no. October, pp. 14–18, 2024, [Online]. Available: https://iafastro.directory/iac/archive/browse/IAC-24/B4/6A/83232/

[2] E. V. Barinova, I. V. Belokonov, and I. A. Timbai, “Technology for Designing the Angular Motion of CubeSat Nanosatellites with a Passive Stabilization System,” Gyroscopy Navig., vol. 15, no. 4, pp. 346–355, 2024, doi: 10.1134/S2075108725700063.

[3] I. V. Belokonov et al., “Lessons Learned of Nanosatellite SamSat-ION Mission: Preliminary Results,” in IAF Space Systems Symposium, Paris, France: International Astronautical Federation (IAF), 2024, pp. 683–697. doi: 10.52202/078372-0067.

[4] E. V. Barinova, E. A. Boltov, N. A. Elisov, and I. A. Lomaka, “Improved estimation of aerodynamic characteristics of a complex geometry nanosatellite,” Eng. J. Sci. Innov., no. 10 (118), 2021, doi: 10.18698/2308-6033-2021-10-2116.

[5] R. Georgevic, “Mathematical Model of the Solar Radiation Force and Torques Acting on the Components of a Spacecraft,” AMERICAN ASTRONAUTICAL SOCIETY, 1971, pp. 33–494. [Online]. Available: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Mathematical+Model+of+the+Solar+Radiation+Force+and+Torques+Acting+on+the+Components+of+a+Spacecraft#0

Adaptive Algorithms for Local Gravity Modeling from Airborne Gravimetry Data Using Spherical Point Masses and L_2-Regularization

M.D. Repin

Lomonosov Moscow State University
Moscow, Russia
mikhail.repin@math.msu.ru

V.S. Vyazmin

Lomonosov Moscow State University
Moscow, Russia
vadim.vyazmin@math.msu.ru

***Abstract*—The paper addresses the problem of local gravity modeling from airborne gravimetry data based on parameterizing the disturbing potential by point masses on a sphere. The task involves determining the magnitudes and depths of the point masses by solving an unconstrained nonlinear minimization problem with L_2-regularization. The solution is obtained via an iterative quasi-Newton algorithm. We also consider an alternative formulation of the problem based on a two-stage minimization of the cost function (first determining the magnitudes of point masses and then estimating their depth). The solution is provided by the least squares method and optimization of a convex single-variable function. The paper presents the results of testing and comparing both algorithms using real airborne and ground gravimetry data.**

Keywords—airborne gravimetry, gravity, L_2-optimization, point masses.

I. Introduction

Airborne gravimetry is focused on measuring Earth's gravity field from an aircraft using stabilized-platform or strapdown airborne gravimeters. The final result of an airborne gravimetry survey is to produce a local gravity field model, which is then used in geological and geophysical interpretation or in geoid modeling [1, 2].

Several methods exist for local gravity modeling from gravimetry survey data (usually assuming data gridding and then providing gravity maps for the survey area), including spline interpolation, fast Fourier transform method and least-squares collocation [3-5]. In this work, we use the point mass method to parameterize the gravity field in the survey area [6]. The point masses are harmonic outside a sphere and have a spatial localization property. However, the method is difficult to implement as numerous parameters (magnitudes, locations and depths of point masses) should be adjusted. Adaptive algorithms for gravity modeling from satellite gravimetry data using point masses were proposed in [7, 8]. For gravity modeling from airborne gravity data, only the linear least-squares formulation was previously considered using other types of spherical radial basis functions [9, 10].

In this study, we formulate the local gravity modeling problem as determining the point mass model parameters (magnitudes and depths) given fixed locations of the point masses (located in a regular grid on the sphere). The problem is reduced to optimization of the cost function (the L_2-norm of the residual vector) with the Tikhonov regularization (L_2-regularization) [11]. Two approaches are proposed: an unconstrained nonlinear optimization and a two-stage optimization constrained by a relation imposed on the unknown parameters. In one approach, the solution is provided by a quasi-Newton algorithm [12] and in the other approach by least squares estimation and single-variable convex function minimization.

The paper describes both approaches and their implementation with the focus on numerical issues. The results of gravity modeling from the airborne gravimetry data collected by a strapdown gravimeter (iMAR) in an aerogeophysical survey in Russia are presented. The accuracy assessment obtained through comparison with ground-truth gravity data is also provided in the paper.

II. Gravity Modeling Using Point Masses and L_2-Regularization

Let δg denote the free-air gravity disturbance. We consider the disturbing potential $T(r)$ at a point with the radius-vector r and will use the spherical Earth approximation. The objective is to determine gravity at arbitrary points outside a reference sphere from δg measurements.

The disturbing potential can be locally represented using the point masses $\Phi_k(r)$ [6]:

$$T(r) = \frac{GM}{R}\sum_{k=1}^{K} c_k \Phi_k(r), \quad \Phi_k(r) = \frac{1}{\| r - r_k \|}, \tag{1}$$

where G is the gravitational constant, M is Earth's mass, R is the radius of reference sphere, r_k is the radius-vector of k-th point mass location (r_k=R–d), d is the unknown depth of a point mass, K is the total number of point masses, and c_k is the unknown coefficient (magnitude). From (1) and the relationship between the gravity disturbance and potential, we derive:

$$\delta g(r) = -\frac{GM}{R}\sum_{k=1}^{K} c_k \frac{\partial \Phi_k(r)}{\partial |r|}. \tag{2}$$

Thus, the problem is reduced to determining the point mass magnitudes c_k and depth d from gravity disturbance measurements given in a survey area. Note that c_k and d are dimensional quantities (measured in meters).

A. Unconstrained Optimization

Let N be the number of gravity disturbance measurements in the survey area and K the number of point masses located at the knots of a regular grid. Define the matrix $\breve{\Phi}_d \in \mathbb{R}^{N\times K}$ containing the radial derivatives of the point masses (2).

We pose the local gravity modeling problem as a minimization problem for the following nonlinear cost function with L_2-regularization:

$$\min_{\mathbf{c},\,d}\left[\,\|\,\Delta\mathbf{g}'-\breve{\Phi}_d\mathbf{c}\,\|_2^2+\lambda\,\|\,\mathbf{c}\,\|_2^2\,\right], \tag{3}$$

where $\Delta\mathbf{g}'\in\mathbb{R}^N$ is the vector of gravity measurements in the survey area, $\mathbf{c}\in\mathbb{R}^K$ is the vector of point mass magnitudes, and λ is the regularization parameter. The unknown parameters are the magnitudes $\mathbf{c}$ and depth d ($K+1$ parameters in total). The solution to (3) is obtained numerically via a quasi-Newton algorithm [12], which requires the initial guess for the unknown parameters. An analysis of the cost function gradient shows that its partial derivative w.r.t. d is of order 10^1, while w.r.t. c_k it is of order 10^{10}. However, depth changes are typically of order 10^3 and magnitude changes are of order 10^5. Therefore, when solving the problem numerically, it is necessary to re-scale the parameters taking these issues into account.

The problem formulation (3) is the unconstrained nonlinear minimization, and the quasi-Newton algorithm convergence heavily depends on the initial guess and scaling of the parameters, which should be repeated for gravity datasets from each new survey. This motivates us to consider an alternative formulation of the gravity modeling problem that involves simpler algorithms.

B. *Two-Stage Optimization*

Using the same cost function, we propose the formulation of the local gravity modeling problem as a two-stage optimization:

$$\min_{d}\,\min_{\mathbf{c}}\left[\,\|\,\Delta\mathbf{g}'-\breve{\Phi}_d\mathbf{c}\,\|_2^2+\lambda\,\|\,\mathbf{c}\,\|_2^2\,\right]. \tag{4}$$

The inner minimization represents a linear least-squares problem with d fixed, while the outer minimization optimizes a single-variable nonlinear function of the unknown parameter d. For gravimetry surveys focusing on mapping local gravity anomalies, the range of d can be limited to values from 0 to 50 km [13]. Within this range, the outer minimization in (4) is reduced to optimization of a convex single-variable function (Fig. 1).

The solution to problem (4) implicitly takes into accounts the relationship between the parameters c_k and d imposed by the inner minimization. Therefore, problem (4) is a constrained minimization of the cost function in the parameter space ($\mathbf{c}$, d). However, unlike problem (3), its solution can be obtained by an algorithm that is more computationally efficient and easier to implement.

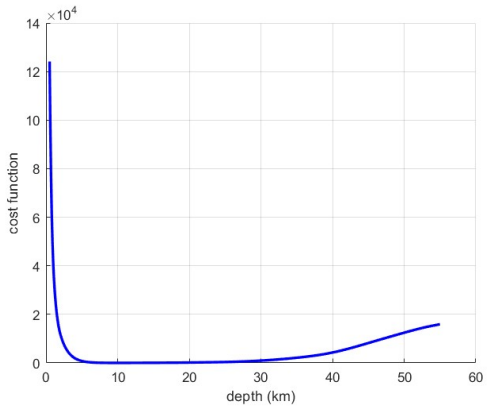


Fig. 1. The dependence of the cost function in (4) on the point mass depth.

C. *Spatial Resolution Determination*

An essential characteristic of a gravity map (or model) is its spatial resolution. We determine it through expansion of the radial derivative of a point mass into a series of Legendre polynomials [6], where the expansion coefficient $k(n)$ takes the following form (Fig. 2):

$$k(n)=(n+1)\left(\frac{R-d}{R}\right)^n. \tag{5}$$

Here n is the Legendre polynomial degree. Hence, the spatial resolution of the point mass model for the selected values of d and R can be determined from the highest degree of the Legendre polynomial at which the coefficient (amplitude) $k(n)$ decays.

III. Numerical Results

The proposed approaches were used for local gravity modeling from high-precision ground gravimetry data (Fig. 3). The constructed gravity model was then upward continued at flight altitude of an airborne gravity survey conducted over the same survey area. The aim of the upward continuation was to evaluate the accuracy of the airborne gravity measurements. The aerogravimetry survey was conducted by Aerogeophysica JSC (Russia) in 2023 using the iCORUS strapdown airborne gravimeter. Raw gravimeter data processing was performed using the methodology and algorithms developed at Lomonosov Moscow State University (the software packages *INS-GNSS* and *IMU-GRAV*) [14].

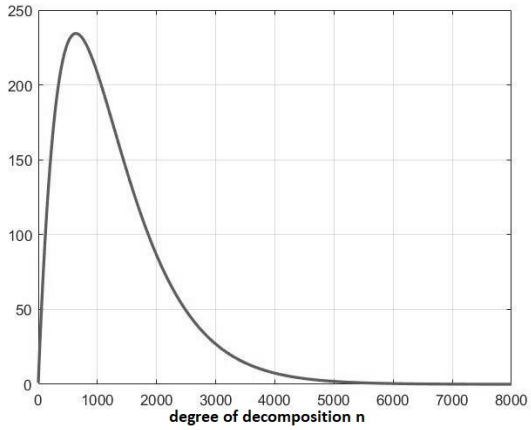


Fig. 2. Values of coefficients at the Legendre polynomials in the spherical harmonic expansion of a point mass.

For gravity modeling from the ground gravimetry data, a lat-lon grid (1.4 by 1.2 km) on a sphere was chosen to locate the point masses. The steps in latitude and longitude are equal to twice the distances between the observation points in these directions (Fig. 3). The regularization parameter values in both approaches (3) and (4) were obtained from the minimum mean square error of the gravity data reconstruction at the observation points. Both approaches yielded an identical value of the regularization parameter: $\lambda=10^3$. The root mean square (RMS) of the residual vector is 0.10 mGal for the first approach and 0.11 mGal for the second approach, with both methods showing zero mean of the residual vector (0.00 mGal). The estimates of the point mass depth provided by both approaches are also identical and equal d=10.5 km. The spatial distribution of the point mass magnitude estimates obtained using approach (3) is shown in Fig. 4.

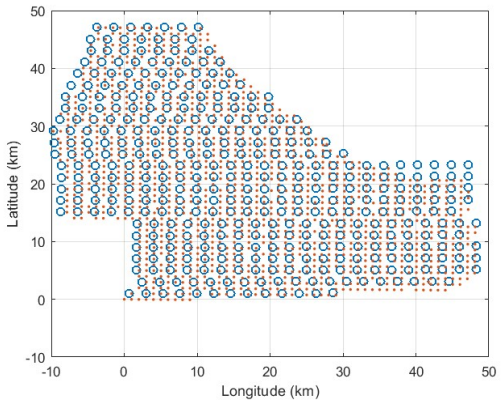


Fig. 3. Observation points in the ground-based survey (red) and the grid of point mass locations (blue).

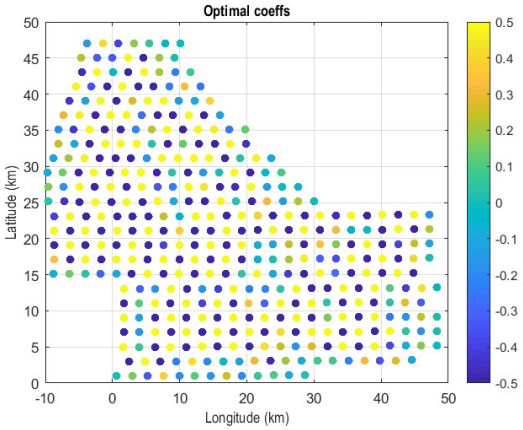


Fig.4. The point-mass magnitudes estimted by approach (3), mm.

To compare both approaches (3) and (4), we examine the difference between the results of the upward continuation of ground gravimetry data at flight altitude obtained using the two approaches (Fig. 5). The RMS is 1.2 μGal and the mean value is 0.3 μGal. This demonstrates that the two gravity maps derived from the original ground-based data using approaches (3) and (4) are almost identical. Therefore, we further consider only one of these two maps, namely, the map provided by (3).

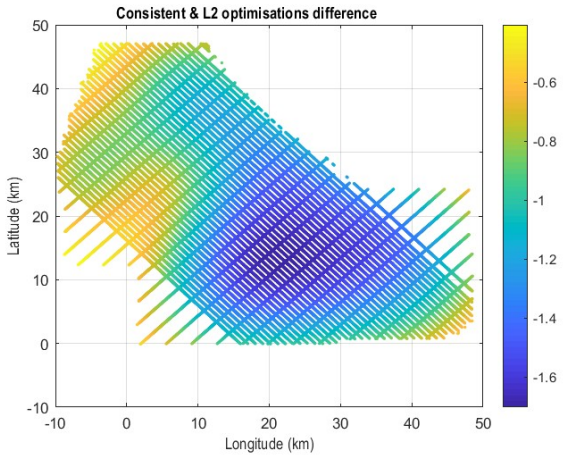


Fig.5. The difference between the gravity disturbances at flight altitude obtained by the upward continuation of ground-based data using approaches (3) and (4), μGal.

Figure 6A displays the airborne gravimetry data, that is, the free-air gravity disturbances at the flight lines obtained through post-processing of raw gravimeter measurements [14]. Figure 6B presents the reference data (the ground gravimetry data upward continued at flight altitude). The difference between the airborne and reference gravity data is shown in Fig. 6C. The RMS is 2.3 mGal and the mean value is 0.3 mGal. These results are consistent with the accuracy of the airborne data obtained from the cross-over analysis, which is 1.9 mGal (RMS error).

The spatial resolution of the reference gravity data (Fig. 6B) was evaluated using (5) with the highest degree of spherical harmonic set equal to n=4500. This resulted in a spatial resolution of approximately 4 km in both directions (along and across the flight lines). In contrast, the spatial resolution of airborne gravity data is determined by two factors: the cutoff frequency of the gravimetric filter and the flight line spacing (the distance between the survey lines) [1]. For the airborne gravity data considered in the paper, we obtain the half-wavelength spatial resolution of 2.5 km along the lines and the line spacing of 1 km. Therefore, the airborne gravimetry data demonstrate significantly higher resolution in both directions comparing to the reference gravity data.

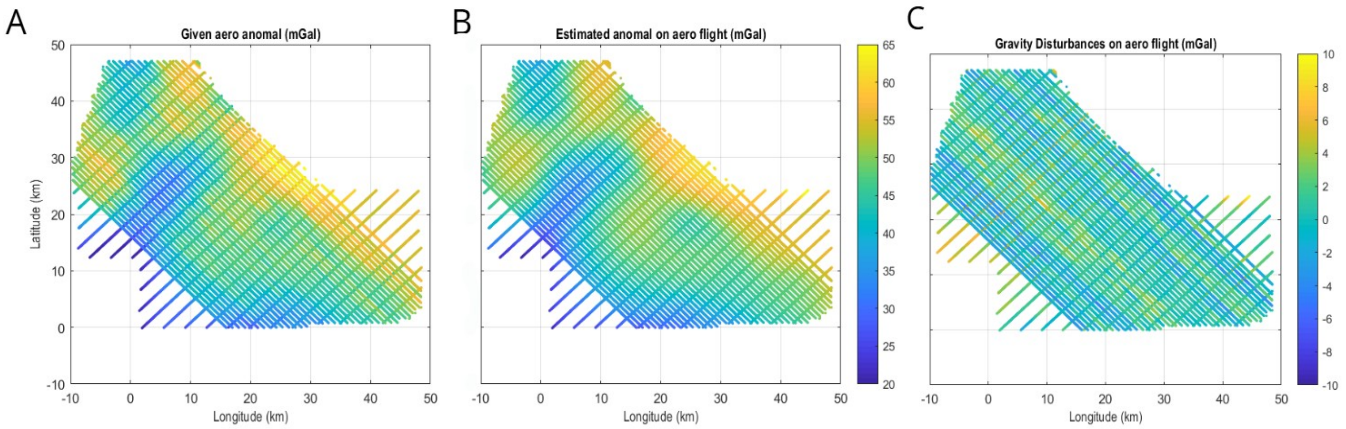


Fig.6. (A): Airborne gravity data, mGal. (B): The upward-continued ground gravimetry data (the reference data), mGal. (C): The difference between the airborne and reference gravity data, mGal.

IV. Conclusions

Two adaptive algorithms for local gravity modeling from gravimetry data are proposed based on parameterization of the anomalous gravity field by the point masses on a sphere. The first algorithm estimates the unknown parameters of the point masses (the magnitudes and depth) through optimization of the L_2-norm of the residual vector with Tikhonov regularization. The optimization problem is nonlinear and solved iteratively using a quasi-Newton algorithm.

The second algorithm solves a constrained optimization problem that implicitly takes into account a relationship between the point mass magnitudes and depth. The advantage of the algorithm is its significantly greater computational efficiency due to the minimization of a convex single-variable function.

Both approaches were tested using airborne and ground gravimetry data and demonstrated comparable accuracy of the results. A key finding from the comparison is the equivalent estimates of the point mass depth obtained by both approaches.

References

[1] Methods and technologies for measuring the earth's gravity field parameters. Peshekhonov V.G., Stepanov O.A. (eds.). Springer, Cham, Switzerland, 2022. https://doi.org/10.1007/978-3-031-11158-7

[2] Vyazmin V.S., Golovin A.A., Kozlov A.V., Repin M.D., Fedorov S.A. Test results from a protoype of a strapdown airborne gravimeter in aerogeophysical surveys. Proceedings of XXXIV conference in memory of the outstanding designer of gyroscopic instruments N.N. Ostryakov. Central State Research Institute "Electropribor", St. Petersburg, 2024, pp. 8-11. (In Russian)

[3] Boyarsky E.A., Koneshov V.N., Kurmaeva A.Sh., Khachiyan T.P. Analysis of marine gravimetric surveys using two-dimensional cubic splines. Gravi-inertial Instruments and Measurements. Institute of physics of the earth of the Academy of Sciences of the USSR, Moscow, 1985, pp. 72-82. (In Russian)

[4] Forsberg R. A new covariance model for inertial gravimetry and gradiometry. J Geophys Res Solid Earth. 1987. V. 92(B2). P. 1305–1310.

[5] Tscherning C.C. Geoid determination by 3D least-squares collocation. In: Sansò F, Sideris M (eds.) Geoid determination. Lecture notes in earth system sciences. V. 110. Springer, Berlin Heidelberg. 2013.

[6] Goossens S., Smith D.E. Gravity degree–depth relationship using point mass spherical harmonics. Geophysical Journal International. 2023. V. 233. P. 1878–1889.

[7] Lin M., Denker H., Müller J. A comparison of fixed- and free-positioned point mass methods for regional gravity field modeling. J. Geodyn. 2019. V. 125. P. 32–47.

[8] Klees R., Tenzer R., Prutkin I., Wittwer T. A data-driven approach to local gravity field modelling using spherical radial basis functions. 2008. J. Geod. V. 82. P. 457–471.

[9] Li X.P. Using radial basis functions in airborne gravimetry for local geoid improvement. J. Geod. 2018. V. 92. P. 471–485.

[10] Vyazmin V.S. Local gravity anomaly determination from airborne gravimetry data using spherical wavelet decomposition. Ph.D. dissertation in Physical and Mathematical Sciences. Moscow, Lomonosov Moscow State University, 2014. 107 p. (In Russian)

[11] Naeimi M., Flury J., Brieden P. On the regularization of regional gravity field solutions in spherical radial base functions. Geophys. J. Int. 2015, 202, 1041–1053.

[12] Nocedal J., Wright S. Numerical optimization. Second edition. Springer New York, NY. 2006. 664 p.

[13] Torge W. Gravimetry.. De Gruyter, Berlin, 1989. 465 p.

[14] Golovan A.A., Vyazmin V.S. Methodology of airborne gravimetry surveying and strapdown gravimeter data processing. Gyroscopy and Navigation. 2023. Vol. 14, no. 1. P. 36–47. https://doi.org/10.1134/S2075108723010029.

Research on the Current Status and Future Development Trends of Intelligent Navigation Technology

Mao Luo
Beijing Institute of Aerospace Control Devices
Beijing, China
che_maomao@126. com

Abstract **With the rapid development of artificial intelligence, the Internet of Things, and big data technologies, intelligent navigation technology has become a key component in modern transportation, aerospace, robotics, and intelligent transportation systems. This paper comprehensively analyzes the current status of intelligent navigation technology, explores its performance in different application scenarios, and predicts future development trends. The study combines empirical analysis with methodological discussions, aiming to provide valuable references for researchers and engineers in related fields.**

Firstly, this paper categorizes and evaluates existing intelligent navigation technologies, focusing on navigation systems based on Global Navigation Satellite Systems (GNSS) , Inertial Navigation Systems (INS) , Machine Vision (MV) , and multi-sensor fusion. Through empirical analysis of performance indicators such as accuracy, reliability, real-time capabilities, and environmental adaptability, the paper reveals the strengths and limitations of these navigation technologies in different application scenarios. For example, GNSS performs well in open environments but suffers from signal blockage in urban canyons and indoor settings; INS, while autonomous, accumulates errors over time. Machine vision and multi-sensor fusion technologies, however, demonstrate higher adaptability and robustness in complex environments.

Secondly, the paper explores key methodological advancements in intelligent navigation technology. In recent years, the application of deep learning and artificial intelligence in the navigation field has become widespread, such as using Convolutional Neural Networks (CNN) for error correction and environmental perception, and employing reinforcement learning to optimize path planning. The paper proposes a multi-sensor fusion framework based on deep learning, which can correct INS errors in real-time and combine geographical feature information provided by machine vision to significantly enhance the accuracy and robustness of navigation systems. Experimental verification shows that this method improves positioning accuracy by approximately 30%in complex environments and exhibits good adaptability under dynamic occlusion and varying lighting conditions.

Lastly, the paper looks ahead to the future development trends of intelligent navigation technology. With the rapid development of 5G communication, the Internet of Things, and autonomous driving technologies, intelligent navigation systems will place greater emphasis on multi-sensor fusion, real-time data processing, and autonomous decision-making capabilities. Moreover, artificial intelligence technologies will further optimize navigation algorithms and enhance the intelligence level of systems. Future research directions may include: (1) developing higher-precision sensors and fusion algorithms; (2) exploring new types of navigation systems based on quantum technology; and (3) studying navigation solutions for extreme environments (such as deep sea, space, and complex urban environments) . In summary, through a comprehensive analysis of the current status and future trends of intelligent navigation technology, this paper presents the latest progress and potential development directions in the field. The study results indicate that intelligent navigation technology is evolving towards higher accuracy, stronger robustness, and broader application scenarios, providing theoretical and technical support for future breakthroughs in intelligent transportation and robotics. .

Keywords— Intelligent navigation; Multi-sensor fusion; Deep learning; Development trends; Empirical research

I. INTRODUCTION

1. 1 Research Background and Significance

With the commercialization and popularization of the Global Positioning System (GNSS) , navigation technology has undergone a revolutionary change in the 21st century. From everyday navigation on smartphones to path planning for autonomous vehicles, precise landing of drones, and deep-space navigation of space probes, navigation technology has become an indispensable part of modern society. However, traditional navigation technologies are facing increasing challenges, such as signal blockage in urban canyons, continuity of indoor-outdoor seamless positioning, real-time requirements in complex dynamic environments, and robustness under extreme conditions.

In recent years, with the rapid development of artificial intelligence, the Internet of Things, big data, and 5G communication technologies, intelligent navigation technology has emerged and rapidly risen. Intelligent navigation technology not only integrates the advantages of traditional navigation systems but also introduces new sensing technologies, data processing methods, and decision-making algorithms, significantly improving the accuracy, reliability, and adaptability of navigation systems. The advent of intelligent navigation technology provides new ideas and methods for addressing the limitations of traditional navigation technologies and lays the foundation for the development of emerging industries such as autonomous driving, drones, and service robots.

The purpose of this study is to systematically review the current status of intelligent navigation technology, analyze the characteristics and limitations of various navigation technologies, explore their performance in different application scenarios, and predict future development trends. Through this research, we aim to provide a comprehensive reference for researchers, engineers, and policymakers in the field of navigation technology, promoting further development and application of intelligent navigation technology.

1. 2 Research Scope and Methods

This study covers several key aspects of intelligent navigation technology, including sensor technology, data

processing algorithms, system integration, and application scenario analysis. The research scope involves the following main areas:

• Classification and Evaluation of Navigation Technologies:A comprehensive analysis of mainstream navigation technologies such as GNSS, INS, MV, and multi-sensor fusion is conducted to understand their principles, characteristics, and application scenarios. A comprehensive evaluation index system is established to quantitatively and qualitatively analyze the performance of each technology.

• Key Technology Progress Research:In-depth exploration of key technologies in intelligent navigation, including deep learning, reinforcement learning, multi-sensor fusion, and real-time data processing. The principles, advantages, and application effects of these technologies in navigation systems are analyzed.

Empirical Analysis and Experimental Verification:An experimental platform is constructed, and comparative experiments are designed to verify the effectiveness and superiority of the proposed multi-sensor fusion framework. The performance improvement in complex environments is quantitatively analyzed.

• Future Trends and Development Directions:Future development trends of intelligent navigation technology are predicted based on technological trends and industry needs, and potential research hotspots and application fields are proposed.

In terms of research methods, this paper adopts a combination of literature review, theoretical analysis, algorithm design, and experimental verification. By systematically reviewing domestic and international research findings, a theoretical framework for intelligent navigation technology is established. New methods and technical solutions are proposed based on this framework, and their effectiveness is verified through simulation experiments and practical tests.

II. Overview of Intelligent Navigation Technology

2. 1 Basic Concepts of Navigation Technology

Navigation technology refers to the use of certain means and methods to determine the position, velocity, attitude, and other information of a carrier (such as vehicles, aircraft, ships, pedestrians, or robots) and to guide the carrier to safely, accurately, and efficiently reach its destination along a predetermined route. Depending on different classification criteria, navigation technology can be divided into various types:

• Classification by Positioning Method:This includes autonomous navigation and non-autonomous navigation. Autonomous navigation relies solely on the equipment carried by the carrier itself, such as INS, without depending on external information. Non-autonomous navigation requires external devices or signal sources, such as GNSS.

• Classification by Working Environment: Navigation can be divided into outdoor and indoor navigation. Outdoor navigation mainly depends on satellite signals, while indoor navigation requires other technologies such as Wi-Fi, Bluetooth, Ultra-Wideband (UWB) , and machine vision.

• Classification by Application Field: This includes land navigation, marine navigation, aviation navigation, and space navigation, each with different requirements for the accuracy, reliability, and real-time capabilities of navigation technology.

• Classification by Technical Features: Navigation technology can be divided into traditional navigation technology and intelligent navigation technology. Traditional navigation technology primarily relies on mathematical models and physical laws, while intelligent navigation technology makes greater use of artificial intelligence, machine learning, and big data analysis.

2. 2 Definition and Characteristics of Intelligent Navigation Technology

Intelligent navigation technology refers to the use of artificial intelligence, machine learning, big data analysis, multi-sensor fusion, and other technologies to achieve higher accuracy, stronger robustness, and broader application scenarios in navigation. Compared with traditional navigation technology, intelligent navigation technology has the following significant characteristics:

• Adaptability: Intelligent navigation systems can automatically adjust navigation strategies according to environmental changes and carrier status, enhancing the system's adaptability to different environments.

• Robustness: Through multi-sensor fusion and data redundancy design, intelligent navigation systems can maintain high positioning accuracy and reliability in complex environments.

Intelligence:Utilizing artificial intelligence and machine learning technologies, intelligent navigation systems can make autonomous decisions and optimizations, improving navigation efficiency and service quality.

• Interconnectivity:Relying on the Internet of Things and 5G communication technologies, intelligent navigation systems can achieve information interaction between vehicles (V2V) , vehicles and infrastructure (V2I) , and vehicles and pedestrians (V2P) , enhancing the overall efficiency of the transportation system.

• Real-time Capabilities: With high-performance computing platforms and real-time data processing algorithms, intelligent navigation systems can quickly respond to environmental changes and provide real-time navigation services.

2. 3 Development History of Intelligent Navigation Technology

The development of intelligent navigation technology can be traced back to the 1980s. With the advancement of computer and sensor technologies, navigation systems began to evolve towards intelligence. Based on technical characteristics and application scenarios, the development of intelligent navigation technology can be roughly divided into the following stages:

• Budding Stage (1980s-1990s) :During this period, with the progress of microprocessors and digital signal processing technologies, Inertial Navigation Systems (INS) began to be miniaturized. Early integrated navigation systems, such as GNSS/INS integrated navigation systems, also emerged. The main feature of this stage

• Development Stage (2000s-2010s) :With the popularization of the Global Positioning System and the maturation of MEMS inertial devices, intelligent navigation technology entered a period of rapid development. During this period, multi-sensor fusion technology gained attention, and Kalman filtering and its improved algorithms were widely used for data fusion in navigation systems. Meanwhile, machine vision technology began to be applied in the navigation field, providing new perception means for intelligent navigation systems.

• Intelligent Stage (2010s to present) :With the breakthroughs in deep learning , reinforcement learning, and big data technologies, intelligent navigation technology has entered a new stage of intelligence. Feature extraction and environmental perception methods based on deep learning have significantly improved the robustness of navigation systems in complex environments. Reinforcement learning has provided new solutions for path planning and decision optimization . In addition, the development of 5G communication and edge computing technologies has enabled intelligent navigation systems to achieve lower-latency data transmission and processing, supporting more demanding real-time application scenarios.

III. Key Technology Analysis of Intelligent Navigation

3. 1 Multi-Sensor Fusion Technology

3. 1. 1 Types of Sensors and Their Characteristics

Multi-sensor fusion technology is one of the core technologies of intelligent navigation systems. It integrates information from different types of sensors to improve the accuracy, reliability, and robustness of navigation systems. Commonly used navigation sensors include:

• Inertial Measurement Unit (IMU) :An IMU, composed of accelerometers and gyroscopes , can measure the angular velocity and acceleration of a carrier, and then calculate the carrier's position, velocity, and attitude. The strengths of IMUs lie in their strong autonomy and high short-term accuracy. However, their main limitation is that errors accumulate over time, leading to increased positioning deviations with prolonged use.

• Global Navigation Satellite System (GNSS) :This includes the GPS of the United States, GLONASS of Russia, Galileo of Europe, and Beidou of China. GNSS can provide high-precision location information with global coverage. However, its main issue is susceptibility to blockage and interference, resulting in a significant decline in performance in urban canyons and indoor environments.

Machine Vision System:By capturing image information of the surrounding environment with cameras and using computer vision techniques for feature extraction and environmental perception, machine vision systems perform well in structured environments.However, they are prone to failure in environments with drastic changes in lighting, lack of texture, or overly similar textures.

• Odometer:Including wheel odometers, LiDAR odometers, and visual odometers, an odometer estimates position by measuring the distance and direction of a carrier's movement.Odometers have high accuracy over short distances but experience cumulative errors as the distance increases.

• Magnetometer:By measuring the Earth's magnetic field to determine direction, magnetometers are often used to assist in heading estimation.The main issue with magnetometers is their susceptibility to interference from surrounding magnetic fields, which limits their accuracy.

• Barometer:By measuring changes in atmospheric pressure to estimate altitude changes, barometers are commonly used for altitude measurement.The accuracy of barometers is significantly affected by weather conditions and the surrounding environment.

3.1.2 Fusion Algorithms and Architectures

The key to multi-sensor fusion lies in effectively integrating information from different sensors to overcome the limitations of individual sensors.The main fusion algorithms include:

• Kalman Filter:The Kalman filter is an optimal linear estimation method widely used for state estimation in navigation systems.Its basic idea is to use the system's dynamic equation and observation equation to recursively calculate the optimal estimate of the state variables.Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) are important improved forms, suitable for nonlinear systems.

• Particle Filter:The particle filter is a nonlinear and non-Gaussian filtering technique based on the Monte Carlo method.It approximates the system's posterior probability distribution using a set of random samples (particles) .Although particle filters perform well in nonlinear and non-Gaussian systems, they have a large computational load.

• Loosely Coupled and Tightly Coupled Fusion Architectures:Loosely coupled fusion involves independent positioning by each sensor, followed by result fusion.Tightly coupled fusion involves fusing the raw sensor data.Loosely coupled fusion is simple to implement but has lower time synchronization requirements;tightly coupled fusion offers higher accuracy but is more complex to implement.

• Deep Learning Methods:In recent years, deep learning has been widely used in multi-sensor fusion.Deep learning models, such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) , can automatically extract high-level features from multi-sensor data, achieving more effective information fusion.

3.1.3 Analysis of Typical Application Scenarios

Multi-sensor fusion technology has demonstrated unique advantages in multiple fields:

• Urban Environment Navigation:In urban environments, GNSS signals are often blocked by high-rise buildings, resulting in decreased positioning accuracy.By fusing data from IMUs, vision, and magnetometers, navigation accuracy in urban environments can be significantly improved.

• Indoor Navigation:Indoor environments lack GNSS signals.Multi-sensor fusion technology combines information from WiFi, Bluetooth, magnetometers, and vision to achieve seamless indoor-outdoor positioning.

• Complex Terrain Navigation:In complex terrains such as mountains and tunnels, a single sensor struggles to provide continuous and reliable positioning information.Multi-sensor

fusion technology, through redundancy design and data complementarity, can significantly enhance the robustness of navigation systems.

Unmanned System Navigation:Unmanned systems such as drones, autonomous vehicles, and robots need to make real-time decisions in dynamically changing environments.Multi-sensor fusion technology provides comprehensive environmental perception capabilities, supporting autonomous navigation and obstacle avoidance.

3.2 Application of Deep Learning in Navigation

3.2.1 Environmental Perception and Feature Extraction

Accurately perceiving the surrounding environment and extracting effective features are crucial for the success of navigation systems in complex environments.Deep learning technology has played an important role in this area:

• Image Feature Extraction:Convolutional Neural Networks (CNN) have a natural advantage in image processing, capable of automatically extracting high-level semantic features from images.In the navigation field, CNNs are widely used for tasks such as road recognition, traffic sign detection, and obstacle identification.For example, classic network structures such as ResNet, VGG, and YOLO have been successfully applied in automotive vision systems.

• Point Cloud Processing:Point cloud data generated by LiDAR contains rich three-dimensional spatial information.Deep learning models such as PointNet, PointNet++, and VoxelNet can effectively process point cloud data to achieve functions like terrain modeling, obstacle detection, and SLAM (Simultaneous Localization and Mapping) .

• Multi-Modal Data Fusion:Deep learning technology can effectively fuse heterogeneous data from different sensors.For example, the combination of CNN and LSTM can handle both image and time-series data simultaneously.Multi-modal fusion networks can integrate visual, radar, and LiDAR data to improve the accuracy and robustness of environmental perception.

3.2.2 Path Planning and Decision Optimization

Path planning and decision-making are core functions of navigation systems, directly affecting the effectiveness and efficiency of navigation.Deep learning technology has brought new solutions to this field:

• End-to-End Path Planning:Traditional path planning methods usually rely on manually designed rules and environmental models.Deep learning can achieve end-to-end learning from sensor input to path output.For example, Deep Reinforcement Learning (DRL) allows agents to interact with the environment in a simulated setting to learn optimal policies, which has shown potential in path planning in complex environments.

• Behavior Prediction and Decision-Making:In multi-agent environments, predicting the behavior of other traffic participants (such as vehicles and pedestrians) and making corresponding decisions is a key challenge for navigation.Prediction models based on deep learning can learn the movement patterns of traffic participants, providing support for safe and efficient decision-making.

• Real-Time Decision Optimization:In dynamic environments, navigation systems need to adjust paths in real-time to deal with sudden situations.Deep reinforcement learning, through online interaction with the environment, can optimize policies in real-time, enhancing the adaptability of navigation systems.

3.3 Real-Time Data Processing and Computing Platforms

3.3.1 Real-Time Data Processing Algorithms

Intelligent navigation systems need to process large amounts of sensor data in real-time and make quick decisions.Therefore, efficient real-time data processing algorithms are crucial:

• Data Preprocessing:Raw sensor data often contains noise and redundant information, which needs to be preprocessed to improve data quality.Common preprocessing methods include filtering (such as Kalman filtering and mean filtering) , noise reduction (such as wavelet denoising) , feature extraction, and normalization.

Feature Selection and Dimensionality Reduction:In high-dimensional data, there are often many redundant and irrelevant features.This not only increases computational complexity but may also affect model performance.Dimensionality reduction methods such as Principal Component Analysis (PCA) , Linear Discriminant Analysis (LDA) , and autoencoders can effectively extract the main features of the data and reduce computational complexity.

• Real-Time Data Fusion:Time synchronization and spatial registration of multi-sensor data are prerequisites for data fusion.Methods such as sliding window techniques, time interpolation, and coordinate transformation can achieve effective fusion of multi-source data.

3.3.2 Edge Computing and Cloud Computing

With the increasing requirements for real-time performance and computing power in intelligent navigation systems, the design of computing architectures has become particularly important:

• Edge Computing:Edge computing moves computing tasks and data storage from the cloud to the network edge closer to the data source, reducing data transmission latency and improving response speed.In intelligent navigation systems, edge computing devices (such as in-vehicle computers and drone onboard computers) can process sensor data in real-time, supporting rapid decision-making and control.

• Cloud Computing:For tasks that require large-scale data processing and model training, cloud computing provides powerful computing and storage resources.Through cloud platforms, complex algorithms can be remotely deployed and updated, and information sharing and collaboration between multiple navigation systems can be achieved.

puting:Fog computing, as a supplement to edge computing and cloud computing, brings computing power closer to terminal devices at the network edge.It achieves better resource sharing and load balancing while ensuring low latency.

3.3.3 Hardware Acceleration and Specialized Chips

To meet the requirements of real-time performance and energy efficiency in intelligent navigation systems, specialized hardware accelerators and chips have developed rapidly:

• GPU Acceleration:Graphics Processing Units (GPUs) are widely used for training and inference of deep learning models due to their powerful parallel computing capabilities.In intelligent navigation systems, GPUs can significantly speed up the processing of image processing and deep learning models.

• FPGA Acceleration:Field-Programmable Gate Arrays (FPGAs) have reconfigurable hardware architectures that can be customized for specific application needs, providing efficient hardware acceleration for navigation algorithms.

• Specialized AI Chips:In recent years, specialized chips for artificial intelligence applications (such as TPUs and NPUs) have emerged.These chips are optimized for deep learning computations and offer higher energy efficiency and computational density, making them suitable for in-vehicle computing platforms and embedded navigation devices.

IV. Empirical Study of Intelligent Navigation Technology

4.1 Experimental Design and Methods

4.1.1 Experimental Platform Construction

This study has built a complete experimental platform for intelligent navigation, which mainly includes the following components:

• Sensor Subsystem:Including various sensors such as GNSS receivers, IMUs, RGB-D cameras, LiDAR, and millimeter-wave radar, to obtain rich environmental information.

• Computing Subsystem:Using the NVIDIA Jetson AGX Xavier edge computing platform, which integrates high-performance GPUs and specialized AI accelerators, to process multi-source sensor data in real-time.

Communication Subsystem:Supporting 5G and Wi-Fi 6 communication protocols to achieve high-speed data transmission with the cloud and other devices.

• Localization and Navigation Algorithm Module:Implementing various sensor fusion algorithms, deep learning models, and path planning methods.

4.1.2 Experimental Scenario Setup

To comprehensively evaluate the performance of the intelligent navigation system, this study has designed several typical experimental scenarios:

• Urban Road Environment:Including open areas, under elevated bridges, tunnel entrances, and densely populated urban areas, to test the system's performance in complex urban environments.

• Indoor Environment:Including shopping malls, office buildings, and underground parking lots, to assess the system's navigation capabilities in environments without GNSS signals.

• Special Environment:Simulating environments such as tunnels, forests, and mountainous areas where GPS signals are weak or completely blocked, to test the system's robustness and adaptability.

• Dynamic Interaction Environment:Setting up moving obstacles and sudden obstacles to test the system's real-time response capabilities and path planning effects.

4.1.3 Performance Evaluation Indicators

To quantitatively analyze the performance of the intelligent navigation system, this study has adopted the following evaluation indicators:

• Localization Accuracy:Measured by comparing the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) between the system's estimated position and the true position (provided by GNSS as a reference) .

• Attitude Angle Error:Comparing the differences between the system's estimated roll, pitch, and yaw angles and the true values, usually evaluated using RMSE.

• Path Tracking Accuracy:Measuring the deviation between the actual driving trajectory and the planned path, commonly assessed using indicators such as average lateral deviation and maximum lateral deviation.

• Real-Time Performance:Recording the end-to-end delay from sensor data acquisition to navigation decision output to evaluate the system's real-time performance.

• Robustness:Analyzing the stability and consistency of system performance through repeated experiments under different environmental conditions.

4.2 Experimental Results and Analysis

4.2.1 Verification of Multi-Sensor Fusion Effect

The experiment compared the performance differences between single-sensor navigation and multi-sensor fusion navigation:

• GNSS and IMU Combined Navigation:In open environments, GNSS/INS combined navigation shows high accuracy, but the positioning error increases rapidly under signal blockage.

• Vision and IMU Fusion Navigation:In environments with rich textures and good lighting conditions, vision/INS combined navigation achieves accuracy comparable to GNSS.However, performance significantly deteriorates in environments with missing textures or drastic changes in lighting.

• Multi-Sensor Fusion Navigation:By fusing data from GNSS, IMU, vision, and LiDAR, the navigation system's performance is significantly improved in various environments.Experimental results show that in urban canyon environments, the localization accuracy of multi-sensor fusion navigation is about 50%higher than that of single-sensor navigation.In indoor environments, compared to pure vision navigation, the position error of the fusion solution is reduced by about 40%.

4.2.2 Performance Evaluation of Deep Learning Algorithms

The deep learning-based navigation algorithms proposed in this study show advantages in several aspects:

• Feature Extraction and Object Recognition:Compared to traditional computer vision methods, deep learning models have increased object recognition accuracy by about 25%in complex scenes, especially under occlusion and varying lighting conditions, showing significantly enhanced robustness.

• SLAM Performance:Compared to traditional methods, SLAM methods based on deep learning are more stable in dynamic environments and textureless areas, with a 30%improvement in the completeness and accuracy of map construction.

• Path Planning Effect:In complex urban environments, path planning algorithms based on deep reinforcement learning can effectively avoid sudden obstacles, with a path length optimization rate of 15%-20%, while ensuring safety.

4.2.3 Analysis of System Real-Time Performance and Robustness

The experiment analyzed the real-time performance and robustness of the intelligent navigation system:

• End-to-End Delay:On the edge computing platform, the average delay from sensor data acquisition to navigation decision output is 80ms, meeting the requirements of most real-time navigation applications.

• Response to Environmental Sudden Changes:When sudden changes occur in the environment (such as the sudden appearance of obstacles or signal blockage) , the system can complete environmental perception and strategy adjustment within 200ms, showing good real-time response capabilities.

• Long-Term Operational Stability:In an 8-hour continuous outdoor navigation experiment, the cumulative positioning error of the multi-sensor fusion navigation system remained within 3%, while the error of the single-sensor navigation system increased significantly over time.

• Extreme Condition Testing:Under extreme conditions such as heavy rainfall, low lighting, and electromagnetic interference, the multi-sensor fusion navigation system can still maintain basic functionality, while the single-sensor navigation system fails severely.

4.3 Discussion of Results

Through the analysis of experimental results, this study draws the following conclusions:

• Multi-Sensor Fusion is an Effective Way to Improve Navigation Accuracy:By integrating information from different sensors, the limitations of individual sensors can be overcome, significantly improving the accuracy and robustness of navigation systems.This is especially evident in urban canyons, indoor environments, and complex dynamic environments, where multi-sensor fusion shows clear advantages.

• Deep Learning Brings Innovation to Navigation Technology:Deep learning technology demonstrates strong capabilities in feature extraction, object recognition, and decision optimization, effectively enhancing the performance of navigation systems in complex environments.

• Edge Computing is Key to Real-Time Navigation:The edge computing platform provides the necessary computing power, enabling complex navigation algorithms to run in real-time on resource-constrained embedded devices.

• System-Level Optimization is More Important than Single Technology Improvement:The performance improvement of navigation systems relies not only on the progress of individual technologies but also on overall system optimization

V. FUTURE DEVELOPMENT TRENDS OF INTELLIGENT NAVIGATION TECHNOLOGY

5.1 Forecast of Technological Development Trends

5.1.1 Development Trends in Sensor Technology

Future sensor technology will evolve towards higher precision, smaller size, and lower cost:

• GNSS Improvements:The next generation of GNSS systems will offer higher accuracy and better anti-jamming capabilities, such as the Beidou-3 and Galileo systems.Meanwhile, advancements in signal processing technology will enhance positioning capabilities in weak signal environments.

• MEMS Inertial Devices:MEMS inertial sensors will develop towards higher precision, lower power consumption, and smaller form factors to meet the diverse needs of consumer and industrial applications.

• New-Type Visual Sensors:New visual sensors, such as event cameras and light field cameras, will gradually be applied in the navigation field to address the limitations of traditional cameras in high-speed motion and extreme lighting conditions.

• Bio-Sensor Fusion:Bio-sensors, which can perceive human physiological characteristics and movement intentions, will be integrated with traditional environmental perception sensors to achieve more natural and intelligent human-machine interaction navigation.

5.1.2 Innovation in Algorithms and Models

The rapid development of artificial intelligence will continue to drive innovation in navigation algorithms:

• Lightweight Deep Learning Models:With the advancement of model compression and knowledge distillation techniques, deep learning models will become more lightweight, making them suitable for deployment on resource-constrained embedded devices.

• Multi-Modal Learning:Future navigation algorithms will focus more on joint learning of multi-modal data to achieve efficient fusion of visual, radar, LiDAR, and inertial data.

• Meta-Learning and Transfer Learning:Meta-learning and transfer learning technologies will enable navigation systems to quickly adapt to new environments and tasks, reducing the dependence on large amounts of labeled data.

• Federated Learning and Privacy Protection:Under the premise of ensuring user privacy, federated learning technology will promote knowledge sharing and collaborative learning across devices and platforms.

5.1.3 Evolution of System Architecture

The architecture of intelligent navigation systems will evolve towards being more open, modular, and interconnected:

• Open Navigation Platforms:Future navigation systems will provide standardized interfaces and protocols to support interconnectivity and functional expansion of devices from different manufacturers.

• Cloud-Edge-End Collaborative Architecture:The collaborative operation of cloud computing, edge computing, and end devices will become mainstream, achieving hierarchical optimization of data processing and dynamic allocation of resources.

• Digital Twin Technology:Digital twin technology will create virtual replicas of the physical world, providing richer training data and simulation environments for navigation systems.

• Quantum Navigation Exploration:Emerging technologies such as quantum inertial navigation and quantum-enhanced positioning will gradually move from laboratories to practical applications, breaking through the precision limits of traditional navigation technologies.

5.2 Potential Challenges and Solutions

5.2.1 Data Security and Privacy Protection

As navigation systems collect and process increasing amounts of personal and environmental data, data security and privacy protection have become significant challenges.Solutions include:

• Data Encryption and Access Control:Employing advanced encryption technologies and strict access control mechanisms to safeguard data during transmission and storage.

• Differential Privacy Technology:Introducing noise during data sharing and analysis to protect individual privacy from being compromised.

Blockchain Technology:Leveraging the decentralized nature of blockchain to enable secure data sharing and trustworthy traceability.

5.2.2 Algorithm Explainability and Reliability

Deep learning algorithms are often regarded as"black boxes, "with decision-making processes that lack transparency.This poses challenges for the certification and regulation of safety-critical navigation systems.Solutions include:

• Explainable AI (XAI) :Developing explainable AI technologies to enhance the transparency of algorithmic decision-making processes.

• Formal Verification:Using mathematical methods to formally verify navigation algorithms and systems, ensuring their correctness and reliability under various conditions.

• Simulation-Driven Development:Conducting large-scale simulations to validate algorithm performance in extreme and rare scenarios, thereby improving system robustness.

5.2.3 Energy Consumption and Sustainability

Intelligent navigation systems, especially those in vehicles and drones, face higher demands for energy efficiency and sustainability.Solutions include:

• Low-Power Chip Design:Developing low-power chips and processors optimized for AI and navigation.

• Energy Harvesting Technologies:Utilizing environmental energy sources such as solar, vibration, and thermal energy to extend device operating time.

• Algorithmic Energy Efficiency Optimization:Designing algorithms and models that optimize energy efficiency to reduce computational resource consumption.

5.3 Outlook on Future Research Directions

5.3.1 Intelligence of High-Precision Maps

With the development of autonomous driving technology, the importance of high-precision maps is becoming increasingly prominent.Future research directions include:

• Real-Time Map Updates:Using vehicle crowdsourcing data and deep learning technologies to achieve real-time updates and maintenance of high-precision maps.

• Semantic Mapping:Constructing semantic maps that contain rich semantic information to provide a deeper understanding of the environment for navigation systems.

• Dynamic Mapping:Integrating real-time traffic, weather, and event information to create dynamic and variable intelligent navigation maps.

5.3.2 Human-Machine Collaboration in Autonomous Navigation Systems

Despite significant progress in autonomous navigation technology, human-machine collaboration remains crucial in complex and extreme environments.Future research directions include:

• Driver Status Monitoring:Using bio-sensors and computer vision technologies to monitor the driver's condition and level of attention in real-time.

• Human-Machine Handover Strategies:Researching natural and safe handover strategies to ensure that control can be safely and smoothly transferred from the system to humans in emergency situations.

• Augmented Reality Navigation:Utilizing augmented reality technology to provide intuitive and natural navigation guidance for drivers and passengers.

5.3.3 Cross-Domain Collaboration and Intelligent Transportation Systems

Intelligent navigation technology will be deeply integrated with intelligent transportation systems to form more efficient and safer transportation networks.Future research directions include:

• Vehicle-Road Collaborative Navigation:Using 5G and V2X communication technologies to enable information exchange and collaborative navigation between vehicles and road infrastructure.

• Swarm Intelligence Navigation:Researching collaborative navigation strategies in multi-agent systems to achieve coordinated path planning and obstacle avoidance for fleets of vehicles, drones, and robots.

• Smart City Navigation:Integrating navigation systems with the traffic management, energy, and public service systems of smart cities to provide more comprehensive and intelligent travel services.

VI. Conclusion

This paper systematically analyzes the current status, key technologies, and development trends of intelligent navigation technology.The study shows that multi-sensor fusion and deep learning are the two main drivers of the development of intelligent navigation technology, while 5G communication, edge computing, and the Internet of Things provide real-time and reliable support for intelligent navigation systems.Through empirical analysis, we have verified the effectiveness of the multi-sensor fusion framework in complex environments, with a 30%improvement in positioning accuracy in urban canyons and indoor environments compared to single-sensor navigation, fully demonstrating the advantages of intelligent navigation technology.

In the future, with the continuous development of artificial intelligence, quantum computing, and new types of sensor technologies, intelligent navigation technology will evolve towards higher precision, stronger robustness, and broader application scenarios.The intelligence of high-precision maps, human-machine collaboration in autonomous navigation systems, and cross-domain collaboration with intelligent transportation systems will become important research directions.We believe that intelligent navigation technology will bring revolutionary changes to human travel and drive the advent of the smart city and autonomous driving era.

References

[1] Li, H., &Zhang, W. (2020) .Advances in multi-sensor fusion navigation technology.Navigation, Positioning and Timing, 7 (3) , 1-10.

[2] Wang, Y., &Cai, Y. (2021) .A review of deep learning-based simultaneous localization and mapping (SLAM) .ISPRS Journal of Photogrammetry and Remote Sensing, 178, 239-258.

[3] Chen, M., &Zhao, G. (2022) .Design of real-time navigation system based on edge computing.Acta Automatica Sinica, 48 (2) , 345-355.

[4] Guo, Y., &Wang, H. (2020) .GNSS-based urban positioning using deep learning.IEEE Transactions on Intelligent Transportation Systems, 21 (8) , 3200-3211.

[5] Liu, Q., &Sun, L. (2019) .Research on error compensation technology of inertial navigation systems.Journal of Chinese Inertial Technology, 27 (4) , 456-462.

[6] Yurtsever, E., et al. (2019) .A survey of autonomous vehicles:Enabling technologies, challenges, and open problems.IEEE Transactions on Intelligent Vehicles, 4 (2) , 1-21.

[7] Sun, B., &Li, M. (2021) .Research on distributed navigation system based on federated learning.Control and Decision, 36 (5) , 1123-1132.

[8] Cai, Y., &Wang, Y. (2022) .Quantum navigation:Principles, challenges and opportunities.Navigation, 69 (1) , 1-15.

[9] Zhou, J., &Wu, F. (2020) .Research on intelligent navigation system based on digital twin.Journal of Astronautics, 41 (7) , 801-810.

[10] Chen, X., &Wang, J. (2021) .Semantic mapping and navigation using deep learning.IEEE Transactions on Robotics, 37 (4) , 1050-1063.

Federated Filtering Methods for Non-Recursive Navigation Data Processing

Yu.A. Litvinenko

ITMO University
the Concern CSRI Elektropribor, JSC
Saint Petersburg, Russia
ya_litvinenko@mail.ru

O.A. Stepanov

ITMO University
the Concern CSRI Elektropribor, JSC
Saint Petersburg, Russia
soalax@mail.ru

***Abstract*— The features of federated filtering (FF) algorithms with a nonrecursive measurement processing scheme are analyzed. It is shown that in this case, unlike the recursive algorithm, conditions can be obtained that ensure the coincidence of estimates generated by the FF with the optimal estimate for a wide class of estimated processes. Note that the calculated covariance matrix generated by the FF will coincide with the real covariance matrix of estimation errors, i.e. the algorithm will be consistent. The results obtained are illustrated by the examples.**

***Keywords*— Kalman filter, nonrecursive processing scheme, federated filtering**

I. INTRODUCTION

It is known that in the case of sensor redundancy, the linear estimation problem can be solved on the basis of two variants: the centralized one and the decentralized variant.

The main advantage of the centralized scheme is the possibility of obtaining an optimal—in the mean-square sense—estimate. The known disadvantages of the centralized variant are their lower reliability and higher computational load [1,2].

In the decentralized scheme, initial processing is implemented in local filters (LF) that generate local estimates. The final (global) estimate is calculated by fusion of local estimates in the master filter. In integrated navigation systems with a distributed structure of sensors, which are commonly called a local system, the decentralized variant is commonly implemented using recursive federated filtering algorithms (FFA).

The main advantages of recursive FFAs usually include a lower computational load compared to the centralized Kalman filter and their immunity to false sensor measurements [1,2]. The main disadvantage of FFA is that, in the general case, these algorithms do not ensure optimal estimates and the covariance matrices calculated in the master filter are not real covariance matrices, that is, the algorithm is not consistent. For some particular cases, the conditions of information sharing are formulated for the so-called consistent tuning of LFs, which allows obtaining an optimal estimate in the master filter [2,7]. However, the authors of this paper are not aware of any general answer to the question about the conditions ensuring the coincidence of the estimates and the calculated covariance matrices generated in the master filter with the results obtained in the optimal centralized Kalman filter.

At the same time, the last decade has seen an increased interest in nonrecursive processing methods, including algorithms based on factor-graph optimization [10]. However, to date, FFAs have been studied only in relation to traditional recursive schemes.

In this regard, the aim of the proposed work is to study the FFAs features in the implementation of nonrecursive measurement processing schemes.

This work was supported by the Russian Science Foundation, project no. 23-19-00626, https://rscf.ru/project/23-19-00626/.

II. PROBLEM STATEMENT. RECURSIVE SCHEME FOR CALCULATION OF THE OPTIMAL ESTIMATE

Let us formulate the estimation problem under study. Assume that we have an n-dimensional Gaussian Markov sequence

$$x_k = \Phi_k x_{k-1} + \Gamma_k w_k, x_0 \in N\{x_0; \bar{x}, P_0\}, w_k \in N\{w_k; 0, Q\}, \quad (1)$$

and the following set of vector measurements

$$y_k^j = h_k^j x_k + v_k^j; \quad v_k^j \in N\{v_k^j; 0, R^j\}, j = 1...m, \quad (2)$$

where k is a discrete moment of time, Φ_k and Γ are known $n \times n$- and $n \times p$-dimensional matrices; y_k^j are local l-dimensional j-th vector measurements; w_k is the Gaussian zero-mean p-dimensional vector of system noise with known covariance matrices Q; v_k^j are Gaussian zero-mean l-dimensional vectors of measurement noise with covariance matrices R^j; h_k^j are the known $l \times n$-dimensional measurement matrices; v_k^j, w_k are assumed to be white-noise sequences, independent of each other and of x_0. For the sake of simplicity, without loss of generality, the dimensions of all j-th measurements y_k^j, $j = 1...m$ are assumed to be the same and equal to l. From here on, $a \in N\{a; \bar{a}, A\}$ is used for the Gaussian random vector a with mathematical expectation $\bar{a}$ and covariance matrix A.

The problem is to determine the optimal—in the mean-square sense—estimate of vector x_k at the moment of time k using measurements (2), accumulated by the current k-th moment of time, and the corresponding covariance matrix P_k of the estimation errors.

It is important to note the following feature of the problem under consideration: a set of m measurements generated by separate local sensors is available at each moment of time k.

If we introduce a composite lm-dimensional vector $y_k^C = \left(y_k^1, y_2^2 ... y_k^m\right)^T$, then measurements (2) can be written as

$$y_k^C = h_k^C x_k + v_k^C, \; v_k^C \in N\{v_k^C; 0, R^C\}, \quad (3)$$

where $v_k^C = \left(v_k^1, v_k^2 ... v_k^m\right)^T$ is an lm -dimensional vector, $R^C = diag\left(R^1, ..R^m\right)$ is an lm -dimensional $lm \times lm$ matrix, $h_k^C = (h_k^1 ... h_k^m)^T$ is an $lm \times n$ -dimensional matrix.

The optimal estimate of vector $\hat{x}_k$ and the covariance matrix P_k can be calculated using the well-known recursive equations of the Kalman filter [11].

In what follows, the scheme for obtaining an optimal estimate, which involves processing of the entire set of measurements $j = 1...m$, is called a centralized scheme.

III. Overview of Recursive Federated Filtering Algorithms

As noted in the introduction, the main idea of the FFA is to obtain local estimates $\hat{x}_k^j$ calculated with only the use of measurements y_k^j , $j = 1...m$ and then, to form estimate $\hat{x}_k^G$, called global, by weighing the local estimates in the master filter. Let us explain the FFA in more detail.

In order to obtain local estimates $\hat{x}_k^j$ in the *j*-th LF, instead of Equation (1) for x_k , we introduce a local shaping filter for the sequence

$$x_k^j = \Phi_k^j x_k^j + \Gamma_k^j w_k^j, x_0^j \in N\left\{x^j; \overline{x}^j, P_0^j\right\}, w_k^j \in N\left\{w_k^j; 0, Q^j\right\}, \quad (4)$$

where Φ_k^j, Γ_k^j are the known $n \times p$ -dimensional matrices; w_k^j are the Gaussian zero-mean *p*-dimensional vectors of system noise with the known covariance matrices Q^j . It is assumed that ν_k^j, w_k^j are white-noise sequences independent of each other and of x_0^j . Further, in each LF, local estimates and their covariance matrices $\hat{x}_k^j, P_k^j$ are designed. The global estimate $\hat{x}_k^G$ and its covariance matrix P_k^G in the master filter are calculated at time *k* as

$$P_k^G = (\sum_{j=1}^m \left(P_k^j\right)^{-1})^{-1}, \quad (5)$$

$$\hat{x}_k^G = P_k^G \sum_{j=1}^m \left(P_k^j\right)^{-1} \hat{x}_k^j . \quad (6)$$

It is known [1-9] that, in the general case, the estimate and the covariance matrix calculated in this manner will be different from the optimal estimate and covariance matrix. In [1,2], the authors formulate the conditions for LF tuning under which the global estimate (6) and covariance matrix (6) will coincide with the optimal estimate and the covariance matrix generated in the centralized filter (CF).

To ensure these conditions, matrices Φ_k^j, Γ_k^j, P_0^j , and Q^j are selected. Note that usually, when considering recursive FFAs, it is assumed by default that for all $j = 1...m$: $\Phi_k^j = \Phi_k$, $\Gamma_k^j = \Gamma_k$, and $\overline{x}^j = \overline{x}$, thus, FFAs tuning is reduced to choosing matrices P_0^j and Q^j . It should also be emphasized that FFAs tuning results in the fact that the estimates obtained in the LFs will be different from the optimal local estimates.

Among all possible LF parameter setups, the consistent tuning is most preferential [2]; in this case, estimate $\hat{x}_k^G$ and the covariance matrix P_k^G coincide with the optimal estimate $\hat{x}_k$ and the corresponding covariance matrix P_k obtained with the use of the centralized processing scheme.

It is known [1,2] that it is generally impossible to ensure a consistent tuning by choosing matrices P_0^j and Q^j . In this regard, it is recommended to reduce the FFA errors, as compared to the optimal CF, by the following information-sharing conditions:

$$\sum_{j=1}^m \left(P_0^j\right)^{-1} = P_0^{-1}; \quad (7)$$

$$\sum_{j=1}^m \left(Q^j\right)^{-1} = Q^{-1} . \quad (8)$$

Although these conditions do not ensure the coincidence of $\hat{x}_k^G$ with the optimal estimate, they allow obtaining guaranteed accuracy of FFA estimation in the sense of inequality

$$P_k^G > D_k^G \geq P_k , \quad (9)$$

where D_k^G is the real covariance matrix of the FFA estimate.

Conditions (7), (8) can also be written in a more convenient form using information-sharing factors β^j [1-8]:

$$\left(P_0^j\right) = \beta^j P_0 ; \left(Q^j\right) = \beta^j Q ; \quad (10)$$

$$\sum_{j=1}^m \left(\beta^j\right)^{-1} = 1 . \quad (11)$$

Note that conditions (7), (8) (or (10), (11)) allow a certain freedom in choosing the settings of the LFs.

As noted, in the general case, conditions (7), (8) do not provide a consistent tuning of LF. This can be easily demonstrated using the example of estimating a scalar random walk sequence.

However, for special cases, such as estimation of time-invariant vector x_k, it can be shown that conditions (7), (8) are sufficient to obtain an optimal estimate by FFAs.

Also, in [2], for the problem considered in this paper, it is shown that consistent tuning can be ensured provided that, in addition to (14), (15), the following condition is met:

$$Q_k^j = \Phi_k P_k^j P_k^{-1} \Phi_k^{-1} Q . \quad (12)$$

In this case, matrices Q_k^j will not only be variable, but also asymmetric; for this reason, it is difficult to fulfill the additional condition (14) in practice.

We should also note the case of the random walk estimation using scalar white-noise measurements. In this case, the consistent tuning of the LFs is achieved by fulfilling, in addition to (7), (8), the following conditions for the information-sharing factors β^j [9]:

$$\left(\beta^1\right)\left(R^1\right)^{-1} = \left(\beta^2\right)\left(R^2\right)^{-1} ... = \left(\beta^m\right)\left(R^m\right)^{-1} . \quad (13)$$

To sum up, it can be noted that for the problem under consideration, to obtain an optimal estimate in the recursive FFA, it is not sufficient to comply with the tuning

requirements for (7), (8). Additional requirements for (12) or (13) are difficult to implement in practice or we need to take into account the variances of measurement noise, which significantly narrows the number of options for possible settings of LFs.

IV. NONRECURSIVE FEDERATED FILTERING ALGORITHMS

Let us now specify the statement of the estimation problem under consideration as applied to a nonrecursive scheme.

First, consider the CF. Following [10], we form a composite *n(k+1)*-dimensional state vector $X = (x_0^T,...,x_k^T)^T$ to be estimated, accumulated by time *k* and including all components x_k for $i = \overline{0.k}.$, and *lmk*-dimensional a composite vector of measurements $Y = ((y_1^C)^T .....(y_k^C)^T)$. In this case, the problem of CF estimation can be formulated as follows: estimate the composite vector X by the measurements accumulated by time *k*:

$$Y = HX + V,\ V \in N\{V;0,\mathbf{R}^{CP}\}, \tag{14}$$

in which $H = \begin{vmatrix} 0_{lm\times n} & h_1^C & 0_{lm\times n} & \cdots & 0_{lm\times n} \\ 0_{lm\times n} & 0_{lm\times n} & h_2^C & \cdots & 0_{lm\times n} \\ \vdots & & & \vdots & \cdots \\ 0_{lm\times n} & 0_{lm\times n} & 0_{lm\times n} & \cdots & h_k^C \end{vmatrix}$ - block matrix of $mlk \times n(k+1)$; $\mathbf{R}^{CP} = diag\{0_{lm\times lm}, R^C,..R^C\}$ is a matrix of $lmk \times lmk$ dimension; $V = (0,(v_1^C)^T,...,(v_\kappa^C)^T)^T$ is an *lmk*-dimensional vector. Here and below, $0_{c\times p}$ is a zero $c \times p$ - dimensional matrix.

In this case, the optimal estimate of the composite vector $\hat{X}(Y)$ and its covariance matrix $\mathbf{P}$ can be calculated as

$$\hat{X}(Y) = \bar{X} + \mathbf{P}H^T(\mathrm{R}^{CP})^{-1}\left(Y - H\bar{X}\right), \tag{15}$$

$$\mathbf{P} = \left(\mathbf{P}_X^{-1} + H^T(\mathbf{R}^{CP})^{-1}H\right)^{-1}, \tag{16}$$

where $\mathbf{P}_X, \bar{X}$ are a priori covariance matrix of estimation errors and mathematical expectation corresponding to the composite vector X [11]. Estimate $\hat{x}_k$ and covariance matrix P_k, which we need, are easy to obtain by extracting the corresponding component from vector $\hat{X}$ and the block from the covariance matrix $\mathbf{P}$.

For what follows, it is important that the problem of an *n*-dimensional sequence estimation at the *k*-th time moment is reduced to the problem of estimating the time-invariant *n(k+1)*-dimensional vector.

Now, let us specify the estimation algorithm, as applied to the nonrecursive scheme, using FFA. In this case, in each LF, we need to estimate the composite vector $X^j = ((x_0^j)^T,...,(x_k^j)^T)^T$ of *n(k+1)* dimension, including the components of x_i^j for all time moments $i = \overline{0.k}$, using the *lk*-dimensional vector measurement $Y^j = ((y_1^j)^T,...,(y_k^j)^T)^T$:

$$Y^j = H^j X^j + V^j,\ V^j \in N\{V^j;0,R^{jp}\},\ j \in \overline{1,m}, \tag{17}$$

where $H^j = \begin{vmatrix} 0_{l\times n} & h_1^j & 0_{l\times n} & \cdots & 0_{l\times n} \\ 0_{l\times n} & 0_{l\times n} & h_2^j & \cdots & 0_{l\times n} \\ \vdots & & & \vdots & \cdots \\ 0_{l\times n} & 0_{l\times n} & 0_{l\times n} & \cdots & h_k^j \end{vmatrix}$ is the matrix of $lk \times n(k+1)$ dimension; $V^j = ((v_1^j)^T,...,(v_k^j)^T)^T$ is a composite *lk*-dimensional vector of measurement errors with a diagonal covariance matrix $R^{jp} = diag(R_1^j....R_k^j)$ of $lk \times lk$ dimension.

Optimal estimates in the LFs of vectors X^j by measurement Y^j can be calculated using the formulas similar to (15), (16):

$$\hat{X}^j(Y^j) = \bar{X}^j + \mathbf{P}^j(H^j)^T\left(R^{jp}\right)^{-1}\left(Y^j - H^j\bar{X}^j\right), \tag{18}$$

$$\mathbf{P}^j = \left(\mathbf{P}_{X^j}^{-1} + (H^j)^T\left(R^{jp}\right)^{-1}H^j\right)^{-1}, \tag{19}$$

where $\mathbf{P}_{X^j}, \bar{X}^j$ are the covariance matrices and mathematical expectations corresponding to the composite vectors X^j [11].

The global estimate and covariance matrix of the composite vector $X^G = ((x_0^G)^T,...,(x_k^G)^T)^T$ are similar to (5), (6):

$$\mathbf{P}^{\mathbf{G}} = (\sum_{j=1}^{m}(\mathbf{P}^j)^{-1})^{-1}, \tag{20}$$

$$\hat{X}^G = \mathbf{P}^G\sum_{j=1}^{m}(\mathbf{P}^j)^{-1}\hat{X}^j. \tag{21}$$

It is easy to see that to obtain an optimal estimate of the FFA, it is sufficient to satisfy the conditions:

$$\sum_{j=1}^{m}\mathbf{P}_{X^j}^{-1} = \mathbf{P}_X^{-1}, \tag{22}$$

$$\bar{X}^j = \bar{X}. \tag{23}$$

This is achieved through the following conditions:

$$\sum_{j}^{m}\left(P_0^j\right)^{-1} = \left(P_0\right)^{-1};\sum_{j}^{m}\left(Q^j\right)^{-1} = Q^{-1}. \tag{24}$$

Let us emphasize the coincidence of conditions (24) for the nonrecursive FFA and conditions (7), (8) for the recursive one. In so doing, note that conditions (24) provide consistent tuning only for the nonrecursive filter. This difference between two schemes can be explained by the need to calculate the prediction and its covariance matrix in the recursive scheme, the calculation of which is performed taking into account not only matrices P_0^j and Q^j but also matrix R^j. In the case of the nonrecursive scheme, no additional calculations are needed.

V. AN EXAMPLE

As an illustration, consider the estimation problem of a random walk using measurements with white-noise errors:

$$x_k = x_{k-1} + w_k, x_0 \in N\{x_0;0,\sigma^2\}, w_k \in N\{w_k;0,q\}, \tag{25}$$

using measurements

$$y_k^j = x_k^j + v_k^j; \quad v_k^j \in N\left\{v_k^j; 0, r^j\right\}; \quad j = 1, 2. \tag{26}$$

Here, x, y^j are scalar values, and σ^2, q , r^j are known variances.

For nonrecursive processing schemes, the corresponding vectors and matrices are written as follows: $X = (x_0, ..., x_k)^T$,

$Y = (y_1^1, y_1^2, ..., y_k^1, y_k^2)^T$, $V = (v_1^1, v_2^2, ..., v_k^1, v_k^2)^T$,

$H = \left(0_{2k\times 1} \quad \tilde{h}\right), \tilde{h} = diag\left(h^C .... h^C\right), h^C = (1 \quad 1)^T$,

$R^{CP} = diag(0, 0, r^1, r^2 .... r^1, r^2)$.

Respectively, for the *j*-th LF, we have

$$X^j = (x_0^j, ..., x_k^j)^T, \bar{X}^j = 0, \; Y^j = (y_1^j, ..., y_k^j)^T,$$

$$V^j = (0, v_1^j, ..., v_k^j)^T, H^j = \left(0_{k\times 1} \quad \tilde{h}^j\right), \tilde{h} = diag\left(h^j ..... h^j\right),$$

$$h^j = 1, R^{jP} = diag\left(0 \;\; r^j ... r^j\right) \quad .$$

Hence, the condition of consistent tuning can be written as

$$\begin{vmatrix} \sum_{j=1}^{2}\left(\left(\sigma^j\right)^{-2} + \left(q^j\right)^{-1}\right) & \cdots & 0 \\ -\sum_{j=1}^{2}\left(q^j\right)^{-1} & \cdots & \\ \vdots & \cdots & -\sum_{j=1}^{2}\left(q^j\right)^{-1} \\ 0 & \cdots & \sum_{j=1}^{2}\left(q^j\right)^{-1} \end{vmatrix} = \tag{27}$$

$$= \begin{vmatrix} (\sigma)^{-2} + (q)^{-1} & \cdots & 0 \\ -(q)^{-1} & \cdots & \\ \vdots & \cdots & -(q)^{-1} \\ 0 & \cdots & (q)^{-1} \end{vmatrix}.$$

It is clear that equality (27) is ensured when

$$\left(\sigma^1\right)^{-2} + \left(\sigma^1\right)^{-2} = (\sigma)^2; \left(q^1\right)^{-1} + \left(q^2\right)^{-1} = (q)^{-1} \tag{28}$$

We have already emphasized the coincidence of the obtained conditions for the consistent tuning with the conditions for the recursive FFAs scheme (9), (8). But at the same time, we note that if only conditions (27) are used, the recursive scheme in this problem of the random walk estimation does not provide an optimal solution, as distinct from the nonrecursive scheme for which these conditions are sufficient.

To illustrate the comparison of the recursive and nonrecursive FFAs in the problem of estimating the random walk (25) by measurements (26), we carry out the simulation under the following conditions: σ =10, q=3, r^1=1, and r^2=10. When constructing FF, we consider two options for setting information distribution coefficients β: with equal coefficients β=2 and with the coefficients satisfying condition (13) for the recursive FF.

The simulation results are shown in Figs. 1 and 2.

Figures 1 and 2 show the graphs of the standard deviation (STD) of the estimation error for the recursive scheme: CF (Opt-R) and federated filter (FF-R) and the coinciding graphs for the nonrecursive scheme: CF (Opt-NR) and (FF-NR). Note that for the recursive federated filter, as shown in Fig. 2, the STD of the CF estimate (Opt-R) coincides with the STD of the federated filter estimate (FF-R) only if the additional condition (20) is met. Nevertheless, if the information-sharing factors are equal (Fig. 1), STD of the recursive federated filter differs from that of the recursive CF.

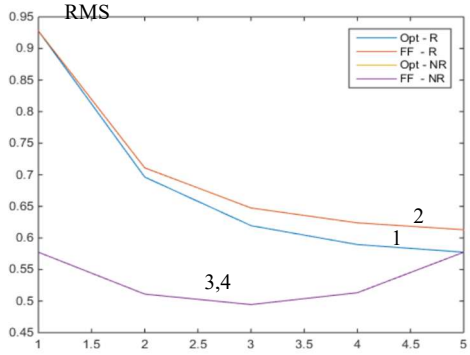


Fig. 1. The RMS errors of the recursive CF (Opt-R,1) and federated filter (FF-R,2) and the nonrecursive CF (Opt-NR,3) and federated filter (FF-NR,4). In FF-R, $\beta^1 = \beta^2 = 2$. Note that Opt-NR and FF-NR coincide, but Opt-R and FF-R are different.

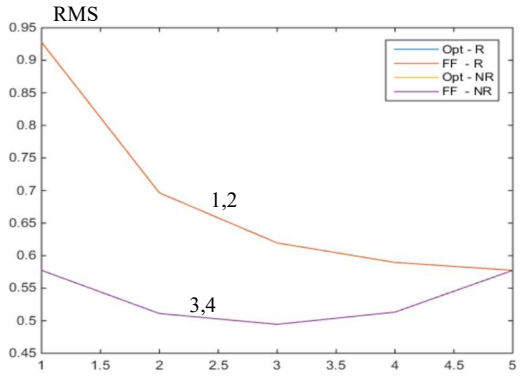


Fig. 2. The RMS errors of the recursive CF (Opt-R,1) and federated filter (FF-R,2) and the nonrecursive CF (Opt-NR,3) and federated filter (FF-NR,4) In FF-NR, the additional condition (20) is met. Note that Opt-NR and FF-NR and Opt-R and FF-R coincide.

Thus, it is clear that the simulation results have confirmed the above conclusions about the possibility of obtaining an optimal estimate using FFAs for a nonrecursive scheme, whereas for a recursive scheme, such an opportunity occurs only if the additional conditions for the FF tuning (13) are taken into account.

In the report, the results obtained are also illustrated by an example of the dead reckoning system correction problem using data on the object's trajectory and measurements of geophysical fields.

VI. CONCLUSION

The federated filtering methods have been considered as applied to the recursive schemes for measurement processing.

The estimation problem solved with the use of federated filtering algorithms is formulated as applied to the nonrecursive scheme; the conditions for the consistent tuning of local filters are given. It has been shown that when the conditions of consistent tuning are met, the estimate obtained

in the nonrecursive FF will coincide, in contrast to the recursive one, with the optimal estimate of the centralized Kalman filter for Markov estimated processes described by a linear shaping filter in the presence of system noise. In this case, the calculated covariance matrix generated by the federated filter will coincide with the real covariance matrix of estimation errors, i.e. the algorithm will be consistent.

The differences between the nonrecursive and recursive schemes, which allow obtaining an optimal estimate using FFAs, lie, first of all, in the fact that the result of the smoothing problem solution on the interval i=0..k in LFs is used to calculate a global estimate. On the contrary, in the recursive scheme, only the result of the filtering problem solution in LF at time k=1 is used.

The results obtained are illustrated using some examples of navigation data processing.

REFERENCES

[1] Carlson N.A. “Federated filter for fault-folerant integrated navigation systems” AGARDograph 331, Aerospace Navigation Systems, June 1995.p. 265-280.

[2] Tupysev V.A. . “Generalized approach to the problem of distributed Kalman filtering”, Proc. AIAA Guidance, Navigation and Control Conference, Boston, 1998,pp. 1097–1116.

[3] Weiwei Lyu, Xianghong Cheng, and Jinling Wang, “Adaptive federated IMM filter for AUV integrated navigation systems”. Sensors 2020, 20, doi:10.3390/s20236806

[4] Pengfei Zhang, Zhenhua Ma, Yin He, Yawen Li, and Wenzheng Cheng. “Cooperative positioning method of a multi-UAV based on an adaptive fault-tolerant federated filter”. Sensors 2023, 23, https://doi.org/10.3390/s23218823

[5] Tarık Ayabakan, Feza Kerestecioglu. “Fault tolerant indoor positioning based on federated Kalman filter”. Journal of Signal Processing Systems,2024 ,273–285

[6] Xu, X., Pang, F., Ran, Y., Bai, Y., Zhang, L., Tan, Z., Wei, C., & Luo, M. “An indoor mobile robot positioning algorithm based on adaptive federated Kalman filter”. IEEE Sensors Journal, 2021?21, 23098–23107

[7] V.A. Tupysev and Yu.A. Litvinenko, “The effect of the local filter adjustment on the accuracy of federated filters”, Proc. of MICNON-2015, pp. 349-354, 2015.

[8] M.S. Selezneva, A.V. Proletarsky, K.A. Neusypin, L. Zhang “Modification of the federated Kalman filter using the observability degree criterion of state variables”, 26th Saint Petersburg International Conference on Integrated Navigation Systems. St. Petersburg: CSRI Elektropribor, 2019.

[9] Li, Z., Gao, J., Wang, J. et al. PPP/INS tightly coupled navigation using adaptive federated filter. GPS Solut21, 137–148 , 2017. https://doi.org/10.1007/s10291-015-0511-z

[10] O.A. Stepanov, A.M. Isaev, A.V. Motorin, Yu.A. Litvinenko, V.P. Zolotarevich, “Recursive and nonrecursive algorithms Applied to navigation data processing: differences and interrelation with factor graph optimization algorithms”, 31st Saint Petersburg International Conference on Integrated Navigation Systems. St. Petersburg: CSRI Elektropribor, 2024.

[11] Oleg A. Stepanov, “Optimal and Suboptimal Filtering in Integrated Navigation Systems, in Aerospace Navigation Systems”, Eds: Alexander V. Nebylov and Joseph Watson, John Wiley & Sons, 2016. https://doi.org/10.1002/9781119163060.ch8

•

Gait Assessment of a Patient with a Modular Foot Prosthesis Using Marker-Based Video Analysis and Electromyography

O.N. Maltseva
Research laboratory "Motion Capture and Simulation Systems"
Saint Petersburg Electrotechnical University "LETI"
Saint Petersburg, Russia
alise_maltceva@mail.ru

***Abstract*—This study presents an instrumented gait analysis of a patient with activity level II-III with an author-developed prosthesis with modular components. The evaluation was performed using marker-based motion capture and surface electromyography to examine the biomechanical and neuromuscular impact of energy-storing midfoot module with different mechanical properties. The findings demonstrate that the prosthetic configuration using a stiffer material improved ankle joint power output and reduced muscle activity in the posterior thigh group compared to the more flexible material, indicating the potential for individualized prosthetic optimization based on the patient's physiological characteristics and functional needs.**

Keywords—Qualisys, foot prosthesis, motion capture, gait, force plate, marker-based video analysis, electromyography, C-motion, biomechanics

I. Introduction

Video motion capture systems—both marker-based and markerless—are widely applied in biomechanical and rehabilitation studies, including gait assessment in patients with cerebral palsy [1] and the evaluation of wearable robotic systems [2]. However, their application in individualized selection of foot prostheses, remains underexplored.

The goal of the study is to investigate the impact of different materials, used in one of the modular components of the developed foot prosthesis, on the patient's gait, considering their varying mechanical properties.

This paper presents an experimental evaluation of an author-developed prosthesis with energy-storing midfoot module using marker-based video analysis, force plates and electromyography (EMG). Previous works, including the author's earlier studies on patient with lower-limb prosthesis gait event detection using multiple force plates [3], as well as the biomechanical modeling and design of 3d-printable modular foot prostheses [4,5] have demonstrated the feasibility of integrating motion analysis into both clinical assessment and prosthesis development workflows. These developments provide a foundation for the current research, aimed at evaluating material-dependent performance of a modular midfoot element in real-world gait conditions.

II. Methods

A. Equipment and Setup

The system included:

- 6 high-speed infrared cameras (Qualisys Miqus M5), operating at 340 frames per second, providing sub-millimeter accuracy for marker tracking;
- 2 force plates (Kistler), embedded in the walkway and operating at 1000 Hz, for synchronized recording of ground reaction forces;
- 2 reference video cameras capturing at 180 frames per second for visual verification;
- Delsys Trigno wireless EMG system;
- A set of 26 (+4 extra) reflective markers was placed according to the Rizzoli Lower Body Marker Set [6].

All components were synchronized using Qualisys Track Manager (QTM) software, which provides real-time marker tracking, integration with analog signals (e.g., EMG), and export of data in standardized formats for further analysis.- Qualisys Track Manager and C-Motion Visual3D for data processing.

This configuration allows high-precision kinematic, kinetic, and EMG analysis in dynamic gait environments.

Due to ongoing patent protection procedures, the detailed geometry of the developed foot prosthesis cannot be disclosed at this stage.

B. Participant

A 34-year-old adult patient with left lower limb amputation and healthy right leg voluntarily participated in this study. The patient's weight was 85 kilograms and height 1.7 meters. The participant demonstrated a relatively stable gait pattern with a gait profile score <8.0 (based on the Gait Profile Score (GPS) scale developed by Qualisys, which quantifies gait quality with lower values indicating more stable and less variable gait patterns). Ethical approval was not obtained for this study as it was considered part of routine clinical evaluation; however, informed consent was provided by the patient.

C. Experimental Protocol

Two interchangeable energy-storing midfoot modules with different mechanical characteristics were tested. Based on stiffness and damping properties, they were conditionally designated as Hflex (higher stiffness) and Mflex (moderate stiffness).

A set of 26 reflective markers was placed according to the Rizzoli Lower Body Marker Set developed by the Istituto Ortopedico Rizzoli for detailed lower-limb kinematic analysis, as described by Cappozzo et al. [6]. Additionally, 4 supplementary markers (2 on the pelvis and 1 on each foot) were used to enhance tracking accuracy of the pelvic and foot segments.

The patient performed 10 walking trials with an author-developed prosthesis with 2 different energy-storing midfoot

modules, to ensure consistency and reliability of the gait analysis results. Each trial was conducted on a level surface, with sufficient rest periods between trials to prevent fatigue.

Surface EMG electrodes were applied in accordance with SENIAM guidelines [7] to record the activity of the m. biceps femoris. EMG data were recorded using a Delsys Trigno wireless EMG system, which provides high-fidelity signals for muscle activation analysis.

The EMG signals were normalized to the patient's body weight (W/kg), allowing for a valid comparison between conditions. To evaluate total muscle activation throughout the gait cycle, the trapezoidal method was used to calculate the area under the curve (AUC) of the EMG signal for each condition (use of prosthesis with Hflex vs. Mflex material)

The formula for calculating AUC:

$$AUC = \sum_{i=1}^{n-1} \frac{EMG_i + EMG_{i+1}}{2} \cdot (\Delta t) \quad (1)$$

where:

EMG_i – EMG values at point i

EMG_{i+1} – EMG values at point $i+1$

Δt – percentage step difference between 2 points

The formula for calculating muscle unloading (MU):

$$MU(\%) = \frac{AUC_{mflex} - AUC_{hflex}}{AUC_{mflex}} \cdot 100 \quad (2)$$

where:

AUC_{mflex} – AUC for condition Mflex

AUC_{hflex} – AUC for condition Hflex

The formula for calculating the average percentage (AU) of MU:

$$AU = \frac{1}{n} \sum_{i=1}^{n} MU_i \quad (3)$$

where:

MU_i – muscle uploading percentage at each individual gait cycle point i

n – total number of data points (0-100)

The study was conducted using an instrumental gait assessment method and clinical movement analysis, proposed by Dr. D.V. Skvortsov [8], as well as kinematic, kinetic, and EMG analysis described by Dr. J. Perry [9].

D. Limitations

The study involved only one participant, which limits the generalizability of the findings. Further studies with larger sample sizes are needed.

Due to the reliance on relatively normal (gait profile score <8,0) gait patterns, the results may differ in cases of atypical gait patterns.

The study was conducted over a short period, and long-term effects of prosthesis use were not assessed.

The participant represented activity level II-III. Results may vary for individuals with different activity levels.

III. Results

The analysis of the goniogram of ankle dorsiflexion in the sagittal plane showed that with the Hflex material, there was less dorsiflexion of the foot at the end of the support phase (25.8°) compared to Mflex (28.5°), which is an important indicator for transitioning from dorsiflexion to plantarflexion.

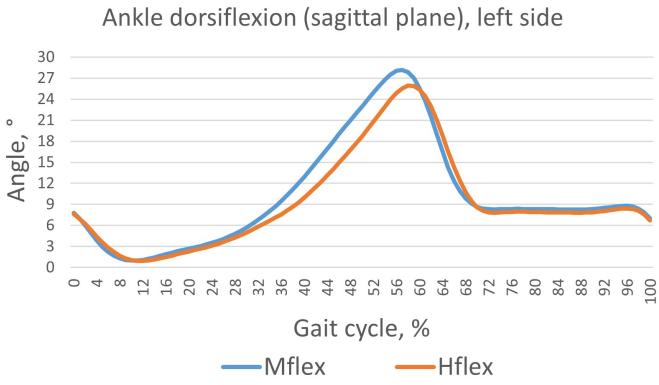


Fig. 1. Ankle dorsiflexion angle (°) in the sagittal plane throughout the gait cycle (%) for prosthesis with Hflex and Mflex energy-storing midfoot module configuration

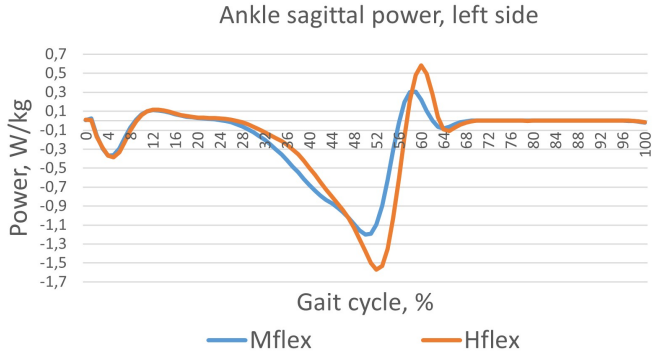


Fig. 2. Ankle power generation (W/kg) in the sagittal plane throughout the gait cycle (%) for prosthesis with Hflex and Mflex energy-storing midfoot module configuration

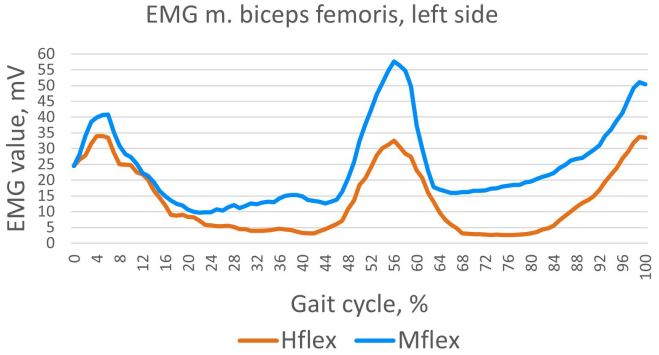


Fig. 3. EMG amplitude (mV) of the m. biceps femoris during the gait cycle (%) for prosthesis with Hflex and Mflex energy-storing midfoot module configuration.

By using materials with different physical-mechanical properties in the prosthesis's modular element, an increase in ankle power generation in sagittal plane at the end of the support phase by 43% was achieved.

The posterior muscle group, specifically the m. biceps femoris, was offloaded during testing with the Hflex material. The AU between the conditions, when using the Hflex energy-storing midfoot module compared to Mflex energy-storing midfoot module, was 50,1%.

IV. Discussion

The results of the present study demonstrate that the material properties of the energy-storing midfoot module in a transtibial foot prosthesis significantly influence key gait parameters. Specifically, the Hflex configuration, characterized by higher elasticity and energy return, provided improved biomechanical outcomes compared to the Mflex.

Kinematic analysis revealed a more favorable ankle dorsiflexion profile in the sagittal plane during the terminal stance phase when using the Hflex module. This suggests improved forward stability at push-off, as the material's elastic properties help control the forward motion of the body and reduce the risk of forward imbalance or falls during the transition from stance to swing phase.

Importantly, surface EMG analysis of the biceps femoris muscle indicated a notable unloading effect with Hflex. The average reduction in muscle activation was 50.1% compared to Mflex, suggesting a lower demand on the posterior muscle group and, potentially, increased gait efficiency and reduced fatigue.

However, unlike many commercial models, the modular nature of the foot prosthesis tested here allows targeted material adjustments without full redesign, providing a pathway toward patient-specific optimization.

While the study provides promising insights, it was limited by the involvement of a single participant. Further studies involving more diverse patient populations and long-term trials are necessary to validate the generalizability of the results and to evaluate the durability of materials under daily use conditions.

References

[1] J. Mazzarella, M. McNally, D. Richie, A.M.W. Chaudhari, J.A. Buford, X. Pan, J.C. Heathcock, "3D Motion Capture May Detect Spatiotemporal Changes in Pre-Reaching Upper Extremity Movements with and without a Real-Time Constraint Condition in Infants with Perinatal Stroke and Cerebral Palsy: A Longitudinal Case Series", Sensors (Basel), 2020 Dec 19;20(24):7312, doi: 10.3390/s20247312.

[2] C.J. Lee, J.K. Lee, "Inertial Motion Capture-Based Wearable Systems for Estimation of Joint Kinetics: A Systematic Review", Sensors (Basel), 2022 Mar 25;22(7):2507, doi: 10.3390/s22072507.

[3] O. N. Maltseva, "Gait cycle detection in patient of lower-limb prosthesis: a comparison of manual and multiple force plates modes," in Proc. Conf. of Young Researchers in Electrical and Electronic Engineering (ElCon), IEEE, St. Petersburg, Russia, 2025, pp. 1829–1833.

[4] O. N. Maltseva, "Design and biomechanical analysis of 3D printable foot prosthesis with replaceable components for transtibial amputee during walking: a case study", M.S. thesis, ETU "LETI," St. Petersburg, Russia, 2024, 107 pp. [in Russian]

[5] O. N. Maltseva, A.Y. Aksenov, R.A. Egorov, Modeling and Manufacturing of a Customized Foot Prosthesis for Biomechanical Gait Assessment via Video Analysis. XVI All-Russian School "Mathematical Modeling and Biomechanics in the Modern University", 2022, p. 62 [in Russian]

[6] A. Cappozzo, F. Catani, U. Della Croce, A. Leardini, "Position and orientation in space of bones during movement: Anatomical frame definition and determination", Clinical Biomechanics, 10(4), 1995, pp. 171-178.

[7] SENIAM, "Surface Electromyography for the Non-Invasive Assessment of Muscle Activity," European Union Biomedical Health and Research Program, 2000. [Online]. Available: http://www.seniam.org

[8] D. V. Skvortsov, Clinical Movement Analysis: Gait Analysis, Moscow: NPTC "Stimul", Ivanovo, 1996, 344 pp. [in Russian]

[9] J. Perry, "Gait Analysis normal and pathological function", P.: SLACK Incorporated, 1992. 524 pp.

Comprehensive Assessment of the GLONASS System Efficiency Using a Simulation Model

T.A. Brovko
Radio systems department
NRU "MPEI"
Moscow, Russia
brovkota@mpei.ru

A.P. Malyshev
Radio systems department
NRU "MPEI"
Moscow, Russia
malyshevap@mpei.ru

V.B. Pudlovsky
Radio systems department
NRU "MPEI"
Moscow, Russia
pudlovskyvb@mpei.ru

***Abstract*—The current criteria for assessing the efficiency of the GLONASS system do not meet modern requirements for the system or its operational conditions. This paper proposes an integrated efficiency criterion that accounts for the characteristics of navigation equipment and their operational conditions within a specified service area of the system. Numerical efficiency assessments were obtained using a simulation modeling method for the GLONASS system.**

Keywords—GLONASS system, navigation support, error, quality assessment, integrated criterion, simulation modeling

I. INTRODUCTION

The quality assessment of GLONASS navigation support is crucial for the effective use of navigation services provided by GNSS receivers (GR). However, evaluating the system's efficiency without considering the performance of GR is inadequate. For instance, it is impossible to assess the impact of interference or the visibility conditions of navigation satellites without accounting for the directional pattern formation specifics of GR.

The aim of this paper is to propose a vector criterion for assessing the quality and efficiency of GLONASS navigation support, which considers the multifactorial influence of various factors on performance metrics.

The article presents a vector of GLONASS system efficiency metrics, including conventional quality indicators. This vector is a multidimensional random variable described by a multivariate joint conditional probability density function (PDF) [3], dependent on the state of the orbital constellation, radio navigation signal propagation conditions, reception conditions, and the impact of various classes of interference. To derive the probability density distributions of quality metrics, the method of statistical experiments using a GLONASS system simulation model is proposed.

II. INTEGRATED EFFICIENCY METRIC OF THE GLONASS SYSTEM

The primary efficiency metrics E and their calculation methodologies, such as positioning error, availability, and continuity of navigation support [4, 5], as well as the interference immunity of GR [1], are well-known and widely covered in literature and regulatory documents [6]. Existing methodologies [2] calculate these metrics independently and do not account for real-world GR operating conditions. This complicates their use for comprehensive navigation condition assessments, such as for aviation users on extended trajectories.

To take into account the factors described above, a vector random variable $\mathbf{X}(\lambda,\varphi,t)$ as shown in expression (1), are introduced:

$$\mathbf{X}(\lambda,\varphi,t)=[\Delta \quad A \quad C]^T, \tag{1}$$

where Δ - is the three-dimensional positioning error, accounting for GR; A - is the availability of usable satellite vehicle (SV) signals at the GR antenna input; C - the probability of obtaining navigation solution.

The vector $\mathbf{X}(\lambda,\varphi,t)$ is a function of space λ,φ at time t characterized by a conditional PDF as shown in expression (2):

$$p(\mathbf{X}(\lambda,\varphi,t)\big|\mathbf{U}(t)) \tag{2}$$

where $\mathbf{U}(t)=[G(t),N(t),J(t)]$, reflects the influence of the following factors on GLONASS system efficiency metrics:

1 The position delusion of precision (PDOP), accounting for all available SV, denoted as G,

2 The number of available SV signals trackable by GR, denoted as N,

3 The signal-to-interference environment, denoted as J.

After collecting sufficient statistics, a multivariate conditional PDF for a single point in space can be obtained. To derive an averaged efficiency assessment for a single point, the multivariate PDF is averaged as shown in expression (3):

$$E(\lambda,\varphi)=\int_{\mathbf{U}}\int_{t} p(\mathbf{X}(\lambda,\varphi,t)\big|\mathbf{U})\,dt d\mathbf{U} \tag{3}$$

For an integrated assessment over a user-specified service area, expression (4) is used:

$$E=\int_{\lambda_{\min}}^{\lambda_{\max}}\int_{\varphi_{\min}}^{\varphi_{\max}} E(\lambda,\varphi)d\lambda d\varphi \tag{4}$$

To derive practically meaningful assessments, the probabilities of exceeding the components of vector $\mathbf{X}$ are calculated using required boundary conditions (5):

$$E_{\delta}(\lambda,\varphi)=\int_{G_{\min}}^{G_{\max}}\int_{N_{\min}}^{N_{\max}} p(\mathbf{X}(\lambda,\varphi,t)\big|\mathbf{U})\,p(\mathbf{X})dGdN \tag{5}$$

The proposed metric allows for calculating the probabilities of exceeding various indicators, such as positioning error or spatial geometric factor, without contradicting methodologies in regulatory documents [2]. A drawback of the proposed method is the need for extensive navigation determinations over long intervals across multiple spatial points. Currently, the GLONASS ground segment's network of navigation receivers is primarily located within Russia [1], complicating real-measurement-based integrated assessments.

However, this integrated assessment can be performed using a GLONASS system simulation model [7], which accurately reflects the current system state, including real data from various Information and Analysis Centers (IAC).

Modern computational capabilities of personal computers and servers enable these calculations within measurable timeframes, making the method practical.

III. USING THE SIMULATION MODEL TO ASSESS THE INTEGRATED EFFICIENCY METRIC

To demonstrate the feasibility of using the simulation model for evaluating efficiency metrics (1) and to highlight differences from existing methods [2], a series of simulation sessions were conducted under various GLONASS operating conditions. Below are the results of a representative scenario:

- GLONASS orbital constellation comprising 24 satellites;
- Airborne UNE traveling from Moscow to Vladivostok;
- Flight duration: 4 hours; evaluation interval for vector $\mathbf{X}$, the vector evaluation interval is 30 seconds;
- Four stationary interference sources along the UNE trajectory with varying power levels.

The locations of interference sources and the flight trajectory are shown in Figure 1.

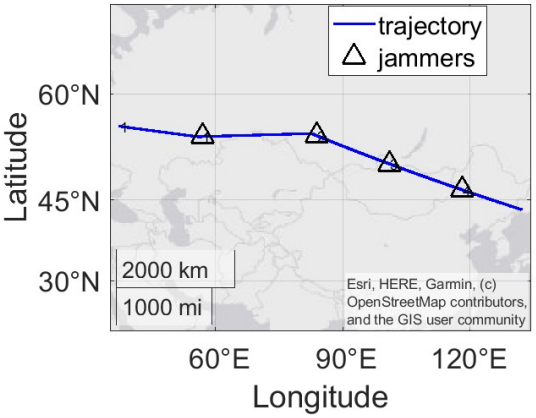


Fig. 1. The locations of interference sources.

Figures 2 and 3 illustrate the temporal changes in component $\mathbf{X}(\lambda,\varphi,t)$ and $\mathbf{U}$, obtained through simulation.

Figure 2 shows changes in the number of available and used satellites for navigation support, and PDOP.

Figure 3 displays the planar coordinate errors along the flight trajectory: errors due to the space segment alone and errors including interference effects.

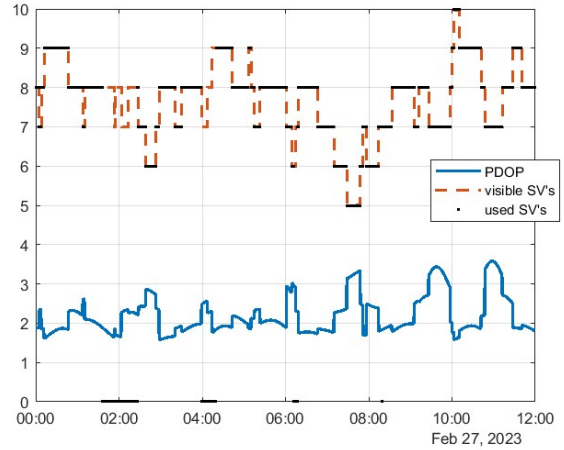


Fig. 2. Changing the number of available and used satellites for navigation support, and PDOP.

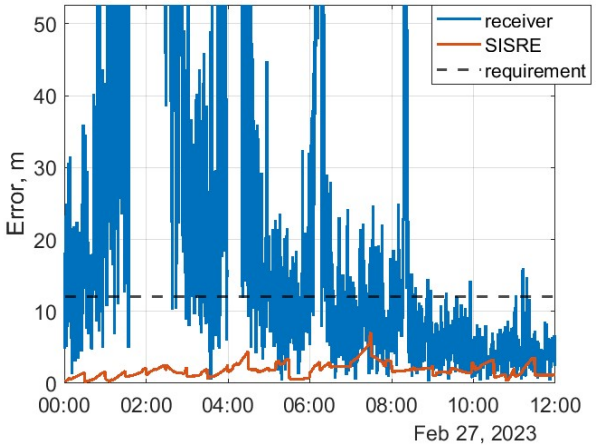


Fig. 3. The planar coordinate errors along the flight trajectory.

Table 1 presents the values of individual components of $\mathbf{X}(\lambda,\varphi,t)$ and the integrated system efficiency assessment, expressed as the percentage of compliance with navigation service requirements over space and time, alongside analogous metrics derived from the GLONASS Open Service Performance Standard methodology [2].

TABLE I. EVALUATION OF THE EFFECTIVENESS OF THE GLONASS SYSTEM

Performance Metric	Percentage of compliance with requirements	
	Integrated Method	***GLONASS Open Service Performance Standard***
Positioning Error (<12 m)	52	100
Availability	100	100
Continuity	т89	100
Efficiency	52	Not assessed

IV. DISCUSSION OF RESULTS

The results in Table 1 and Figure 3 correspond to simulations of GR without integration with inertial navigation systems or spatial interference suppression algorithms. Additionally, the system efficiency value was

derived for a single GR on one trajectory. For assessments involving multiple GR types and a specified service area, the efficiency value may be higher despite low individual GR performance. Thus, the proposed efficiency assessment, accounting for GR specifics and operational conditions within the service area, appears more adequate compared to methodologies in [2].].

V. CONCLUSION

The efficiency metrics of the GLONASS system should be evaluated collectively, justifying the use of the proposed vector-based efficiency indicator.

Modern conditions necessitate including UNE interference immunity in the efficiency vector, as it significantly impacts the continuity of navigation support.

Analytical assessment of GLONASS efficiency across interdependent criteria is complex. Therefore, employing a GLONASS simulation model for this purpose is advisable.

REFERENCES

[1] GLONASS. Modernization and development prospects. Monograph / Edited by A.I. Perov. Moscow, Radiotekhnika Publ., 2020. – 1072 p.: ill. ISBN 978-5-93108-198-4

[2] Standard of Performance Characteristics of Open Service (STECHOS) Revision 2.2 (6.2019) Korolev 2019

[3] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[4] A. Tyulin, V. Betanov, “Guaranteed coordinate-time and navigation support for consumers of satellite system information” Legal Informatics, 2020, vol. 2, pp. 4-16 DOI: 10.21681/1994-1404-2020-2-04-16.

[5] Fundamentals of GLONASS Integrity Assurance / A. G. Milkovsky et al. // Information Technologies. - 2015. - Vol. 21, No. 8. - Pp. 595-601.

[6] GOST R 52865-2009. Global navigation satellite system. Parameters of radionavigation field. Technical requirements and test methods.

[7] T. A. Brovko, A. P. Malyshev, V. B. Pudlovsky, P. A. Baranova. Using a GLONASS Simulation Model to Assess the Quality of Coordinate-Time Determinations // Modern Problems of Radio Electronics. 2024. Pp. 38-41.

Comparison of Criteria for Choosing the Optimal Constellation of Navigation Satellites for Location Determinations in Conditions of Abnormal Use of the Navigation Field of the GNSS

V.I. Baburov

«Navigator» JSC
Saint Petersburg, Russia
info@navigat.ru

N.V. Ivantsevich

«Navigator» JSC
Saint Petersburg, Russia
info@navigat.ru,

BGTU VOENMEH
Saint Petersburg, Russia
sciencebstu@bstu.spb.su

N.V. Vasilyeva

«Navigator» JSC
Saint Petersburg, Russia
info@navigat.ru

***Abstract*—A comparison is made of the information characteristics of the working constellations of the GNSS navigation satellites, corresponding to different criteria for choosing a working constellation, during navigation definitions in emergency situations.**

Keywords—navigation satellites systems, working constellations of the GNSS navigation satellites, emergency situations, simulation modeling.

I. INTRODUCTION

Satellite radio navigation systems (GNSS) GLONASS and GPS are currently used in various fields of practice [1-3]. They were created on the basis of the scientific and technical reserve that existed in the world in the second half of the last century. The relevant questions then were about the computing capabilities that can be used in consumer navigation equipment. Therefore, during the development of GPS and at the first stage of the introduction of satellite location methods, in particular, when choosing a specific location algorithm, the criterion of optimal choice of a working constellation consisting of four navigation spacecraft (NS) from among those that fall within the range of radio visibility of the consumer was applied. This criterion is applicable when working on one GNSS. If the navigation field of two systems is available, for example, GLONASS and GPS, then there should be at least five NS in the working constellation. In some versions of consumer navigation equipment, such algorithms are still implemented. The question of choosing a working constellation of the NS is also relevant due to the fact that navigation fields from four navigation satellite systems are currently available: GLONASS, GPS, Galileo and BeiDou [4, 5], which significantly increase the computing load on the computer. On the other hand, work is underway to explore the possibilities of using the navigation field of satellite systems for location determination in emergency situations. The redundancy of the navigation field can be used to solve this problem. The criterion for choosing a working constellation depends on the specifics of specific emergency situations.

The article compares the information characteristics of the working constellations of the GNSS navigation satellites, corresponding to different criteria for choosing a working constellation, during navigation definitions in emergency situations.

II. ABNORMAL SITUATIONS WHEN USING THE NAVIGATION FIELD

Abnormal situations of receiving navigation satellite signals are situations that differ from those listed in the Interface Control Documents for satellite navigation systems [6, 7]. Such situations include the following conditions for receiving satellite signals on the consumer:

- the presence of reflected signals, for example, from the aircraft body and from other objects; form a working constellation by increasing the permissible minimum elevation angle of the spacecraft. This leads to the exclusion from the constellation of all spacecraft "visible" by the consumer with low elevation angles;
- roll–pitch motion of the object being determined or violation of the condition of permissible orientation of the axis of symmetry of the radiation pattern of the receiving antenna; - cases of shadowing of parts of the surrounding space in the upper hemisphere;
- interference with the reception of individual satellite signals

and other situations.

These situations lead to a transformation of the working constellations, to a change in the geometric factor and, consequently, the accuracy of navigation definitions on the consumer.

III. INVESTIGATION OF THE COMPOSITION AND INFORMATION CHARACTERISTICS OF OPERATIONAL CONSTELLATIONS OF NAVIGATION SATELLITES IN EMERGENCY SITUATION

Location information received from navigation satellites of one or two navigation systems will be characterized by the amount of information and Shannon entropy. These information characteristics will change in case of emergency situations. A change in these characteristics uniquely

determines a change in the accuracy of determining coordinates according to the satellite system. In turn, the accuracy of navigation definitions is proportional to the geometric factor (GF) of the working constellation of navigation satellites. It can be shown that under certain conditions, the error of relative navigation definitions is also determined by the geometric factor of the working constellation of satellites. The change in the entropy of the working constellations of navigation satellites in these abnormal situations will be equal to the change in the entropy of the geometric factors of these working constellations. Therefore, to compare the information characteristics of the working constellations selected according to the compared optimality criteria, the geometric factor of the NS working constellation was further used.

A working constellation consisting of four NS, with relative navigation definitions based on one GNSS, is estimated in [8] as significantly dependent on its constituent NS. The results of calculations are given, according to which a change in such a working constellation, even by one NS, can cause a four-fold increase in the measurement error of coordinates. Therefore, it is concluded that it is necessary to force the selection of identical NS constellations on two interacting consumers with relative navigation definitions. Taking into account the fact that during inter-flight navigation, the probability that the satellite constellations will differ by no more than one NS, the factor of deterioration in the accuracy of navigation definitions will vary from two to four. Therefore, it is concluded that it is necessary to force the selection of a working constellation that is the same on interacting objects. To implement this method, data exchange between interacting consumers will be required. Thus, the implementation of such a method requires, in addition to the standard navigation equipment on the consumer, additional software, hardware and a data transmission line.

We study the effect of the transformation of the NS working constellation as a result of the exclusion of individual satellites from it on the accuracy of absolute and relative navigation definitions. To compare the information capacity of one navigation satellite in a working constellation consisting of four navigation satellites and in a working constellation including all navigation satellites entering the radio visibility zone, mathematical simulation using the Monte Carlo method was performed. The situations "NS_i - Consumer" were modeled, i = 1, 2, ..., the number of navigation satellites located in the radio visibility zone of the consumer [9]. The visible satellites were determined for the set values of α_{min}, which are the permissible minimum elevation angles of the satellites. The initial data for modeling were taken from the IAC website [10]. Geometric factors PDOP were calculated for the constellations of all navigation satellites visible from the consumer. The consumer was located at a point with random coordinates. The measurements were carried out at a random time from the repeatability interval of the satellite configuration. Then one or more satellites were excluded from the working constellation and geometric factors were determined again. Based on these data, the coefficient of deterioration of the geometric factor k was calculated with a decrease in the number of satellites in the working constellation by n NS. The simulation results are presented in table 1. In the header of the fourth column, the expression $P_1 = P(Nn < 5)$ means the probability that the number of visible NS after the exclusion of randomly selected n satellites became less than acceptable. In the fifth column, $P_2 = P(PDOPn > 6)$ means the probability that after the elimination of n NS, the value of the geometric factor in position became more than six. These conditions correspond to the Interface Documents for the systems [6, 7].

TABLE I. THE STATISTICS OF GF GLONASS+GPS WITH THE EXCEPTION OF NS

n	Stat	$PDOP_n$	P_1	P_2
0	Mean RMS k	1.148 0.117 1.000	0.000	0.000
1	Mean RMS k	1.193 0.133 1.039	0.000	0.000
2	Mean RMS k	1.245 0.154 1.084	0.000	0.000
3	Mean RMS k	1.303 0.179 1.134	0.000	0.000
4	Mean RMS k	1.373 0.213 1.194	0.000	0.000
5	Mean RMS k	1.458 0.261 1.270	0.000	0.000
6	Mean RMS k	1.560 0.335 1.354	0.000	0.000

Then the maximum number of satellites excluded from the constellation was determined, which had the same effect as when changing the working constellation in the form of an optimal four navigation satellites per satellite. As follows from the data shown in table 1, the random exclusion of even six satellites from the constellation of all visible GLONASS+ GPS satellites results in a deterioration of GF by only 35.4%. At the same time, in the remaining constellation, the conditions corresponding to the interface documents are still fulfilled with a probability of at least 0.9992%. The considered option corresponds to abnormal situations caused by the presence of natural or artificial interference to the signals of individual NS. One of the ways to protect against interference coming into consumer navigation equipment at small angles relative to the local horizon is to use in consumer navigation equipment an artificial increase in the permissible elevation angle of navigation satellites above the horizon of the consumer $\alpha_{min.}$ For the practical implementation of this method, it is enough to change only one setting, which sets the limit for the permissible elevation angle of the satellites. At the same time, the working constellation will include fewer NS. However, as follows from table 1, in the NS constellations, when working on two GNSS, the exclusion of even six NS does not lead to a significant deterioration in the accuracy of navigation definitions. The research of this method was carried out by simulation modeling. The results are presented in table 2. From the data shown in table 2, it follows that when working on two systems, GLONASS + GPS, it is possible to increase the minimum allowable angle of the NS to α_{min} to 25°, and when working on one system – no more than 15 °. From the data shown in table 2, it follows that when working on two systems, GLONASS + GPS, it is possible to increase the

minimum allowable angle of the NS to α_{min} to 25°, and when working on one system – no more than 15 °.

TABLE II. GF STATISTICS AT DIFFERENT PERMISSIBLE ELEVATION ANGLES OF THE NS

α_{min}	**GNSS**	Mean (GF)	RMS (GF)	P_2
5°	GLONASS	1.77	0.21	0.000
	GLONASS + GPS	1.20	0.10	0.000
10°	GLONASS	2.13	0.44	0.000
	GLONASS + GPS	1.44	0.17	0.000
15°	GLONASS	2.59	0.69	0.003
	GLONASS + GPS	1.74	0.27	0.000
20°	GLONASS	3.18	0.94	0.013
	GLONASS + GPS	2.18	0.42	0.000
25°	GLONASS	3.91	0.92	0.055
	GLONASS + GPS	2.77	0.52	0.000
30°	GLONASS	4.61	0.75	0.296
	GLONASS + GPS	3.69	0.70	0.063

IV. CONCLUSION

Thus, it is established that the conclusions obtained for the information capacity of one navigation satellite in different working constellations corresponding to different criteria differ significantly. It is useful to take the data obtained into account when forming the "appearance" of the algorithmic support for the navigation equipment of the GNSS consumer operating in absolute or relative mode, under conditions different from the standard ones specified in the Interface Control Documents for satellite navigation systems.

REFERENCES

[1] Global'nyi plan aeronavigatsii 2016–2030 (Global Air Navigation Plan for 2016-2030), Mezhdunarodnaya organizatsiya grazhdanskoi aviatsii (International Civil Aviation Organization), Ed. 5, 2016.

[2] Osnovy gosudarstvennoj politiki Rossijskoj Federacii v Arktike na period do 2020 goda i dal'nejshuyu perspektivu (Fundamentals of the state policy of the Russian Federation in the Arctic for the period up to 2020 and beyond) (approved by the President of the Russian Federation on 18.09.2008 N Pr-1969).

[3] GLONASS. Principy postroeniya i funktsionirovaniya (GLONASS. Principles of construction and functioning), Edited by A.I. Perov, V.N. Kharisov. 4th ed. revision. and additional M.: Radio Engineering, 2010.

[4] Kozadaev K.V., Mitrakhovich I.O. Search for the minimum required constellation of satellites for positioning tasks in global navigation satellite systems, in: Informatika, vychislitel'naya tekhnika i upravlenie. Informaciya i kosmos (Informatics, Computer Engineering and Management. Information and Space), no.1, 2019, pp. 104 -110.

[5] Zhilinsky V.O., Gagarina L.G. Algorithms for selecting navigation spacecraft when solving a navigation problem, in: Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta (Bulletin of the Voronezh State Technical University), vol. 17, no 6, 2021, pp. 43–55.

[6] Global'naya sputnikovaya navigacionnaya sistema (Global Satellite Navigation System GLONASS, An interface control document), Revision 5.1. – Moskva, 2008.

[7] ICD-GPS-200. Technical Characteristics of the NAVSTAR GPS.

[8] Orlov, V.K., Gerchikov, A.G., Chernyavskij, A.G., Lokal'nye radiotekhnicheskie sistemy mezhsamoletnoij navigatsii (Local Radio Engineering Systems of Inter-aircraft Navigation: Izd-vo LETI, 2011.

[9] Baburov, V.I., Ivantsevich, N.V., Vasil'eva, N.V., Panov, E.A., Sovmestnoe ispol'zovanie navigatsionnykh polei sputnikovykh radionavigatsionnyh sistem i setei psevdosputnikov (Joint Use of Navigation Fields of Satellite Radio Navigation Systems and Pseudo-Satellite Networks), St. Petersburg: Izd-vo Agentstvo RDK-Print, 2005.

[10] http:/www.glonass-iac.ru/glonass/

Research on Integrated Navigation Algorithm Based on Radio Navigation Station

Yang Guoliang
Xi'an Flight Automatic Control Research Institute, Aviation Industry Corporation of China
Xi'an ShaanXi,710076, China
yanggl009@avic.com

Zhang Yangkang
Xi'an Flight Automatic Control Research Institute, Aviation Industry Corporation of China
Xi'an ShaanXi,710076, China
zhangyk033@avic.com

Zu Zhaozi
Xi'an Flight Automatic Control Research Institute, Aviation Industry Corporation of China
Xi'an ShaanXi,710076, China
zuzz@avic.com

***Abstract*—Performance based Navigation (PBN) requires higher flight accuracy and integrity of civil aircraft. PBN includes two basic navigation specifications: Area Navigation (RNAV) and Required Navigation Performance (RNP). In the concept of PBN operation, the aircraft can fly along any desired path and the navigation accuracy meets the requirement of Required Navigation Performance (RNP). The navigation mode's accuracy based on radio navigation station, depends on the station selection strategy due to the relatively fixed position of the navigation station to a great extent. Therefore, it is very important to study the accuracy of integrated navigation algorithm based on radio navigation station in any flight path under different station selection strategies.**

Airborne navigation equipment of civil aircraft in the route area includes Inertial Navigation System (INS), Global Positioning System (GPS), Air Data System (ADS), Distance Measure Equipment (DME), VHF Omnidirection Range (VOR) and so on, which the position mainly depends on the location of INS and GPS separately, and the location of DME/ DME, VOR/ DME joint positioning. Factors affecting the joint positioning position include the distance between the navigation station and the aircraft, the angle between the navigation station and the aircraft, the positioning performance of the navigation station and the navigation duration of the navigation station, etc.

In this paper, the station selection strategy is studied in the route area using machine learning algorithm, and the radio navigation positioning result is combined with the INS and GPS navigation results to verify the influence of different station selection strategies on the integrated navigation algorithm. The test result of Beijing-Shanghai flight route shows that the station selection strategy in this paper has better precision and continuity indicator, which can meet the requirement of RNP-2.0.

Keywords—integrated navigation algorithm, radio navigation, performance based navigation

I. INTRODUCTION

Radio Navigation as the cornerstone of modern aviation navigation technology, has undergone multiple stages of evolution since its application in the early 20th century. Due to its high signal strength and wide availability range, it still serves as an important means in aviation navigation technology. The traditional radio navigation method requires the aircraft to fly towards or departure radio navigation station to maintain a good signal strength, but to a certain extent, it loses the performance indicators of flight[1]. Performance-based Navigation (PBN) is a new generation of navigation concept proposed by the International Civil Aviation Organization (ICAO)[2]. PBN takes flight navigation Performance as the core requirement, rather than relying on specific radio navigation stations(NAVAID)[3].

PBN integrates multiple navigation methods and advanced navigation management algorithms, enables aircraft to fly along the optimal flight path while maintaining the best navigation performance indicator[4]. The integrated Navigation method based on radio navigation stations is an important means to achieve the operation of PBN. While making full use of advantages of radio navigation stations, it can be combined with other airborne navigation methods, such as Inertial Navigation System (INS), Global Navigation Satellite System (GNSS), and so on, which can improve the utilization rate of airspace, reduce flight costs and enhance flight safety, to a certain extent[5].

II. RADIO NAVIGATION STATION

As one type of modern aviation navigation technologies, radio navigation stations, due to their specific geographical locations, enable aircraft to selectively use them at various flight stages. Generally, radio navigation stations include the following categories: Very High Frequency Omnidirectional Range beacon station (VOR), Distance Measuring Equipment (DME), Non-Directional Beacon (NDB), Instrument Landing System (ILS), etc.

The DME navigation station, through the query-response mechanism, transmits response signals after a fixed delay after receiving the tuning instructions from the airborne equipment. After receiving the response signal, the on-board equipment calculates the total time of query-response, and then calculates the slant distance between the aircraft and navigation station. The VOR navigation station transmits modulated signals of reference phase and variable phase. The on-board receiver can calculate the azimuth of the aircraft relative to the VOR station based on the phase difference of the received signals.

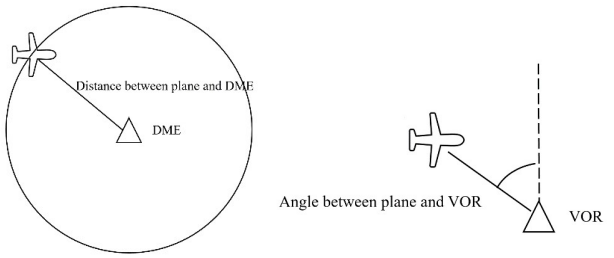


Fig. 1. The left figure shows the measured slant distance between the aircraft and the DME navigation station under the query-response mechanism, and the right figure shows the azimuth of the aircraft relative to the VOR station calculated by the on-board equipment through the measurement of the phase difference.

The DME navigation station can return the slant distance information between the station and aircraft. As long as the slant distance information of three navigation stations can be

obtained simultaneously, the position of the aircraft can be uniquely determined. When with two navigation stations, it is necessary to make calculations based on the aircraft's position in order to calculate the real-time position of the aircraft more accurately. The VOR navigation station has certain errors in positioning due to the azimuth measurement information it returns. However, it can be combined with the DME navigation station to achieve navigation station positioning by integrating distance and azimuth measurement information. Meanwhile, the VOR navigation station can also be used as directional information source. DME and VOR navigation stations can be used for aircraft positioning because which return distance and azimuth measurement information respectively, and have also become common method of navigation and positioning in radio navigation.

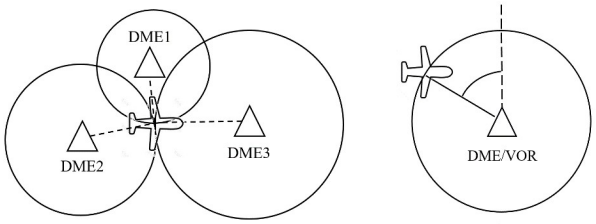


Fig. 2. The left figure is the principle of the aircraft achieving navigation and positioning through three DME navigation stations, and the principle of the aircraft achieving navigation and positioning through the DME/VOR co-located navigation station is showed on the right.

Early radio navigation methods required aircraft to fly towards or departure from the station to ensure that the radio navigation signal strength is sufficient, so that navigation and positioning could be carried out based on the information returned by the radio navigation station. Once the direction is deviated from, it will cause the signal of the navigation station to weaken and thereby affect the navigation accuracy. Then which can lead to the actual flight path of the aircraft being composed of segment broken line, which seriously wastes airspace resources and also affects the performance of the aircraft.

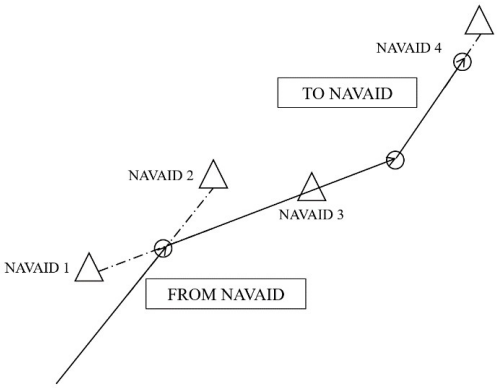


Fig. 3. Schematic diagram of flying towards or departure from the station. In the first section of the flight path, the aircraft is flying towards NAVAID 2, which means towards the station. In the second section of the flight path, it is flying departure from NAVAID 1, which means departure from station.

With the development of modern navigation technology, flight performance has drawn increasing attention from researchers. As a result, PBN operation emerged, and radio navigation methods have been able to play a greater role under the PBN concept due to their high precision and stability. With the development and maturation of the PBN theory, the requirements for the use of radio navigation methods have been relaxed. Aircraft can no longer fly towards or departure navigation stations, but also can choose the most suitable flight route according to the flight destination. Appropriate navigation stations can be choosing at any position during the flight as the navigation source undoubtedly greatly improves economic benefits and airspace utilization[6].

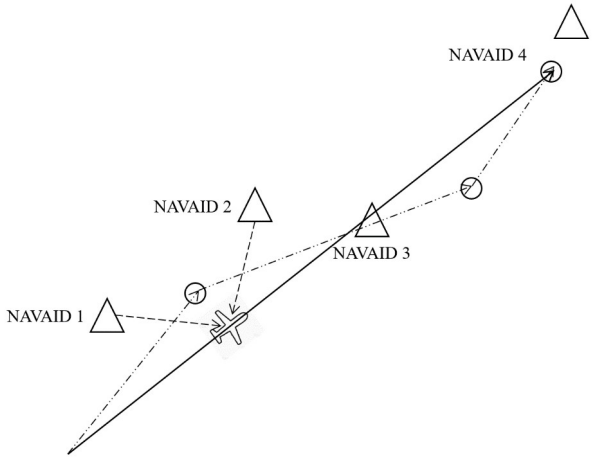


Fig. 4. Schematic diagram of the PBN operation. The aircraft no longer has to fly towards or departure navigation stations but can directly fly to the destination, during which the most suitable navigation stations can be selected for position positioning.

Under the PBN operation system based on radio navigation stations, how to select the most suitable navigation stations, which is the selection strategy of the navigation station, will be crucial to the actual flight effect. Suitable navigation stations can improve navigation accuracy and integrity to a certain extent, while choosing unsuitable ones will seriously affect navigation performance. Generally speaking, suitable navigation stations have the characteristics of being close to the aircraft and being less affected by propagation errors.

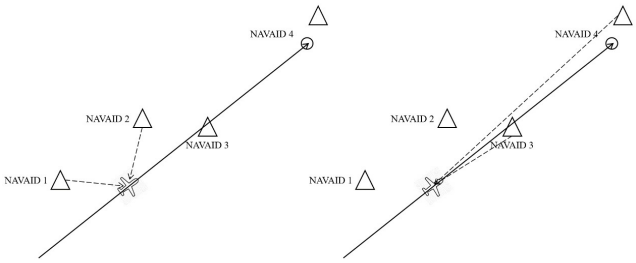


Fig. 5. Comparison diagram of navigation stations selection strategy. Navigation stations selected in the left figure are more suitable. The distance between the navigation stations is shorter, and the error caused by propagation has a smaller impact on the positioning result. Navigation stations selected in the right figure are less suitable. On the one hand, the distance of selected navigation stations is relatively far. On the other hand, even a slight error in measuring the slant distance may lead to a significant positioning deviation.

III. GROUND SIMULATION VERIFICATION

Under the PBN operation system based on radio navigation stations, how to select the most suitable navigation station has become a key factor determining the positioning accuracy. Therefore, a large number of

experimental studies need to be conducted in the early stage to determine the optimal station selection principle. The high cost of actual flights test determines that it is very necessary for carrying out some functional verifications in the ground simulation environment.

By integrating DME and VOR navigation stations in the ground simulation environment, the radio navigation scene during the flight can be restored to a certain extent, while the inquiry response mechanism of the airborne receiver is realized by simulating the feedback information of the ground station through the actuator.

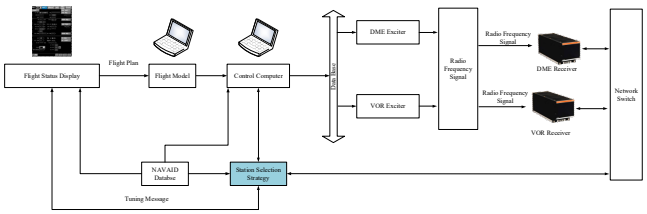


Fig. 6. Schematic diagram of ground simulation environment. The flight scene is driven by the flight plan and flight simulation model. The selected navigation station information is sent to the exciter, and the feedback signal simulated by the exciter from the radio navigation station is sent to the receiver. The receiver then sends out the measurement information, ultimately achieving the simulation of radio navigation scene in the ground environment.

The ground simulation environment can be combined with the flight status of the aircraft to achieve verification in both static and dynamic scenarios. Throughout the entire process, the exciter simulates radio frequency signal of the desired navigation station and sends it to the receiver. The receiver tunes the specific navigation station based on the analog signal of the exciter and finally performs positioning calculation based on the measurement information returned by the receiver. The expected navigation station is the most suitable screened out from the database based with station selection strategy. Combined with the results of aircraft positioning calculation, the advantages and disadvantages of the station selection strategy are analyzed and evaluated.

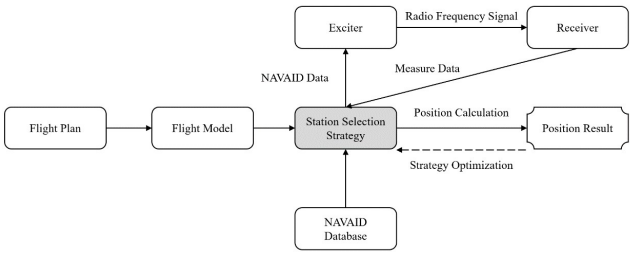


Fig. 7. Flowchart of the station selection optimization strategy. The aircraft positioning of the complete flight plan is realized in the ground simulation environment. Based on the positioning results, the station selection strategy is optimized in return, ultimately making the positioning results as close as possible to the real time position of flight simulation model.

For the DME navigation station, when conducting position calculation, the slant distance information between the aircraft and navigation station returned by the receiver should be adopted. Therefore, the accuracy and stability of its data are crucial for positioning. In the ground simulation environment, by measuring the return distances of different navigation stations at the same position and the return distances of the same navigation station at different positions, with conducting multiple repeated tests, the accuracy and stability evaluation results of the return distance can be obtained.

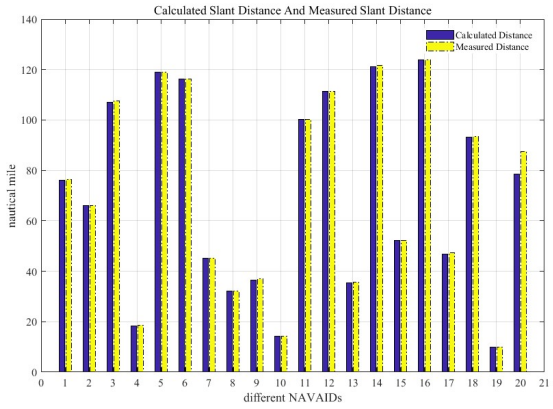


Fig. 8. Graph of DME receiver measured distance and calculated distance error results. A comparative analysis is made between the measured and calculated distances of 20 navigation stations with the same aircraft position. The maximum error is 10.82%, and the minimum one is 0.00436%.

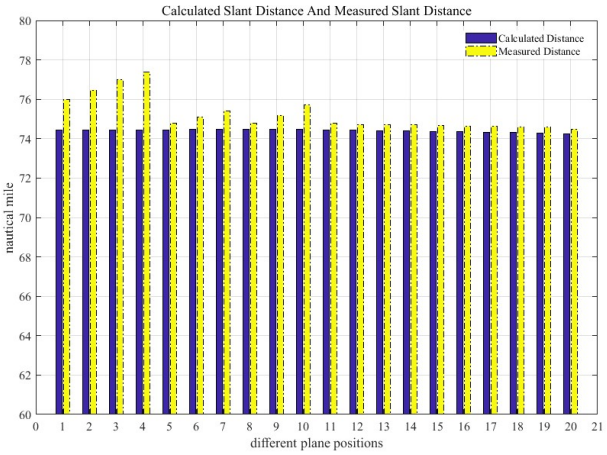


Fig. 9. Graph of DME receiver measured distance and calculated distance error results. Based on the comparison between the measured and calculated distances of the same navigation station with 20 different aircraft positions. The maximum error is 3.88%, and the minimum one is 0.30%.

For the VOR navigation station, when conducting position calculation, the azimuth information of the aircraft and the navigation station returned by the receiver should be adopted, and then combined with the slant distance information returned by the DME receiver for position positioning. In the ground simulation environment, by measuring the return azimuth of different navigation stations at the same position and the return azimuth of the same navigation station at different positions, with conducting multiple repeated tests, the accuracy and stability evaluation results of the return azimuth results can be obtained.

The actual errors of radio navigation stations mainly consist of signal attenuation and angular dispersion errors caused by the propagation of radio signals in the air. In the ground environment, this part of the error is mainly manifested as the radio frequency signal propagation error between the exciter and receiver. This part of the error

mainly affects the positioning results and accuracy indicators.

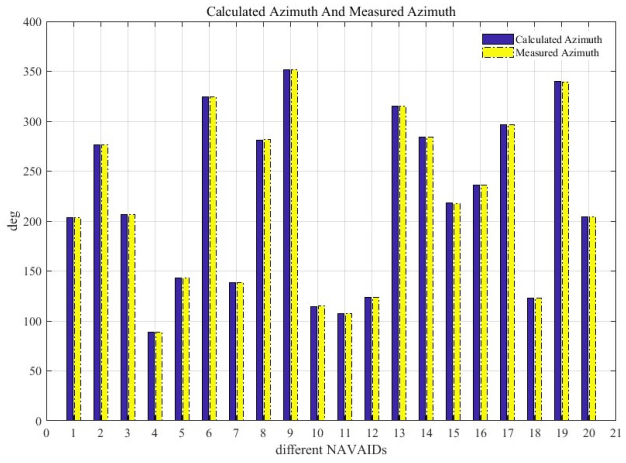


Fig. 10. Results of VOR receiver measured azimuth and calculated azimuth error. The measured azimuth and calculated azimuth of 20 navigation stations at the same aircraft position are compared and analyzed. The maximum error is 0.27%, and the minimum one is 0.00687%.

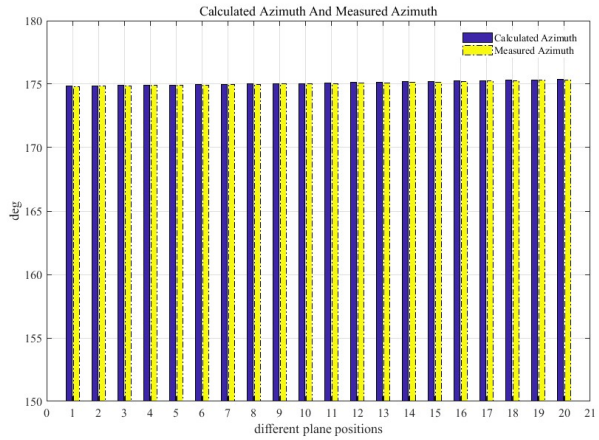


Fig. 11. Results of VOR receiver measured azimuth and calculated azimuth error. The measured azimuth and calculated azimuth of the same navigation station and 20 different aircraft positions are compared and analyzed. The maximum error is 0.0206%, and the minimum error is 0.0193%.

IV. Integrated Navigation

The integrated navigation method based on radio navigation stations includes the combination of the navigation method based on radio navigation stations with INS, GNSS and other methods. Generally speaking, the accuracy of the navigation method based on radio navigation stations lies between INS and GNSS method. On the one hand, it can be used as a modification of INS because its positioning accuracy is not affected by time. On the other hand, it can serve as a supplement to GNSS, making up for the deficiency that satellite navigation may have poor positioning integrity due to reasons such as the atmospheric environment [7].

When the navigation method based on radio navigation stations is combined with INS, it can maintain high precision navigation and positioning for a long time to a certain extent. Its principle relies on the correction of the positioning results of INS based on the positioning results of radio navigation stations. Usually, the positioning result of INS can be corrected based on the error between INS and radio navigation after continuous positioning for a certain period of time. When the error is greater than that of radio navigation, the positioning result of radio navigation can be used to correct the positioning result of INS.

The high precision feature of GNSS is often used as a supplement to other navigation methods. For radio navigation methods, which can be combined with GNSS to achieve more accurate positioning results. Usually, based on the error of the radio navigation method, when the error exceeds a certain threshold, GNSS positioning position can be adopted as the reliable positioning result, thereby making up for the deficiency of the large positioning error of the radio navigation method in some cases.

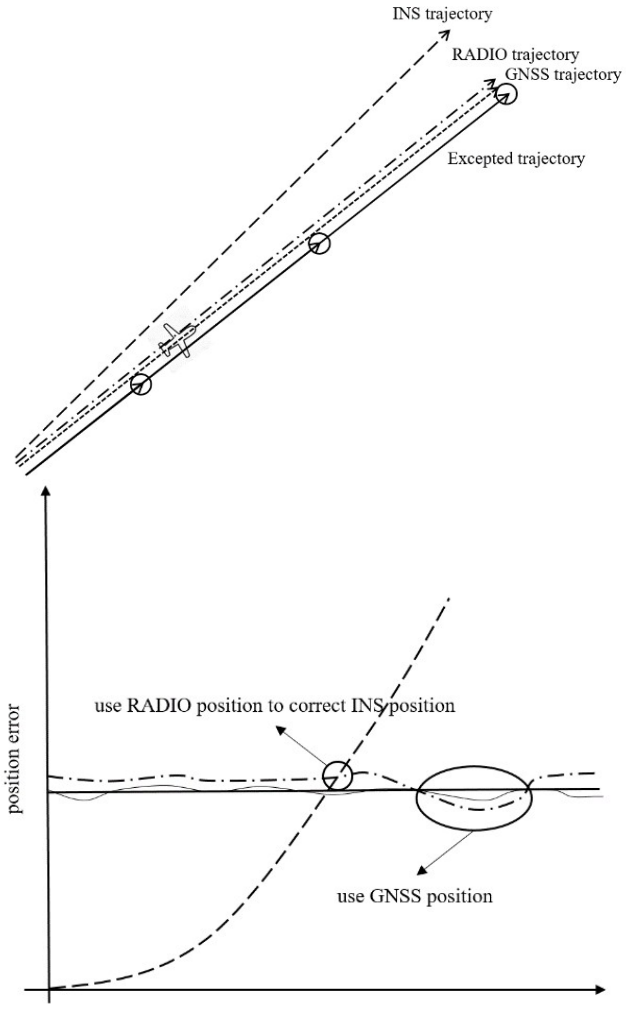


Fig. 12. The principle of integrated navigation method. The left figure shows the deviation diagrams of radio navigation flight trajectories, inertial navigation trajectories, satellite trajectories and expected trajectories. The right figure shows the error distribution of each navigation mode. When the inertial navigation error exceeds the radio navigation error, the radio navigation position is used to correct the inertial navigation position. When the radio navigation error is compared with the satellite navigation error, the satellite navigation position is used as the positioning position.

V. Simulation Test

When the navigation method based on radio navigation stations is actually used at the airborne, it will be affected by various situations such as the station selection strategy, the

distribution of navigation stations and flight route. And the station selection strategy can have a significant effect on improving the flight performance when the flight route and the distribution of navigation stations are relatively fixed. Taking a flight route from Beijing Capital International Airport to Shanghai Hongqiao International Airport as an example, experimental analyses are conducted on the positioning results and error indicators under different station selection strategies.

TABLE I. LIST OF NAVIGATION STATIONS SELECTED BY RADIO NAVIGATION STATION SELECTION STRATEGY

Number	DME1 Frequency(MHz)	DME2 Frequency(MHz)
1	113.6	117.2
2	117.7	112.1
3	114.7	117.7
4	115.2	113.7
5	112.1	116.6
6	114	112.8
7	116.85	113.7
8	117.15	114.3
9	113.5	113.7
10	117.15	112.8
11	112.8	115.3
12	113.3	112.8
13	115.75	115.3
14	113.8	109.8
15	117.05	112.5
16	117.05	113.1
17	115.15	115.3
18	115.9	113.8
19	113.2	113.8

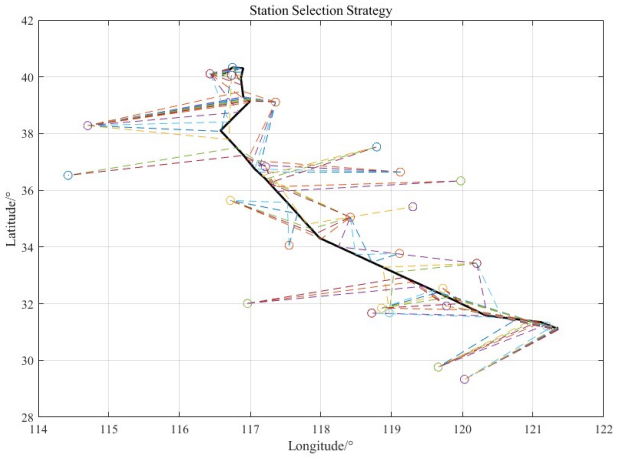


Fig. 13. The selection strategy of radio navigation station. According to the flight route from Beijing Capital International Airport to Shanghai Hongqiao International Airport, the experiment is carried out in the ground simulation environment. The station selection strategy can select navigation stations in the whole flight route, and 19 DME navigation station pairs are selected, and the switching times and positioning accuracy are suitable.

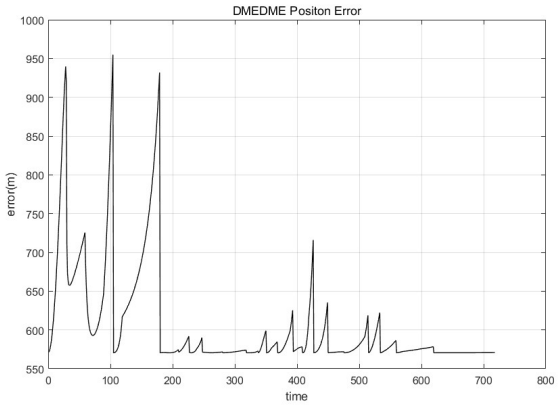


Fig. 14. Positioning error of integrated navigation method based on radio navigation station. The positioning error does not exceed 600m in ground simulation test. In some times, less navigation stations and stations change errors cause error to increase, but the whole can meet the operation requirements of RNP-2.0 in route area.

VI. CONCLUSION

This paper studies the integrated navigation method based on radio navigation station under PBN operating system, analyzes and summarizes the characteristics of radio navigation station, and compares the DME ranging error and VOR angle measurement error in the ground simulation environment. The results show that the simulation environment error can support the validation of integrated navigation algorithm based on radio navigation station. At the same time, the integrated navigation method based on radio navigation station is introduced and tested on Beijing-Shanghai flight route. The verification results can meet the requirements of navigation and positioning on the whole flight route, and the positioning error can meet the requirements of RNP-2.0.

REFERENCES

[1] Park S, Lee H. Transition from Ground-Based to Performance-Based Navigation: Global Trends and Challenges. Transportation Research Part A: Policy and Practice, 2021,143, 1-15.

[2] ICAO. Performance-Based Navigation (PBN) Manual (Doc 9613). 4th Edition. International Civil Aviation Organization, 2013.

[3] Liu, J. The Evolution and Current Status of Aeronautical Radio Navigation Systems. IEEE Transactions on Aerospace and Electronic Systems, 2017, 53(4), 1727-1740.

[4] Zhao, Y, Zhang, W. Performance-Based Navigation (PBN) and its Application in Civil Aviation. Journal of Navigation, 2018, 71(3), 567-585.

[5] FAA. National Airspace System Navigation Strategy. Federal Aviation Administration, 2020.

[6] Park S, Lee H. Transition from Ground-Based to Performance-Based Navigation: Global Trends and Challenges. Transportation Research Part A: Policy and Practice, 2021, 143, 1-15.

[7] Zhang, L. Integration of Traditional Radio Navigation and PBN in NextGen and SESAR. Aerospace Science and Technology, 2022.

Calculation of Secondary Navigation Parameters Based on the Data of the Short-Range Radionavigation System

E.S. Zemlyanyy
Ph.D, Assoc. prof.
Head of navigation sector
JSC RPKB
Ramenskoye, Russia
ezemlyany@rpkb.ru

M.V. Tektov
1 cat. Mathematician
JSC RPKB
Ramenskoye, Russia
mtektov@rpkb.ru

K.A. Kravtcov
1 class Navigatior
JSC Gromov Flight Instityte
Zhukovsky, Russia
pochta_adres@mail.ru

***Abstract*— A method is proposed for determining secondary navigation parameters, namely, the ground speed and the actual track angle, according to the data of the short-range radionavigation system.**

***Keywords*— navigation measurements, radio beacons, position coordinates**

I. INTRODUCTION

The short-range radionavigation system (SRRS) is designed for the formation and subsequent output of the following navigation parameters to the onboard equipment complex [1] of an aircraft:

- azimuth and range to the radio beacon in the navigation mode;
- deviations from the course and glide paths in the landing mode.

The report examines the processing of data from the SRRS received in the navigation mode, these data are considered to be primary.

II. THE MAIN PART

Currently, the problem of forming and calculating the current coordinates of the aircraft in the onboard equipment according to the SRRS data has been solved [2]. The calculation is carried out as follows: according to the primary navigation data, as well as according to the previously known coordinates of the SRRS radio beacon, the current coordinates of the aircraft are calculated [3]. This problem is well known and is being solved in many aircraft and helicopter control systems. At the same time, the task of determining the so-called secondary navigation parameters, such as ground speed and actual track angle according to the SRRS data, has not been solved at the moment and is relevant.

The purpose of the work is to study and develop methods for generating secondary (indirect) navigation parameters based on the SRRS data, such as ground speed and the actual track angle of the aircraft.

The closest analogue for comparison may be the satellite navigation system (SNS), which generates both primary parameters - the coordinates of the aircraft, and secondary navigation parameters – the actual track angle and ground speed. At the same time, when calculating such parameters from the SRRS, there are some specific features, namely: the frequency of updating data on azimuth and range channels differs.

The following method is proposed for calculating the actual track angle and ground speed:

1) Based on the azimuth and range to the radio beacon and the previously known coordinates of the radio beacon, the coordinates of the aircraft location are calculated: latitude and longitude.

2) When changing coordinates at each moment of time, by solving the inverse geodetic problem, the bearing and the distance between the initial and final coordinates of the aircraft are calculated for each iteration.

3) The resulting distance is used to calculate the ground speed, and the bearing is used to calculate the actual track angle.

4) By applying several filters sequentially, the final calculations of the actual track angle and ground speed are formed.

Figure 1 shows spherical angles in a flat view, which determine the actual track angle and ground speed.

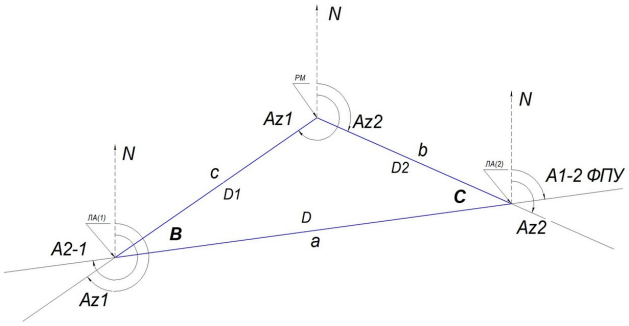


Fig. 1. For the algorithm for calculating secondary parameters

Notation in the drawing:

AC (1, 2) is the position of the aircraft on the first (second) measuring clock of the algorithm; Az1 (2) is the azimuth measured by the radar on the first (second) measuring cycle of the algorithm; D1 (2) is the range measured by the radar on the first (second) measuring cycle of the algorithm. The parameters a, b, and c are spherical ranges [4] used in the final calculations.

In accordance with the information communication protocol, the SRRS outputs the azimuth and range in both filtered and unfiltered form. The results of calculations and flight experiments have shown that unfiltered data should be

used in the formation of secondary navigation parameters. In addition, it is necessary to take into account the peculiarities of filtering angular parameters, since the transition [0..360] affects the operation of filters.

III. Flight Experiment Results

The developed algorithms were implemented in the on-board software. Based on the results of flight tests, graphs of secondary navigation parameters were obtained

- Figure 2 shows a graph of changes in the calculated ground speed;
- Figure 3 shows a graph of changes in the calculated track angle.

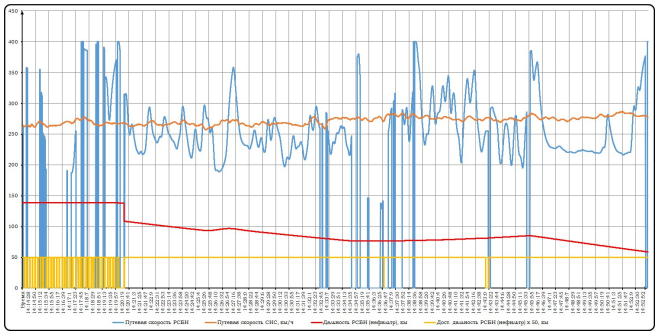


Fig. 2. Calculated ground speed

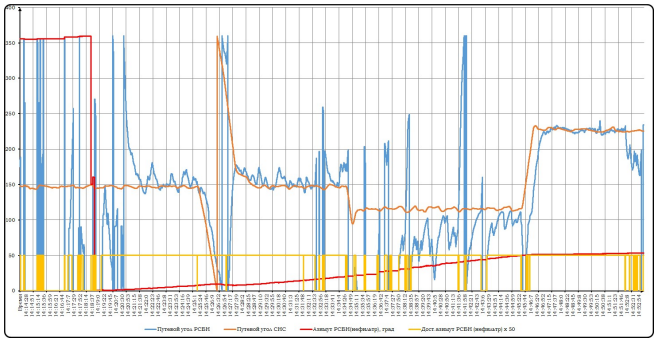


Fig. 3. Calculated track angle

IV. Conclusion

The fundamental possibility of determining secondary navigation parameters (ground speed, actual track angle) by SRRS data is shown.

The experimental results showed that the "azimuth filtered" and "range filtered" parameters calculated and issued by the SRRS introduce errors in coordinate measurements, therefore, the use of the "azimuth unfiltered" and "range unfiltered" parameters of the SRRS has a higher priority.

In addition, it was found that it is necessary to take into account the instability of signal reception in the immediate vicinity of the beacon. The distance from the beacon also directly affects the quality of the calculation.

It is advisable at a separate stage to evaluate the errors of the SRRS in determining azimuth and range and compensate for these errors.

References

[1] Augustov L.I., Babichenko A.V., Orekhov M.I., Shkred V.K. Navigation of aircraft in near-Earth space, M: Facets of success, 2022

[2] Kucheryavy A.A. On-board information systems. A course of lectures. Ulyanovsk: UlSTU, 2004

[3] Lebedev M.I. Airplane navigation. A training manual for pilots and navigators of civil, military transport and strategic aviation. Stavropol: Lit., 2003.

[4] Morozov V.P. Course of spheroidal geodesy, Moscow: Nedra Publ., 1979.

About the Implementation of the Direction Finding Function in the On-Board DAAR of the UGV

S.S. Golubev
JSC Radar mms
Saint Petersburg, Russia
golubev_ss@radar-mms.com

A.V. Nemov
JSC Radar mms
Saint Petersburg, Russia

K. Y. Cheredeev
JSC Radar mms
Saint Petersburg, Russia
cheredeev_kju@radar-mms.com

Due to the development of radio electronics and the increasing complexity of the electromagnetic environment, there is a tendency to use L-band digital antenna arrays in control systems (CS) of unmanned ground vehicles (UGV), railway locomotives, on special vehicles of the Ministry of Internal Affairs, etc. With the help of these devices, a significant increase of the signal-to-noise ratio at the input of the receiving equipment is achieved. As a result, the accuracy and reliability of determining coordinates and receiving information on board the UGV are improved and, accordingly, deviations from the set route and the scheduled departure time to the destination are reduced. It is advisable to give the on-board navigation DAAR the function of bearing sources of powerful, exceeding the intrinsic noise of the receivers of signal sources (SS) in the L frequency range, among which there may be "parasitic" radiation that interferes with driving. These radiations can potentially be affected by satellite television, in the form of harmonics from transmitters of the meter-high aircraft landing system, aircraft emergency transmitters, side bands of the spectrum of transmitters of radio relay communication lines, amateur communication stations and devices "jamming" anti-theft systems used by criminals. Then the signals of the listed bearing-localized SSare filtered by known methods in the DAAR.

During the construction of the equipment, bearings are calculated in a complex digital field using a nonlinear pseudo-estimation of the angular spectrum from the MUSIC family of estimates. This assessment is effective when measuring the bearings of targets that are not solvable by the Rayleigh criterion, a consistent signal model, and a sufficient signal-to-noise ratio.

The purpose of the material is to present the results of a study of a DAAR construction option that meets the conditions of installation on a UGV, proposals on the DAAR structure and algorithms for measuring bearings in the L-frequency bands.

The results are the structure of the on–board radar for measuring bearings, an analytical description of the high-resolution bearing estimation, and a comparison with the monopulse method.

Practical significance – the theoretical and practical results obtained can be used in the design of on-board direction finders of "parasitic" radiation sources designed to operate in a complex electromagnetic environment.

Keywords – direction finding, DAAR structure, UGV, pseudo-estimation of the angular spectrum of MUSIC.

I. INTRODUCTION

Due to the development of radio electronics and the significant complication of the electromagnetic situation, there is currently a tendency to use L-band DAAR in UGV and other vehicles. With the help of these devices, a significant increase in the signal-to-noise ratio at the input of the receiving equipment is achieved. As a result, the accuracy and reliability of determining coordinates and receiving information on board the UGV are improved and, accordingly, deviations from the set route and the designated time of reaching the target are reduced. It is useful to give the on-board DAAR the function of bearing powerful, exceeding the intrinsic noise of the SS receiver. If the direction finding is performed with a resolution of several closely spaced SS, the accuracy of UGV driving will increase, since signals from interfering SS localized in the bearing can be compensated by known methods in the DAAR, or such SS can be eliminated by organizational measures.

Carts for vehicles have specific features, in particular, increasing the size of the antenna system beyond 180x180 mm, and the total weight of the DAAR beyond 1.0 kg will make the equipment not competitively capable.

To improve the characteristics of determining angular coordinates, a promising solution is to use combined algorithms in the DAAR using the monopulse method at the first stage and the so-called super-resolution estimates of angular coordinates at the second stage. In practice, antenna systems are limited to a small number of radiators (no more than 8).

II. TASK

The following tasks are set in the material:

- to substantiate a variant of the DAAR structure with a highly efficient bearing measurement function for high-power emitting SS in the L-band that meets the installation conditions for UGV and other venichles;
- define high-resolution targets by angle;
- describe the high-resolution SS direction finding algorithm in comparison with the monopulse method.

The error in estimating the bearing is investigated as the scattering of the estimate of the angular parameter, which is set by the direction of incidence on the AS plane of the front of the analyzed spatially coherent signal received against the background of quasi-white channel noise when exposed to an interfering spatially coherent signal (signals). The wave front is considered flat, the SS is located in the far zone. The signals can be random and at the same time statistically independent.

III. RESOLUTION CAPACITY

In a complex electromagnetic environment, in the presence of interfering SS and a small relative opening of the speaker, classical direction finding methods, usually quasi-optimal in the situation of a single SS operating against the background of the receivers' own noise, in particular the monopulse method, cannot qualitatively measure the angular parameters under the action of interfering signals. This fact is explained by the limitation of the resolution by the interval

according to the Rayleigh criterion [1]. This interval is inversely proportional to the opening size of the speaker. Recall that the measure of resolution according to the Rayleigh criterion or the resolution interval is the half-width of the main lobe of the DP AS at the zero level [1].

The solution is to apply super-resolution bearing estimates. In the literature [2], estimation data generation methods are understood as methods for determining the directions to the SS, which make it possible to measure bearings with an angular separation of the SS less than the Rayleigh resolution interval. It is more appropriate to call these methods and, accordingly, estimates not super-resolution but high-resolution.

The deterministic Rayleigh resolution criterion only qualitatively characterizes the effectiveness of high-resolution estimates, and does not allow them to be compared with each other with a fixed AS opening.

In the developed approach, the resolution is proposed to be described by scattering the estimates of the bearing of the SS against the background of the action of the interfering SS. This scattering characterizes the statistical resolution of estimates, in contrast to the Rayleigh criterion, which is deterministic. The results of applying both criteria are comparable, since they determine the resolution in terms of angle measurement.

IV. DAAR STRUCTURE

When building the equipment, direct digital conversion of a multi-channel radio signal from high to video frequency in a radio frequency ADC is used. Bearings are calculated in a complex digital field using a nonlinear pseudo-estimation of the angular spectrum from the MUSIC family of estimates. This estimate is effective when measuring the bearings of targets that are not solvable by the Rayleigh criterion, with a sufficient signal-to-noise ratio.

Figure 1 shows the structure of a promising data center for UGV equipment. The radio signals received by the 8-channel microwave path are pre-separated by bandpass filters in the L1 and L2 frequency ranges and combined at the ADC inputs into a single signal. Additional filtering by frequency ranges is performed at the video frequency in RF ADC processors. The estimation and accumulation of correlatiom matrix (CM), its decomposition according to its own system, and digital signal processing (DSP) for estimating bearings are parallelized to eliminate the need for scanning with a "beam" through space. DSP is implemented in FPGAs, for example, the Xilinx series. The ADC data is exchanged with the FPGA via high-speed JESD204B ports.

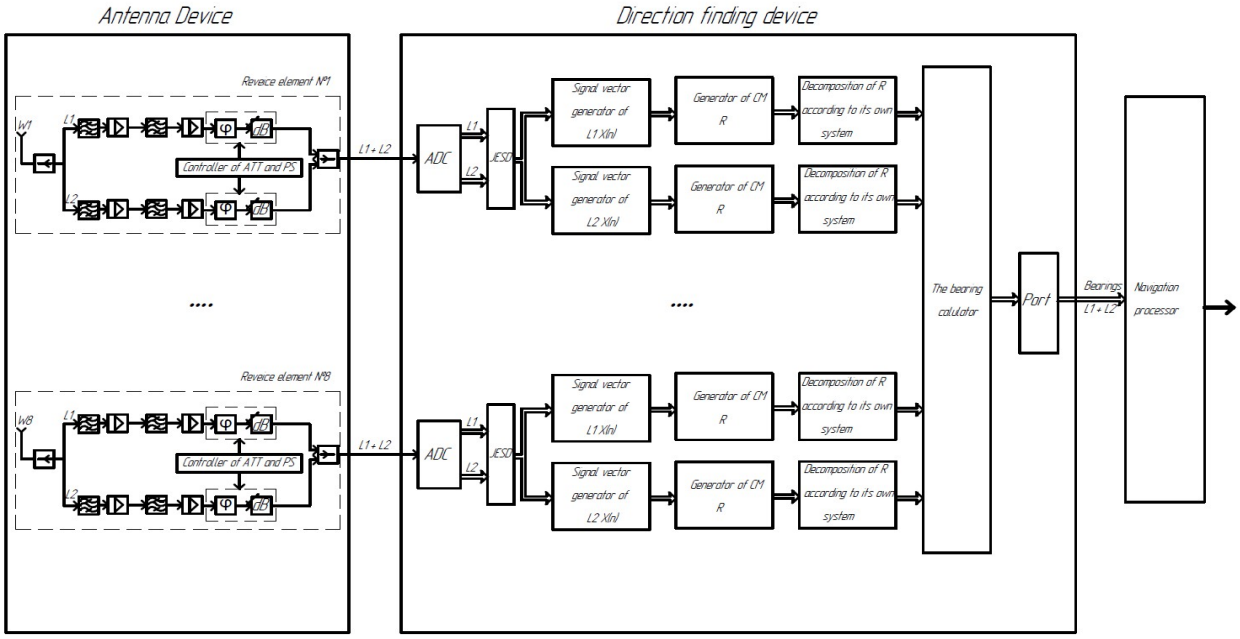


Fig. 1. L-band DAAR structure with bearing measurement function

V. HIGH-RESOLUTION BEARING MEASUREMENT ALGORITHMS

Under the assumption of a point-type SS, we consider the use of high-resolution pseudo estimates of the angular parameter. One of the most effective is the assessment of MUSIC (Multiple Signal ClassifSSation) [3] and its modifications. MUSIC operates in the noise subspace of its own data matrix system, which can be an estimated CM of a spatial signal received by a spatially spaced speaker from several RE. For comparison, the traditional bearing estimation formed by the monopulse method is also considered. Below, this method is called the weight summation (WC) method.

The signal model

Let the received RE AS signal be a narrow-band space-time signal. Let's imagine a signal propagating in the far zone from a point SS in the form of a complex signal

$$\dot{s}(t) = \dot{S}(t) \cdot e^{j2\pi \cdot f_0 \cdot t}, \quad (1)$$

where:

$$\dot{S}(t) = abs\left(\dot{S}(t)\right) \cdot e^{j \arg\left(\dot{S}(t)\right)} = A(t) \cdot e^{j\Phi(t)} \quad (2)$$

– complex signal amplitude $A(t)$ and $\Phi(t)$ are the amplitude and phase envelopes of the received oscillation; f_0 – the carrier frequency of the signal oscillation.

The real signal $s(t)$ is related $\dot{s}(t)$ to the ratio:

$$s(t) = \mathrm{Re}\{\dot{s}(t)\} \tag{3}$$

Assuming $A(t)$ and $\Phi(t)$ to be slowly changing functions over the time τ_{M-1} of delays in the propagation of the signal through the opening of the speaker, the expression for the complex spatial signal in the speaker takes the following form:

$$\begin{cases} \dot{s}_1(t) = \dot{s}(t) = \dot{S}_1(t)\cdot\exp\{j2\pi f_0 t\} \\ \dot{s}_2(t) = \dot{s}(t-\tau_1) = \dot{S}_2(t)\cdot\exp\{j2\pi f_0\tau_1\}\cdot\exp\{j2\pi f_0 t\} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dot{s}_M(t) = \dot{s}(t-\tau_{M-1}) = \dot{S}_M(t)\cdot\exp\{j2\pi f_0\tau_{M-1}\}\cdot\exp\{j2\pi f_0 t\} \end{cases} \tag{4}$$

Let's imagine (4) as a column vector of complex envelopes in the n-th time frame:

$$\begin{bmatrix} \dot{S}_1(n) \\ \dot{S}_2(n)\cdot\exp\{j2\pi f_0\tau_1\} \\ \dots\dots\dots\dots \\ \dot{S}_M(n)\cdot\exp\{j2\pi f_0\tau_{M-1}\} \end{bmatrix} \tag{5}$$

We will consider the background noise to be an additive white Gaussian random process, including RE noises, amplifier noises recalculated to the antenna input of the equipment, , frequency instabilities of generators, etc. For the vector of complex amplitudes of the input process X(n) on the AS (the vector of the amplitude-phase distribution of the signal), omitting the initial oscillation phases that are the same for all RE, from (5) we get:

$$\mathbf{X}(n) = \begin{bmatrix} A_1(n) \\ A_2(n)\cdot\exp\{j\mu(\varphi,\theta)c\tau_1(\varphi,\theta)\} \\ \dots\dots\dots\dots \\ A_M(n)\cdot\exp\{j\mu(\varphi,\theta)c\tau_{M-1}(\varphi,\theta)\} \end{bmatrix} + \mathbf{n}_{uu} \tag{6}$$

where: $\mu(\varphi,\theta)$ – spatial frequency of the signal; c – the speed of light; φ, θ – azimuth and additional angle of the SS location respectively.

It is assumed that the AS is located in the horizontal plane and its horizontal axis is oriented to the north.

For a linear equidistant speaker, the complex amplitudes on the RE are determined by the expression:

$$S_i(\mu) = A_i\exp(j\frac{2\pi}{\lambda}\sin\theta di) = A_i\exp(j\mu di) \tag{7}$$

where: $i = 1,\dots,M$; M – the number of RE AS; λ – wavelength; d – the distance between neighboring RE equidistant AA; $\mu = 2\pi/\lambda\sin\theta$.

The maximum likelihood method

Multi-channel estimation of the angular parameter is performed using a quasi-deterministic signal against the background of point noise-like interference with a known direction of arrival and white noise. Let's use an equidistant linear antenna array in a narrow-band space-time system.

We use an approximation of the general asymptotic expression to estimate the maximum likelihood of the angular parameter of a signal with an unknown amplitude and initial phase against the background of interference with a known spatial correlation[4].

The logarithm of the likelihood ratio:

$$\hat{\mu} \to \max_{\mu} \frac{|R|^2}{q^2},\ q^2 >> 1, \tag{8}$$

where:

$R = \frac{1}{2}\int_{-\infty}^{\infty}\mathbf{X}^T(t,\mu,\nu)\mathbf{W}^*(\mu,\nu)W^*(t)dt$ — the complex correlation integral;

$\mathbf{W}(\mu,\nu) = \mathbf{R}_n^{-1}\mathbf{S}(\mu)$ — the weight vector of optimal spatial processing;

$W(t) = 2S(t)$ — the weight function of optimal time processing; $\mathbf{X}^T(t,\mu,\nu)$ — column vector of the received AS signal; μ – spatial frequency of the signal; ν – spatial frequency of interference; $\mathbf{R}_n$ — interference correlation matrix; $S(\mu)$ — the amplitude-phase distribution of the signal over the speaker opening, given in general by formulas (6); $\dot{S}(t)$ — the law of time variation of complex signal oscillation amplitudes, common to all antenna elements; $q^2(\mu)$ — signal-to-noise ratio.

$$S_i(\mu) = \exp(j\frac{2\pi}{\lambda}\sin\theta di),\ i = 1..M \tag{9}$$

where: M — number of antenna elements of AS; λ — wavelength; d — the distance between neighboring RE; μ – spatial frequency of the target; θ – the angle of incidence of the signal wave on the phased array; $\mu = \frac{2\pi}{\lambda}\sin\theta$; $q^2 = 2E_0\mathbf{S}^T(\mu)\mathbf{W}^{-1*}\mathbf{S}^*(\mu)$,

where: $E_0 = \frac{1}{2}\int_{-\infty}^{\infty}|S(t)|^2dt$.

The complex correlation integral R is represented as:

$$\mathbf{R} = \mathbf{X}^T\mathbf{W}^*(\mu,\nu) \tag{10}$$

where:

$X^T = \left\|\frac{1}{2}\int_{-\infty}^{\infty}X_i(t,\mu,\nu)W^*(t)dt\right\|$, which means separating of spatial and temporal processing.

Thus, before spatial processing begins, the received RE signal $X_i(t,\mu,\nu)$ undergoes optimal temporary processing.

Then from (8) we get:

$$\hat{\mu} \to \max_{\mu} \frac{\left|\mathbf{X}^T \mathbf{R}_n^{-1*} S^*(\mu)\right|^2}{S^T(\mu)\mathbf{R}_n^{-1*} S^*(\mu)} \tag{11}$$

Composition $\mathbf{X}^T \mathbf{R}_n^{-1*}$ in the numerator (6) performs interference compensation, multiplying by $S^T \mathbf{R}_n^{-1*}$ corresponds to the inter-element accumulation, and normalization to the denominator is necessary in order to eliminate the shift of the maximum of the bottom of the DAAR due to interference compensation. Interference compensation means the formation of a dip in the bottom of the DAAR in directions determined solely by amplitude-phase distributions. $S(\nu)$, independent of the intensity and type of point interference. A priori knowledge or estimation of the matrix is needed $\mathbf{R}_n$.

For the case of a single source of interference and the noise intensity of the receiving channels, which is much lower than the interference intensity from the interfering SS, we obtain from (11):

$$\hat{\mu} \to \max_{\mu} \frac{\left|\mathbf{X}^T [\mathbf{I} - \frac{1}{M}\mathbf{S}^*(\nu)\mathbf{S}^T(\nu)]\mathbf{S}^*(\mu)\right|^2}{1 - \left|\frac{\sin\frac{\pi}{2}M(\sin\alpha - \sin\nu)}{M\sin\frac{\pi}{2}(\sin\alpha - \sin\nu)}\right|^2} \tag{12}$$

If the interference matrix $\mathbf{R}_n$ is not known, cancelling of the joint optimal estimation of several angular parameters, taking into account the signal structure, assuming a small delay in signal propagation through the RE and finding the maximum by scanning the space from (12), we obtain a monopulse algorithm for measuring the bearing (13). The algorithm is based on the method of weight summation (WS). Operations are performed to determine the angular coordinate of the SS:

$$\mu_{u} = \max_{\theta} \left| \sum_{i=0}^{M-1} x_i \cdot e^{-j\mu id} \right| \tag{13}$$

and

$$\theta_{u} = \arcsin\left\{ \frac{\mu_{u} \cdot \lambda}{2 \cdot \pi} \right\} \tag{14}$$

where: x_i — estimation of the complex oscillation amplitude at the output of the i-th RE, μ and θ — the variable spatial frequency and target bearing, M is the number of receiving channels.

The MUSIC method

The method is based on dividing the data contained in the estimated CM of the received spatial signal vector into two vector subspaces – the signal subspace and the noise subspace, and on the orthogonality property of the eigenvectors associated with these subspaces [5].

Correlation $M \times M$ - the matrix R in the case of K uncorrelated signals in complex white noise has the form::

$$\mathbf{R} = \left(\sum_{k=1}^{K} S_k \cdot e_k + \sigma_{uu} \cdot e_{uu} \right) \cdot \left(\sum_{k=1}^{K} S_k \cdot e_k + \sigma_{uu} \cdot e_{uu} \right)^H = \\ = \sum_{k=1}^{K} S_k{}^2 e_k e_k{}^H + \sigma_{uu}{}^2 \cdot I \tag{15}$$

where: I – the identity matrix, $e_k = \begin{bmatrix} 1 \\ \exp(j\frac{2\pi \cdot d}{\lambda}\sin(\theta)) \\ \cdots \\ \exp(j(K-1)\frac{2\pi \cdot d}{\lambda}\sin(\theta)) \end{bmatrix}$ – the k-th control (signal) vector.

Let's represent CM **R** as the sum of CM signals and CM noise:

$$R = S + \sigma^2 \cdot I \tag{15}$$

Let's assume that the number SS is $K < M$. In this case, the signal matrix will have the following decomposition according to its own system:

$$S = \sum_{m=1}^{M} \lambda_m v_m v_m{}^H = V \cdot \Lambda \cdot V^H, \tag{16}$$

where: V − matrix of eigenvectors; Λ − diagonal matrix of eigenvalues.

Considering that the rank of the matrix S is equal to K, the diagonal matrix of eigenvectors contains K nonzero elements, which correspond to K of the M eigenvectors matrix. Note that eigenvectors matrix are orthonormal, that is:

$$v_i{}^H v_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{17}$$

The representation of the identity matrix in terms of orthonormal eigenvectors matrix has the form:

$$I = \sum_{m=1}^{M} v_m v_m{}^H \tag{18}$$

Substituting (19) and (17) into (16), we obtain the decomposition of the autocorrelation matrix R by its eigenvalues:

$$R = \sum_{k=1}^{K} \lambda_k v_k v_k{}^H + \sigma^2 \sum_{k=1}^{M} v_k v_k{}^H = \\ = \sum_{k=1}^{K} \left(\lambda_k + \sigma^2\right) \cdot v_k v_k{}^H + \sigma^2 \sum_{k=K+1}^{M} v_k v_k{}^H \tag{19}$$

Hence, on the eigenvectors $v_{K+1}, \ldots, v_M$ the noise subspace of the matrix R is "stretched" on, and on the vectors $v_1, \ldots, v_K$ – the signal subspace.

Since all eigenvectors matrix are orthogonal to each other, the vectors from the subspace of the signal (including the vector s_k) will be orthogonal to (approximately taking into account the effect of noise and exactly in the absence of it) eigenvectors matrix from the subspace of noise:

$$s_k{}^H\left(\sum_{m=K+1}^{M}\alpha_m v_m\right)\approx 0 \qquad (20)$$

Using property (21), we obtain an estimate of the angular spectrum:

$$P_M(\mu)= =\left(e(\mu)^H\cdot\left(\sum_{m=K+1}^{M}v_m v_m{}^H\right)\cdot e(\mu)\right)^{-1} \qquad (21)$$

It is proposed to form an estimated CM as follows:

$R = X^*(n)\cdot X^T(n)$, where $X(n)$ – column vector of the n-th sample of the spatial signal on the M RE of the form (3), $(...)^*$ – complex matrix conjugation.

Then we obtain the CM estimate by averaging the matrices over all sampling intervals:

$$\hat{\mathbf{R}} = \frac{1}{N_t}\sum_{n=1}^{N_t}R_n = \frac{1}{N_t}\sum_{n=1}^{N_t}x(n)\cdot x(n)^H \qquad (22)$$

Based on the obtained CM estimate, the matrix is decomposed according to its own system $\hat{\mathbf{R}}$ and taking into account (22), an estimate of the angular spectrum is constructed:

$$I(\theta)=\frac{1}{\left|\sum_{l=l_0+1}^{M}e(\theta)^H\cdot v_l\right|^2}, \qquad (23)$$

where: l_0 – the number of eigenvector subspaces of the signal;

v_l – l -th eigenvector matrix $\hat{\mathbf{R}}$;

$e(\theta)=\begin{bmatrix}1\\ \exp(j\frac{2\pi\cdot d}{\lambda}\sin(\theta))\\ \cdots\cdots\cdots\cdots\cdots\\ \exp(j(M-1)\frac{2\pi\cdot d}{\lambda}\sin(\theta))\end{bmatrix}$ – a scanning vector of M elements having the structure of a control vector in the expression (15).

Number l_0 does not necessarily coincide with the true number of signals K, then the signal model is inconsistent.

Estimation of directions on the SS: bearings, k — numbers of local maxima of the pseudo-estimation of the angular spectrum (24).

Spatial scanning can be performed once per CM accumulation interval. Scanning can be avoided by switching to the modification of MUSIC – the Root-MUSIC method, the description of which goes beyond the limits of this material. The advantage is the high resolution of closely spaced SS with sufficient signal-to-noise ratios. However, the number of signals in the input process should not exceed the value of eigenvectors corresponding to the subspace of signals: $K \le M - l_0$.

However $l_0 \ge 2$, otherwise the method is inoperable. Reliable information about the number of SS is desirable.

VI. The Results Of The Study Of Bearing Estimation Methods

Since the MUSIC evaluation is nonlinear and the analytical calculation is complex, the analysis is performed on a statistical software model.

The signal model is defined as the vector sum of the complex amplitudes of a random spatially coherent analyzed signal and a spatially coherent interfering signal in the form of a signal delayed by a variable angle, attenuated to a level of 0.4. A background of quasi-white, channel-independent additive noises of the same intensity is created. A linear equidistant antenna array is used.

With a different *M* (NFAR is indicated in the figures) and a different signal-to-noise ratio (SNR is shown in the figures), the maximum normalized scattering of the bearing estimate (Emax) was recorded during the change of the viewing angle to a point interfering SS. The normalization was performed to the width of the radiation pattern of the DAAR at the level of 0.707 in free space. The receiving channels were modeled as identical.

Figure 2 shows the scattering surface of the sun estimation method. Figure 3 shows a similar surface for the MUSIC method. In this case, the parameter taken into account in formula (24) is $l_0 = M - 2$, the actual number of signals in the received oscillation is also 2, the signal model is consistent.

It can be seen from Figures 2 and 3 that the MUSIC bearing estimate begins to outperform the HR estimate with an SNR of more than 20-25 dB.

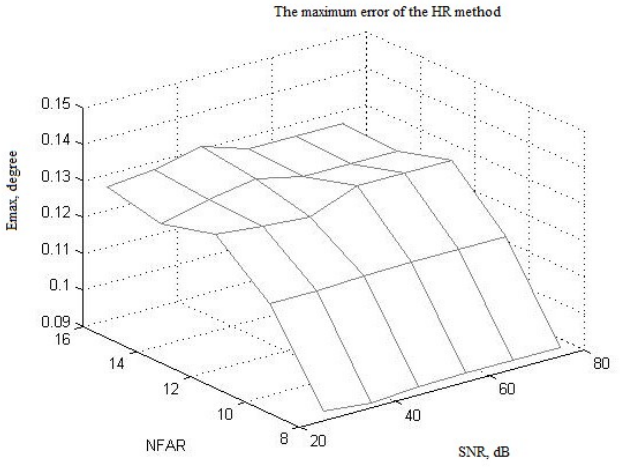


Fig. 2. The maximum error of the HR method

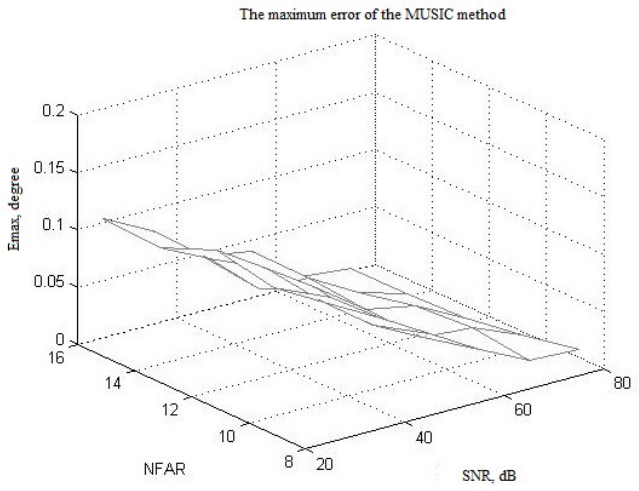


Fig. 3. The maximum error of the MUSIC method

VII. CONCLUSION

This material describes the structure of the DAAR in the variant of bearing estimation, which is possible for use in promising UGV CS and other civil ground vehicles, and shows the advantage in accuracy and resolution of the high-resolution SS assessment of the MUSIC type over the traditional monopulse method in a situation of interfering IC. The advantages of MUSIC are evident with a signal-to-noise ratio of more than 25 dB.

REFERENCES

[1] Filkelstein M.I. Fundamentals of radar. Moscow: Radio and CommunSSations, 1983. 536 p.

[2] Zotov S.A., Makarov E.S., Nechaev Yu.B. Super-resolution methods in radio direction finding problems // ScientifSS journal "Innovative and information processes and technologies in Society and EconomSS" / Edited by Professor Yu.A. Korchagin. Voronezh: RNCE, 2006, No. 3, pp. 12-26.

[3] Marple, Jr., S.L. Digital spectral analysis and its applSSations: Translated from English– Moscow: Mir, 1990, 584 p.

[4] Kulikov E.I., Trifonov A.P. Evaluation of signal parameters against interference, Moscow: Soviet Radio, 1978. 296 p.

[5] Nemov A.V., Kuznetsova E.V. Joint use of high-resolution frequency estimates // Automation and telemechanSS Vol.8, 2003. pp.135-144.

Research on Time Slicing Square Root Kalman Filter under the Limitation of Hardware Resource

Liu Zhuo
School of Microelectronics, Xidian University
Xi'an , China
liuzhuo0403@163.com

Yan Gong-min
College of Automation Northwestern Polytechnical University
Xi'an , China
yangongmin@163.com

Shan Guang-bao
School of Microelectronics, Xidian University
Xi'an , China
gbshan@xidian.edu.cn

Huang Wei
Xi'an North Jierui Optoelectronics Technology Co., Ltd
Xi'an , China
bfjr029@163.com

Luo Yu
Xi'an North Jierui Optoelectronics Technology Co., Ltd
Xi'an , China
bfjr029@163.com

Yang Xiao-kang
College of Automation Northwestern Polytechnical University
Xi'an , China
yangxiaokang495@163.com

As the widespread application of unmanned vehicles such as drones continues to prosper, electro-optical pods and their navigation have become indispensable perception systems. In order to further enhance the integration of the systems while reducing cost and minimizing physical dimensions and weight, optimizing the hardware resource configuration of the inertial navigation system is of critical significance. For conventional inertial navigation system, its implementation is complicated and requires a large amount of computation and storage resources due to the use of Kalman filters in inertial navigation algorithms to obtain accurate state estimation and solve navigation information. This has restricted its practical applications such as cost reduction, lightweight, and further integration with other onboard system processing units for navigation systems, which must be addressed by optimizing algorithms based on application scenarios.

This study focuses on the application of low-cost single precision floating-point arithmetic units in inertial navigation systems. By reducing the time update frequency of the mean square error matrix and dispersing the time update computation of Kalman filter to multiple navigation cycles, the real-time performance of the navigation system and the optimized balance of computing resources are realized. Our proposal realizes the real-time performance of MINS data acquisition, navigation computation and output, while balancing the computation complexity of combined filtering update. It can accomplish a complete time update computation within one Kalman filtering time update cycle. Simulation experiments have verified that our proposal reduces the hardware resource requirements for navigation computation by over 30%. The simulation results suggest that our proposal not only reduces the hardware resource requirements of the navigation system but also achieves comparable accuracy performance as a double precision arithmetic unit on a single precision floating-point arithmetic unit platform. Specifically, the heading and attitude angle errors are less than 0.1 arcminutes, and the positioning error is less than 0.1 meters, which has crucial practical value for integrated application of electro-optical and navigation systems of small unmanned aerial vehicles.

Keywords—Kalman filter; inertial navigation algorithm; mean square error matrix; low-cost arithmetic unit; square root filtering

I. INTRODUCTION

As an efficient recursive mathematical algorithm, Kalman filter/filtering (KF) has been commonly used to dynamically estimate the unknown state of a system from observation data containing noise. The rationale of KF is to minimize the covariance matrix of state estimation, which involves iteratively predicting and updating (correcting) the system model and observations through two steps. As an optimal recursive estimation method, KF can effectively integrate multi-source sensor data (such as IMU, GNSS, DR, magnetometer, etc.) with system dynamics model to realize optimal state estimation in complicated environment with noise interference. It has been widely applied in diverse fields, including but not limited to navigation, control, and signal processing [1].

In conventional airborne INS/GNSS integrated navigation systems, the computation process of navigation is usually based on angular velocity and linear acceleration output by gyroscopes and accelerometers. Specifically, the state parameters such as attitude, velocity, and position are obtained through integration, and the difference between the position, velocity, and attitude of INS and GNSS (differential GNSS) is combined and filtered as an observation. The computation complexity of conventional KF is directly proportional to the cube of the state space dimension, which leads to a sharp increase in computation in high-dimensional problems. For airborne navigation applications, the data update frequency of the navigation system shall reach 100Hz or higher in order to satisfy the requirements of high maneuverability and real-time performance. To realize accurate error estimation, KF with shorter computation cycles can significantly reduce hardware computation load for navigation computers [2].

Under complicated scenarios with combined navigation methods, it is of great necessity to establish high-dimensional error models to solve issues such as lever arm errors and multi-sensor fusion [3,4]. After further expanding the state variables in multi-sensor fusion systems [5-7], the computation complexity of conventional KF in such scenarios will exhibit cubic growth. Under this circumstance, it is urgent to decompose large matrix operation into low-dimensional submatrix iterations through time sliced square

root filtering, and achieve square-level over compression of computation complexity while ensuring numerical stability.

Low-cost navigation system commonly adopts MEMS-IMU as the core for integrated navigation. In order to compensate for the accuracy disadvantage of MEMS-IMU, it is necessary to design IMU arrays and improve navigation accuracy through multi-IMU data fusion [7]. To compensate for device errors and improve positioning accuracy, one of the commonly used approaches is to add installation errors and internal lever arm errors as states in the integrated navigation filter, which inevitably increases the computation complexity of KF [8]. Therefore, low-cost navigation system must reduce the computation complexity of KF. Square root filtering possesses higher numerical stability. When conduct filtering, the square root of the covariance is updated, and the positive definiteness of the matrix is prone to maintain and the stability is stronger even with single precision [10,11]. Therefore, it is more suitable for computing platforms with limited hardware resources.

In response to the above problem, researchers foreign and domestic have proposed various improvement methods. In 2020, Yan *et al.* proposed a sequential filtering method that directly extends the state mean square error matrix for prediction and measurement updates, which effectively reduced the hardware resource demand of KFs [2]. By optimizing the operation sequence of state mean square error matrix, this approach significantly reduced the usage of computation and storage resource, making it particularly applicable for hardware-resource-constrained embedded system. In 2020, Madhukard *et al.* [9] reported a method of unscented KF, which provided a simple nonlinear approach to estimate system state by capturing the most important statistical properties of prior random variables adopting Sigma focus. In addition, related researches have noticed that the major computation resource consumption of airborne INS/GNSS integrated navigation systems comes from computing the state mean square error matrix and error covariance matrix. Therefore, optimizing the above computation process is the key to lowering hardware burden.

This study proposes a time-sliced square root KF method that decomposes the root mean square error prediction process into multiple time slices for sequential computation, reducing the frequency of error covariance computation. The method effectively realizes high-precision navigation computation while possessing excellent real-time performance and computation efficiency. Compared with conventional KF, the proposed method exhibits higher applicability under hardware-resource-constrained conditions, providing a new solution for practical engineering applications.

As a classic filtering technique, KF possesses critical theoretical significance and practical application value for the optimization of hardware resource constraints. The key contributions of this study can be summarized as follows. Firstly, the proposed solution markedly minimizes the computation complexity of updating high-dimensional state matrices in practical navigation applications, thereby reducing the hardware computing power demand of the system on navigation computers. Secondly, according to the practical bandwidth design requirements of the system, the time separation is adjusted to ensure the accuracy and real-time performance of the integrated navigation. Finally, it has significant practical value in the developing trend of low-cost and high integration design of navigation and optoelectronic detection system hardware.

II. Conventional KF Model and Its Computation Complexiaty

The navigation solution based on conventional KF first begins with initialization, i.e. giving the initial values of the carrier state $\hat{\boldsymbol{X}}_0$ and the error covariance $\hat{\boldsymbol{P}}_0$. The second step is to use the IMU acceleration integration to predict position and velocity, and to predict attitude based on angular velocity information. Let the system state equation be

$$\hat{\boldsymbol{X}}_k = \boldsymbol{F}_k \hat{\boldsymbol{X}}_{k.\ k-1} + \boldsymbol{\Gamma}_{k-1} \boldsymbol{W}_{k-1}, \tag{1}$$

where $\hat{\boldsymbol{X}}_k$ denotes the state vector estimation, $\boldsymbol{F}_k$ denotes the system equation, $\boldsymbol{\Gamma}_k$ denotes the system error allocation matrix, $\boldsymbol{W}_{k-1}$ denotes the system noise vector.

Assuming that the measurement is one-dimensional, the observation equation can be expressed as

$$Z_k = \boldsymbol{H}_k \boldsymbol{X}_k + V_k, \tag{2}$$

where Z_k denotes the measurement, $\boldsymbol{H}_k$ denotes the observation matrix, V_k denotes the measurement noise.

By using *priori* estimation to predict the state and mean square error from time k-1 to k, $\hat{\boldsymbol{X}}_k$ and $\boldsymbol{P}_k$ can be expressed as

$$\boldsymbol{K}_k = \boldsymbol{P}_{k/k-1} \boldsymbol{H}_k^T \left(\boldsymbol{H}_k \boldsymbol{P}_{k/k-1} \boldsymbol{H}_k^T + R_k \right)^{-1} \tag{3}$$

$$\hat{\boldsymbol{X}}_k = \hat{\boldsymbol{X}}_{k/k-1} + \boldsymbol{K}_k (Z_k - \boldsymbol{H}_k \hat{\boldsymbol{X}}_{k-1}^k) \tag{4}$$

$$\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k)\, \hat{\boldsymbol{P}}_{k-1}^T, \tag{5}$$

where $\boldsymbol{K}_k$ denotes the filtering gain matrix, R_k denotes the measurement noise variance.

The third step is to introduce other observations (such as GNSS) through *a posteriori* estimation to determine the Kalman gain $\boldsymbol{K}_k$, and adaptively adjust the weights of different observations in the error covariance matrix to update the one-step prediction results.

For the five updated Equations (1) to (5) of the conventional KF, the n^2 multiplication computation required to multiply an $n \times n$ matrix by an $n \times 1$ vector are denoted as S (i.e. $S = n^2$), which performs as the unit for analyzing the computation complexity. Therefore, the computation cost of the state update Equation (1) is $1\mathrm{S}$. The multiplication computation cost in Equation (2) is $2n\mathrm{S}$, and that in Equation (3) is $m\mathrm{S}$, where m denotes the dimension of the measurement vector (m-dimensional vector measurement can be rewritten as a sequential filtering form of m scalar measurements). The computation cost in Equation (4) is negligible, and that in Equation (5) is $m\mathrm{S}$. As a result, estimating the magnitude of multiplication computation for a complete update of conventional KF in one iteration is $2(n+m)\mathrm{S}$.

III. Improved Filtering Method and Its Computation Complexity

A. Time-sliced filtering algorithm

The analysis of the computation complexity of Equations (1) to (5) suggest that during the process of performing

conventional Kalman filtering, updating Equation (2) requires $2nS = 2n^3$ multiplication operations, which underlies the major factor restricting the efficiency of KF operations.

Take a conventional 15-dimension integrated navigation computer as an example. Assuming that the update frequency of IMU is 100Hz, only accomplishing the update of Equation (2) requires the computer to perform 6750 multiplication operations within 10ms. For airborne integrated navigation, the time update frequency of its combined filtering only needs to be no less than 10Hz to satisfy the requirements, for which reason it is unnecessary to strictly follow the update frequency of IMU for KF time update.

To realize 10Hz time update, the time update process is sliced, i.e., a complete time update process is first dispersed into 10 IMU samples and then accomplished. This not only reduces the computation complexity of time updates, but also satisfies the frequency of time updates required by integrated navigation. The transformation process from conventional KF to time-sliced filtering is shown in Fig. 1.

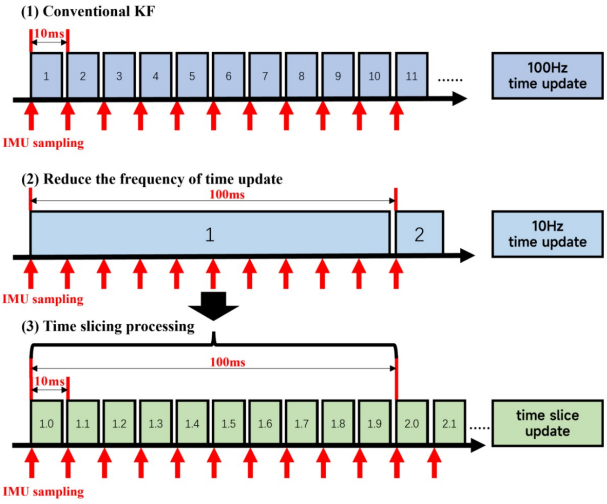


Fig. 1. Comparison between regular time update and sliced time update (10 slices)

As shown in Fig. 1, conventional time update method utilizes the data updated by inertial navigation system for time update after IMU sampling, under which circumstance the KF data processing time is limited to 10ms. If the time update frequency is reduced to 10Hz, computing time update is extended to 100ms. However, the above complete time update process will be frequently interrupted by the IMU sampled inertial navigation update, based upon which the original time update process can be sliced. If the process is fragged into 10 time update slices and the data processing within each slice is accomplished sequentially after IMU sampling, the processing cycle of event update can be extended without affecting the original IMU data processing, thereby solving the problem of large computation complexity in time update.

In order to allocate $2(n+m)S$ calculations reasonably, ensuring time update allocation is required.

B. *Square root filtering algorithm*

Let the square roots of the mean square error matrices $\boldsymbol{P}_{k-1}$, $\boldsymbol{P}_{k/k+1}$, and $\boldsymbol{P}_k$ be $\boldsymbol{\Delta}_{k-1}$, $\boldsymbol{\Delta}_{k/k+1}$, and $\boldsymbol{\Delta}_k$ respectively, we obtain

$$\boldsymbol{P}_{k-1} = \boldsymbol{\Delta}_{k-1}\boldsymbol{\Delta}_{k-1}^T \tag{6}$$
$$\boldsymbol{P}_{k/k+1} = \boldsymbol{\Delta}_{k/k+1}\boldsymbol{\Delta}_{k/k+1}^T \tag{7}$$
$$\boldsymbol{P}_k = \boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^T. \tag{8}$$

The equations for updating the KF time can be rewritten as

$$\begin{aligned}\boldsymbol{\Delta}_{k/k+1}\boldsymbol{\Delta}_{k/k+1}^T &= \boldsymbol{\Phi}_{k/k-1}\boldsymbol{\Delta}_{k-1}\boldsymbol{\Delta}_{k-1}^T\boldsymbol{\Phi}_{k/k-1}^T + \\ &\quad \boldsymbol{\Gamma}_{k-1}\boldsymbol{Q}_{k-1}^{\frac{1}{2}}\left(\boldsymbol{Q}_{k-1}^{\frac{1}{2}}\right)^T\boldsymbol{\Gamma}_{k-1}^T \\ &= \begin{bmatrix}\boldsymbol{\Phi}_{k/k-1}\boldsymbol{\Delta}_{k-1} & \boldsymbol{\Gamma}_{k-1}\boldsymbol{Q}_{k-1}^{\frac{1}{2}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Delta}_{k-1}^T\boldsymbol{\Phi}_{k/k-1}^T \\ \left(\boldsymbol{Q}_{k-1}^{\frac{1}{2}}\right)^T\boldsymbol{\Gamma}_{k-1}^T\end{bmatrix},\end{aligned} \tag{9}$$

where $\boldsymbol{Q}_{k-1}^{\frac{1}{2}}$ denotes the square root of $\boldsymbol{Q}_{k-1}$.

Perform QR decomposition with respect to $\begin{bmatrix}\boldsymbol{\Delta}_{k-1}^T\boldsymbol{\Phi}_{k/k-1}^T \\ \left(\boldsymbol{Q}_{k-1}^{\frac{1}{2}}\right)^T\boldsymbol{\Gamma}_{k-1}^T\end{bmatrix}$,

we obtain $\boldsymbol{\Delta}_{k/k+1}^T$, thus accomplishing time update.

By substituting Equations (6)-(8) into Equation (5), we obtain the covariance measurement update equation, which is expressed as

$$\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^T = \boldsymbol{\Delta}_{k/k-1}\left(\boldsymbol{I} - \rho_k^{-2}\boldsymbol{\Delta}_{k/k-1}^T\boldsymbol{H}_k^T\boldsymbol{H}_k\boldsymbol{\Delta}_{k/k-1}\right)\boldsymbol{\Delta}_{k/k-1}^T, \tag{10}$$

where $\rho_k^2 = \boldsymbol{H}_k\boldsymbol{\Delta}_{k/k-1}\boldsymbol{\Delta}_{k/k-1}^T\boldsymbol{H}_k^T + R_k$.

Let $\boldsymbol{a}_k = \boldsymbol{\Delta}_{k/k-1}^T\boldsymbol{H}_k^T$, the KF covariance update equation is expressed as

$$\begin{cases}\rho_k^2 = \boldsymbol{a}_k^T\boldsymbol{a}_k + R_k \\ \gamma_k = \rho_k\left(\rho_k \pm \sqrt{R_k}\right) \\ \boldsymbol{\Delta}_k = \boldsymbol{\Delta}_{k/k-1}\left(\boldsymbol{I} - \gamma_k^{-1}\boldsymbol{a}_k\boldsymbol{a}_k^T\right) \\ \boldsymbol{K}_k = \boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^T\boldsymbol{H}_k^TR_k^{-1}\end{cases}. \tag{11}$$

The flowchart of square root filtering is shown in Fig. 2.

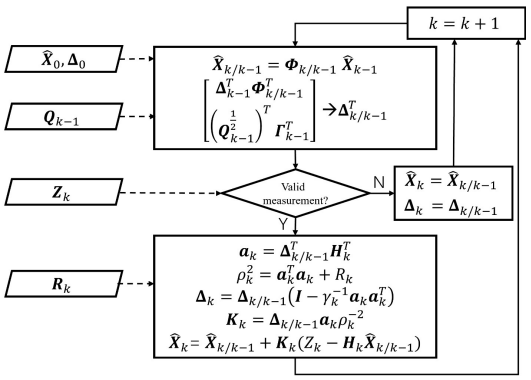


Fig. 2. Flowchart demonstrating the update process of square root filtering

C. *Time-sliced square root KF*

By combining time-sliced filtering algorithm with the square root filtering algorithm while adopting time-sliced update method to compute square root filtering, the time-sliced square root KF can be obtained, which is a kind of KF algorithms applicable under hardware-resource-constrained conditions.

The time update of square root filtering requires performing QR decomposition with respect to $\begin{bmatrix} \boldsymbol{\Delta}_{k-1}^T \boldsymbol{\Phi}_{k/k-1}^T \\ \left(\boldsymbol{Q}_{k-1}^{\frac{1}{2}}\right)^T \boldsymbol{\Gamma}_{k-1}^T \end{bmatrix}$. Let $\boldsymbol{A}_i$ be the i-th column vector of matrix $\boldsymbol{A}$, R_{ij} be the i-th row and j-th column element of matrix $\boldsymbol{R}$. The process of decomposing matrix $\boldsymbol{A}$ adopting the improved Gram-Schmidt method is presented as follows.

$\boldsymbol{R}=\boldsymbol{0}_{n\times n}$
for i=1,2,…,n
  $\boldsymbol{R}_{ii}=\sqrt{\boldsymbol{A}_i^T\boldsymbol{A}_i}$
  $\boldsymbol{A}_i=\boldsymbol{A}_i/R_{ii}$
  for j=i+1,i+2,…,n
    $R_{ij}=\boldsymbol{A}_i^T\boldsymbol{A}_j$
    $\boldsymbol{A}_j=\boldsymbol{A}_j-R_{ij}\,\boldsymbol{A}_i$
  end
end

The time slicing processing of square root filtering can be realized by allocating the QR decomposition process to several IMU sampling periods, to which the schematic diagram demonstrating this process is shown in Fig. 3. Allocate the computation of state update, QR decomposition and measurement update into several IMU sampling periods to reduce the computation complexity of filtering update.

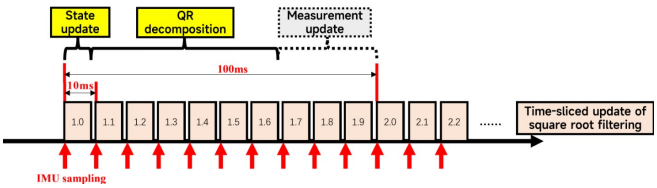


Fig. 3. Update process of square root filtering (10 time slices)

IV. SIMULATION EXPERIMENT

We utilized PSINS toolbox to conduct flight trajectory design, inertial sensor output simulation, inertial navigation computation, and INS integrated navigation simulation [12]. Simulate the flight plane trajectory of a small unmanned aerial vehicle (UAV), as shown in Fig. 4. The UAV accelerated from a standstill at an acceleration of 5m/s^2 to 50m/s and flied at a constant speed heading north. Then, it first turned 90° to the left and flied heading to the west, and then turned 450° to the right and flied heading to the north. Finally, it decelerated until landing and resumed to a standstill. Throughout the entire simulation, the total flight distance of the UAV was about 80km, and the simulation travel time was about 1800s.

Set the frequency of the IMU sampling and INS computation to 100Hz, with the gyroscope random constant drift of 0.1°/h, the angle random walk coefficient of 0.01°/$\sqrt{\text{h}}$, the accelerometer random constant zero bias of 200ug, the velocity random walk coefficient of 10ug/$\sqrt{\text{Hz}}$, the INS initial alignment horizontal misalignment angle of 1′, the azimuth misalignment angle of 30′. The GNSS output frequency is set to 1Hz, and its position measurement noise standard deviation is 5m (1σ).

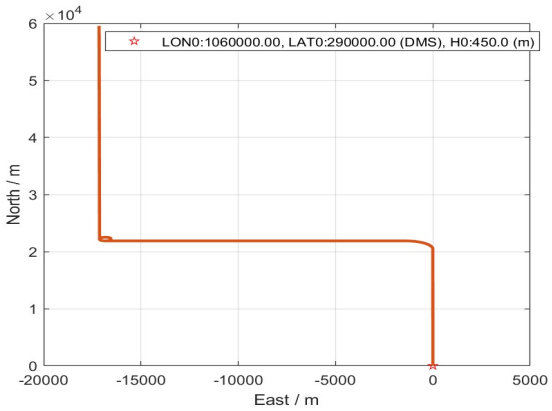


Fig. 4. UAV flight plane trajectory

We conducted comparative analysis of simulation data using both conventional combined navigation and the proposed time-slicing square root combined navigation algorithm. For the former approach, the INS navigation solution and KF data types are both double precision floating-point types. As for the latter approach, the INS navigation adopts a double precision floating-point data type, in which the attitude quaternion and latitude and longitude representation use double precision numbers. The KF adopts a single precision data type, and the number of time slicing is set to 10, suggesting that one KF update is accomplished in 10 INS navigation solutions.

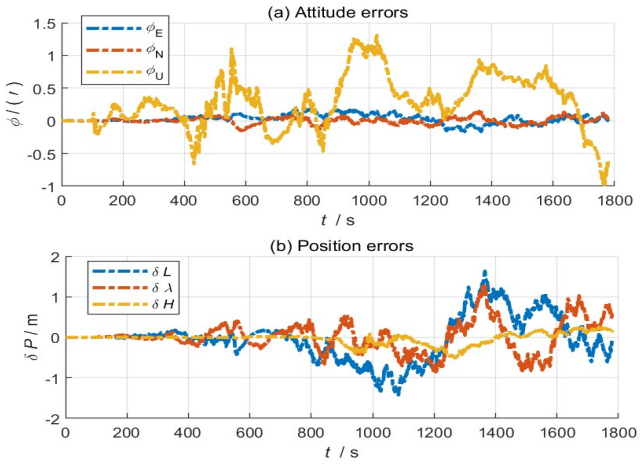


Fig. 5. Navigation error adopting the proposed method

We adopted conventional dual precision data combined navigation results as reference benchmark, to which the misalignment angle and relative positioning error of the proposed filtering method are shown in Fig. 5. It can be observed that the single precision KF error gradually increases at the beginning, and the fluctuation amplitude of the error gradually stabilizes after about 800 seconds. Due to the randomness of numerical solution errors, the state estimation errors exhibit characteristics similar to random walks.

The azimuth misalignment angle error ϕ_U in Fig. 5(a) is $0.49'$ (RMS), while the latitude error δL and longitude error $\delta\lambda$ in Fig. 5(b) are 0.54m and 0.38m (RMS), respectively. The azimuth error is basically within $1'$, and the horizontal position error is mostly within 1m. The single precision time sliced KF algorithm is capable of satisfying the application requirements of medium-precision integrated navigation for UAV.

Further simulation suggests that the number of time slices underlies the major reason attributing to the error of the proposed method, which means the error can be noticeably reduced if the time slice data becomes smaller. Only in higher precision integrated navigation applications can the significant performance of square root filtering be demonstrated.

V. CONCLUSION

The simulation results suggest that adopting the proposed time-sliced square root KF algorithm on the platform with single precision floating-point processor, the accuracy of real-time output navigation information of integrated navigation approximates the post data processing accuracy of the integrated navigation system adopting double precision floating-point arithmetic unit, under which circumstance the attitude and heading errors are less than 1', and the positioning error is less than 1m.

The proposed time-slicing square root Kalman filtering algorithm can effectively improve the computation efficiency of hardware-resource-constrained navigation computing platform, and the navigation results adopting single precision Kalman filtering are close to that of dual precision combined navigation. The proposed method provides a feasible optimization solution for small UAV navigation system with limited resources.

REFERENCES

[1] Yan G, Weng J. Strapdown inertial navigation algorithm and integrated navigation principle(2nd Edition.)[M]. Xi'an: Northwestern Polytechnical University Press, 2023.

[2] Yan G, Y. Deng. Review on practical Kalman filtering techniques in traditional integrated navigation system[J]. Navigation Positioning and Timing. 2020, 7(2): 50-64

[3] W. Jun, J.N. Liu, K. Kou, et al. Online estimation algorithm of multi-lever arm effect for integrated pod SINS/GNSS navigation. [J]. Journal of Chinese Inertial Technology. 2021, 29(2), 184-190

[4] B. Yang, H. Yu, C. Liu, X. Wei, Z. Fan and J. Miao, "A New Compensation Method of Dynamic Lever-Arm Effect Error for Hypersonic Vehicles," in IEEE Access, vol. 10, pp. 59878-59888, 2022.

[5] Z. Wang, J.J. Liu, W. Z. Zhang, et al. Adaptive filtering algorithm for UVA integrated navigation in complex environments[J]. Transducer and Microsystem Technologies. 2025, 44(01): 135-140.

[6] Z. Fang. Research on key technologies of multi-source information fusion for autonomous navigation and simulation for long-endurance UAVs[D]. Nanjing University of Aeronautics and Astronautics,2014.

[7] Lan, Jianping, Kaixuan Wang, Sujing Song, Kunpeng Li, Cheng Liu, Xiaowei He, Yuqing Hou, and Sheng Tang. "Method for measuring non-stationary motion attitude based on MEMS-IMU array data fusion and adaptive filtering." Measurement Science and Technology 35, no. 8 (2024): 086304.

[8] Y. Deng, G. M. Yan and X. K. Yang, "Research on Dis-tributed Attitude System Based on MEMS Inertial Sen-sor Network," 2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, Russia, 2020, pp. 1-4.

[9] P. S. Madhukar and L. B. Prasad. State Estimation using Extended Kalman Filter and Unscented Kalman Filter[C]. International Conference on Emerging Trends in Communication, Control and Computing (ICONC3), Lakshmangarh, India, 2020, pp. 1-4.

[10] Chen, Lin, Wentao Yu, Guoyang Cheng, and Jierui Wang. "State-of-charge estimation of lithium-ion batteries based on fractional-order modeling and adaptive square-root cubature Kalman filter." Energy 271 (2023): 127007.

[11] Chen, Liping, Xiaobo Wu, António M. Lopes, Lisheng Yin, and Penghua Li. "Adaptive state-of-charge estimation of lithium-ion batteries based on square-root unscented Kalman filter." Energy 252 (2022): 123972.

[12] Yan Gongmin. www.psins.org.cn(2025-04-12).

Research of Properties of Characteristics of the Correlation-Extreme Navigation System, Based on Radio Altimeter and Digital Terrain Map

G.V. Smykov

Flight-Navigation and Information Measuring Systems Department Moscow Aviation Institute (National research university)
Moscow, Russia
egor.smikov@yandex.ru

K.K.Veremeenko

Flight-Navigation and Information Measuring Systems Department Moscow Aviation Institute (National research university)
Moscow, Russia
nio3@mai.ru

Abstract – The growing demand for enhanced navigation accuracy through autonomous methods excluding reliance on external information sources while minimizing structural modifications to existing aircraft platforms, necessitates the development of innovative solutions. As an initial step, this study proposes the implementation of a correlation-extreme navigation system integrated into the onboard software, without requiring physical alterations to the airframe. The system architecture incorporates a radio altimeter and a digital terrain map to facilitate precise navigation.

Keywords — correlation-extreme navigation system, inertial navigation systems, radio altimeter, aircraft, bilinear interpolation algorithm

I. INTRODUCTION

Currently, correction of onboard inertial navigation systems (INS) of aircraft vehicles using optical and radiotechnical systems is one of the primary methods for ensuring navigational accuracy during flight in the absence of updates from satellite navigation systems (SNS) [1, 2]. Based on the commonly accepted classification and depending on the onboard equipment of an aircraft vehicle (AV), three types of correlation-extreme navigation systems (CENS) can be considered: Type I - point-based area measurements, Type II - line-based area measurements, Type III - frame-based area measurements [3].

To address the problem of correcting the coordinates computed by the INS in the absence of updates from external sources, the use of a Type I CENS (point-based field measurements) is proposed, as most AVs are equipped with a radar altimeter (RA) and are capable of loading an onboard digital terrain map (DTM). The primary objective is to enable integration of CENS into existing AV with minimal structural modifications.

The proposed system integrates an INS, RA, and an onboard DTM, specifically one that does not require additional preprocessing, unlike in conventional algorithms [4]. This study utilizes a standard DTM in ".sxf" format.

From the algorithmic perspective, a heuristic one-dimensional search-based CENS is proposed for position estimation [3].

For working with terrain elevation grids, two algorithms are considered: Bresenham's line algorithm [5] and the bilinear interpolation algorithm. Bresenham's algorithm is chosen due to its low computational burden compared to bilinear interpolation.

The emphasis of this work lies in utilizing already deployed onboard systems while minimizing any hardware or software changes.

II. PROPOSED APPROACH

As mentioned, a heuristic one-dimensional search-based CENS is to be developed for position estimation, incorporating the INS, RA, and DTM in standard ".sxf" format, along with a processor running a real-time operating system (RTOS). Absolute altitude derived from an external air pressure sensor (e.g., pitot-static system) or equivalent is used as the reference height. The system architecture is schematically illustrated in Figure 1.

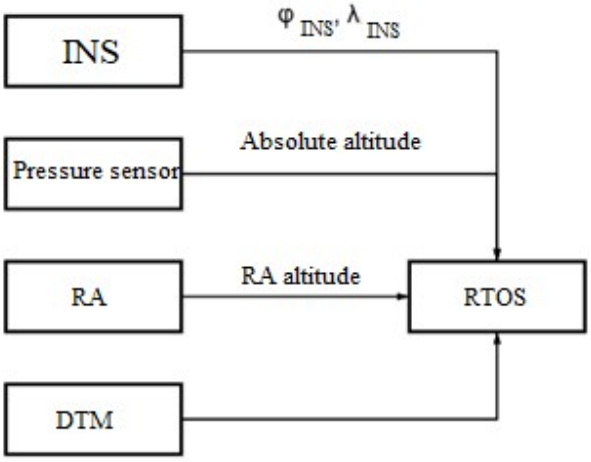


Figure 1 Basic Architecture of the Correlation-Extreme Navigation System

Due to the use of a standard DTM, primarily intended for display purposes, and given that the minimal horizontal resolution of terrain elevation points (TEP) is typically 50 meters, algorithms for grid modification are necessary to reconstruct a valid trajectory.

As a result of the analysis, two algorithms were selected: Bresenham's line algorithm, which determines the discrete set of raster points approximating a straight line between two endpoints [5] and Bilinear interpolation, which estimates intermediate values of a function based on known values at four surrounding points.

Key findings from the comparative analysis:

Bresenham's algorithm is suitable for coarse trajectory estimation, but it introduces significant distortion in the reconstructed trajectory when compared to the true path derived from the DTM. The difference is shown in Figure 2.

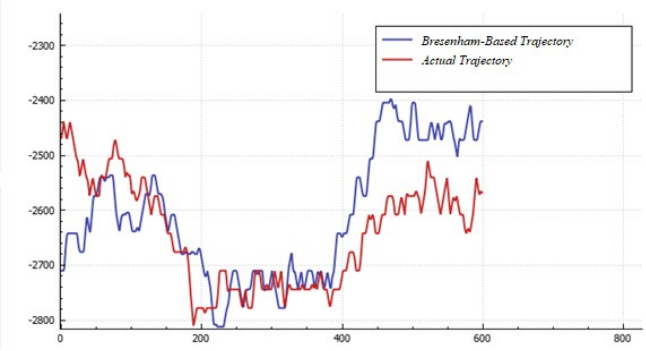


Figure 2 Discrepancy Between Actual and Bresenham-Based Trajectory

Bilinear interpolation enables more accurate trajectory reconstruction, inducing only minor vertical profile errors (point altitude) but imposes significantly higher computational demands. The comparison is presented in Figure 3.

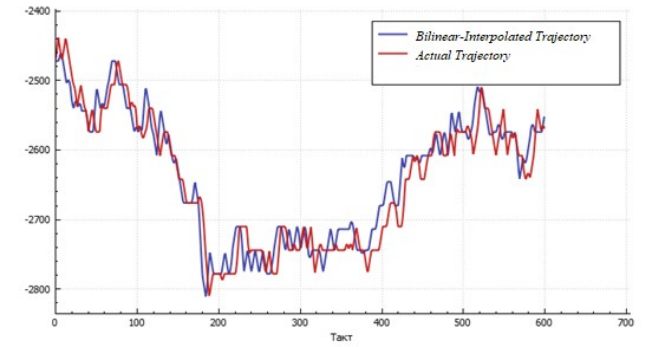


Figure 3 Discrepancy Between Actual and Bilinear-Interpolated Trajectory

III. RESULTS

During algorithm debugging, it was found that the onboard computational complex of the AV has significant performance limitations, making it infeasible to compute the entire CENS optimization process within a single algorithm cycle. Several optimization strategies were proposed and implemented to reduce computational complexity:

1. Task segmentation - The process of identifying the optimal solution is divided into discrete computational cycles. Thus, the final solution becomes available only after a delay needed to process terrain data points. This does not affect the accuracy of coordinate estimation, as positional error evolves over time and exceeds the total duration of the CENS algorithm cycle.

2. Early trajectory pruning - During weight coefficient calculation, any trajectory whose weight exceeds that of the previously found optimal trajectory is discarded early. This eliminates the need to evaluate the entire trajectory from a given map point. No accuracy degradation was observed with this approach.

3. Restricted search domain - Optimization is performed only within a region defined by the current coordinate estimation error (root-mean-square deviation, RMSD). Any trajectory exceeding this boundary is excluded from consideration. Although this significantly reduces the search space, it may result in erroneous localization if the true position lies outside the assumed RMSD radius.

After successful emulator-based testing and hardware-in-the-loop validation, the system was approved for field trials. The results of field testing are summarized in Table 1.

TABLE I. FIELD TEST RESULTS

Calculation No.	Coordinate Type	CENS Estimate	Reference (SNS)	Positioning Error
1	Latitude	43°19' 39.6″	43°20' 33.03″	8249 m
	Longitude	43°31' 49.55″	46°37' 55.97″	
2	Latitude	43°32' 50.42″	43°32' 41.21″	371 m
	Longitude	43°03' 53.94″	43°03' 59.22″	
3	Latitude	42°55' 18.58″	42°55' 23.14″	179 m
	Longitude	44°23' 13.33″	44°23' 05.99″	
4	Latitude	43°27' 30.06″	43°27' 22.65″	270 m
	Longitude	43°11' 02.52″	43°10' 56.2″	
5	Latitude	43°16' 46.2″	43°17' 34.89″	4997 m
	Longitude	43°32' 40.74″	43°28' 34.75″	
6	Latitude	43°12' 07.41″	43°12' 08.61″	87 m
	Longitude	43°35' 42.15″	43°35' 45.66″	
7	Latitude	43°07' 43.09″	43°09' 02.14″	3211 m
	Longitude	43°38' 58.49″	43°37' 26.49″	

Calculations No. 1, 5, and 7 yielded unsatisfactory correction results due to operations over terrain with low information content (i.e., high terrain correlation radius). Under such conditions, the algorithm's performance significantly deteriorates, leading to a considerable increase in total positioning error.

The remaining cases were conducted over moderately to highly informative terrains (i.e., low to medium terrain correlation radius), resulting in minimal positioning errors. Calculations No. 2 and 4 encountered minor radio altitude errors due to multipath effects over terrain with low correlation.

IV. CONCLUSION

Future work involves adapting the proposed algorithms to specific onboard processing systems, taking into account their technical constraints, and further developing the CENS to accommodate operations over terrain with large correlation radius.

REFERENCES

[1] Navigation of Aerial Vehicles in Near-Earth Space. Avgustov L.I., Babichenko A.V., Orekhov M.I., Sukhorukov S.Ya., Shkred V.K., Edited by Prof. G.I. Dzhandzhgava, Doctor of Technical Sciences. Moscow: Nauchtekhizdat, 2015.

[2] Veremeenko K.K., Zheltov S.Yu., Kim N.V., Kozorez D.A., Krasilshchikov M.N., Sebryakov G.G., Sypalo K.I., Chernomorsky A.I. Modern Information Technologies for UAV Navigation and Guidance. Ed. by M.N. Krasilshchikov, G.G. Sebryakov. Moscow: FIZMATLIT, 2009. 556 p. ISBN 978-5-9221-1168-3.

[3] Fundamentals of Navigation Using Geophysical Fields. Beloglazov I.N., Dzhandzhgava G.I., Chigin G.P. Edited by Krasovsky A.A. Moscow: Nauka, Main Editorial Office for Physical and Mathematical Literature, 1985.

[4] Methods for Structuring Onboard Databases of Terrain Relief. Dzhandzhgava G.I., Sazonova T.V., Shelagurova M.S. - Izvestia of Tula State University. Technical Sciences, 2015, Issue 11, Part 2.

[5] Rogers D. "Procedural Elements for Computer Graphics." Moscow: Mir, 1989, pp. 54-63. ISBN 5-03-000476-9.

Autonomous Navigation Method When Flying over Mountainous and Flat Terrain

O.I. Sauta

St. Petersburg State University of Aerospace Instrumentation
St. Petersburg
sauta-oleg@yandex.ru

I.V. Savelieva

JSC Navigator
St. Petersburg
saviris@yandex.ru

E.P. Vinogradova

St. Petersburg State University of Aerospace Instrumentation
St. Petersburg
kate_v@rambler.ru

A.V. Zhmurin

St. Petersburg State University of Aerospace Instrumentation
St. Petersburg
qqwwww2006@gmail.com

Abstract — The paper considers a method for improving the accuracy of correlation-extreme algorithms, which provides a solution to navigation problems of unmanned aerial vehicles in the absence of signals from the global navigation satellite system. The considered method uses digital terrain matrices of the earth's surface, measurements from on-board radio altimeters and coarse speed sensors. An optimal filtering algorithm is proposed to improve the accuracy of navigation using microelectromechanical sensors. The results of studies of the accuracy of correlation-extreme algorithms in simulating flight over mountainous and lowland terrain are presented.

Keywords — correlation-extreme navigation system, digital elevation matrix, microelectromechanical system, Kalman filter.

I. INTRODUCTION

In recent years, unmanned aerial systems (UAS) have become an integral part of aerial aviation. Cargo delivery, monitoring of industrial facilities, rescue operations in emergency situations – the possibilities of using UAS are almost limitless. However, a reliable navigation system is needed for their effective use.

The main source of UAS location data is currently the global Navigation satellite System (GNSS), whose data is used to correct UAS coordinates formed by integrating acceleration and angular velocity signals received from on-board inertial navigation systems (INS) [1]. Most UAS currently use microelectromechanical systems (MEMS) as an INS. Due to the low power of GNSS signals, the latter can be lost in the presence of industrial electromagnetic interference, due to magnetic storms, when the on-board GNSS antennas are shaded by relief elements, or as a result of deliberate radio electronic suppression [2]. The loss of GNSS signals leads to a rapid increase in errors in determining the coordinates of the UAS and, as a result, to the inability to solve navigation problems.

Therefore, determining the location of the UAS using alternative methods that do not require the use of GNSS is an urgent task. One of such alternative methods is the method of correlation-extreme navigation, using measurements of on-board radio altimeters (RA) and digital elevation matrices (DEM) [3].

II. SETTING THE TASK

Suppose that the on-board navigation system of the UAS includes: a computing device in whose memory a digital elevation matrix (DEM) for the intended flight area is stored; an ANN built using coarse MEMS sensors, and a radio altimeter. Then, if the software of the computing device implements the coordinate search method described in [3], which is based on the criterion of the minimum difference between the actual and simulated flight altitude in the area of the most likely location of the UAS (the CENS algorithm), then, as shown in [4], when simulating the flight of the UAS for 120 seconds in the range at altitudes from 1000 m to 2000 m above sea level, the estimate of the error of the UAS coordinates with a probability of 0.95 will be 550 m and 110 m when using coarse or accurate speed sensors, respectively.

The purpose of this report is to develop an algorithm to improve the accuracy of a correlation-extreme terrain navigation algorithm using coarse speed measurements by applying optimal filtering equations.

III. DESCRIPTION OF THE ALGORITHM

Let us use by V_N and V_E the northern and eastern components of the UAS ground speed vector, which are formed by the INS using MEMS sensors. Then, the corresponding errors of these components δV_N and δV_E can be represented by a linearized model described by the expressions [5]:

$$\begin{aligned}\delta V_N &= \frac{g\delta\omega_N}{\omega_S^2}(1-\cos\omega_S t)+\frac{\delta a_N}{\omega_S}\sin\omega_S t + N(0,\sigma_N^2),\\ \delta V_E &= -\frac{g\delta\omega_E}{\omega_S^2}(1-\cos\omega_S t)+\frac{\delta a_E}{\omega_S}\sin\omega_S t + N(0,\sigma_E^2),\end{aligned} \quad (1)$$

where $\delta\omega_N, \delta\omega_E$ - errors caused by MEMS gyroscope drift, g – acceleration due to gravity at sea level, $\delta a_N, \delta a_E$ - accelerometer measurement errors, ω_S - Schuler frequency, $N(0,\sigma_N^2)$, $N(0,\sigma_E^2)$ - random "white" noise with zero mathematical expectation and constant variance, the same for the northern and eastern components of the velocity error, t - time.

If the state vector of the system proposed for constructing an algorithm for solving the problem of estimating coordinates is represented by a model in the state space of the form:

$$X(t) = \left[\varphi(t), \lambda(t), V_N(t), V_E(t)\right]^T, \quad (2)$$

where $\varphi(t), \lambda(t)$ - current geodetic latitude and longitude of the flight, $V_N(t), V_E(t)$ - current Northern and Eastern components of flight speed, t – current time, T – transposition symbol,

and the time derivative of the state vector, omitting the time symbol, is defined as:

$$\dot{X} = \left[\frac{V_N}{R_1}, \frac{V_E}{R_2 \cos\varphi}, \delta\dot{V}_N, \delta\dot{V}_E \right]^T \quad (3)$$

where R_1 - radius of curvature of the meridian, R_2 - radius of curvature of the first vertical,

and the equation of the corresponding discrete system and the equation of the measuring instruments can be written in the form:

$$\begin{aligned} x_i &= \Phi_{i,i-1} x_{i-1} + G_{i,i-1} w_{i-1} + \Gamma_{i,i-1} u_{i-1} \\ z_i &= H_i x_i + v_i, \end{aligned} \quad (4)$$

where $x_i = x(t_i)$ vector of the system state at a point in time t_i, $\Phi_{i,i-1}$ - system transition matrix, $G, \Gamma_{i,i-1}$ - input matrices, w_{i-1} - process noise vector, u_{i-1} - control vector $(-V_{N_i} + \hat{V}_{N_i}, -V_{E_i} + \hat{V}_{E_i})^T$, H_i - matrix of dimensions, z_i - vector of measurements, v_i - measurement noise vector,

then the optimal filtering algorithm for the system (4) will have the form [6, 7]:

$$\begin{aligned} x_{i/i-1} &= \Phi_{i/i-1} x_i + \Gamma u_i, \\ P_{i/i-1} &= \Phi_{i/i-1} P_{i-1} \Phi_{i/i-1}{}^T + G_{i/i-1} Q_{i-1} G_{i/i-1}^T, \\ K_i &= P_{i/i-1} H_i^T (H_i P_{i/i-1} H_i^T + R_i)^{-1}, \\ \hat{x}_i &= x_{i/i-1} + K_i (z_i - H_i x_{i/i-1})_i, \\ P_i &= (E - K_i H_i) P_{i/i-1}. \end{aligned} \quad (5)$$

When constructing the algorithm, it is taken into account that the components of the measurement vector represent the latitude and longitude of the UAS, calculated by finding the minimum difference between the measured and simulated radio altitude around the point with the predicted coordinates $\varphi_i^{CENS}, \lambda_i^{CENS}$ [4].

The correction period should be chosen so that it corresponds to the DEM discreteness and the current flight speed. Choosing too small a measurement period does not lead to an increase in accuracy, but only increases the computational complexity. On the other hand, choosing an excessively large period leads to the loss of data necessary for correction, which reduces the accuracy of navigation determinations [8].

IV. SIMULATION RESULTS

To assess the impact of the proposed optimal filter on the accuracy characteristics of the CENS algorithm considered in [3], a digital simulation of the UAS flight over the Kola Peninsula was performed in an area with flat and mountainous terrain. The studies used DEMs with a horizontal resolution of one and three arc seconds, DEM-1 and DEM-3, respectively. The UAS flight simulation was performed for 30 minutes with a constant ground speed of 30 m/s. Gyroscope drift errors $\delta\omega_N, \delta\omega_E$ were assumed to be equal to 1 deg/hour; accelerometer errors $\delta a_N, \delta a_E$ were assumed to be equal $3 \cdot 10^{-5}$ м/с^2.

The error in determining the location of the UAS was calculated at intervals of 1 s using the formula:

$$\Delta_i = \sqrt{\left(\varphi_i^{CENS} - \varphi_i^{true}\right)^2 + \left(\lambda_i^{CENS} - \lambda_i^{true}\right)^2}, \quad (6)$$

where $\varphi_i^{true}, \lambda_i^{true}$ - the true coordinates of the current location.

Fig. 1 shows the position estimation error when simulating a UAS flight at an altitude of 1500 m and using DEM-1 without applying the filtering algorithm (5).

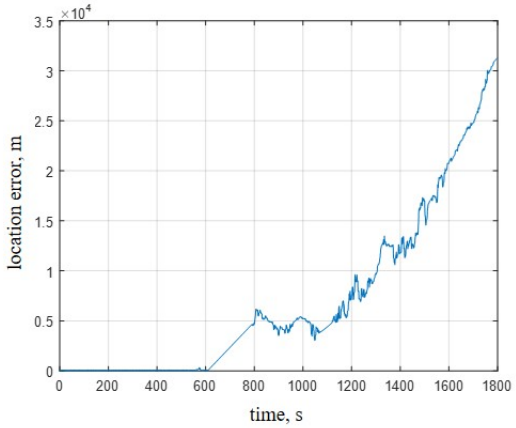


Fig. 1. Location error without filtering

Fig. 2 shows the simulation results for the same initial data, but using the proposed filtering algorithm (5).

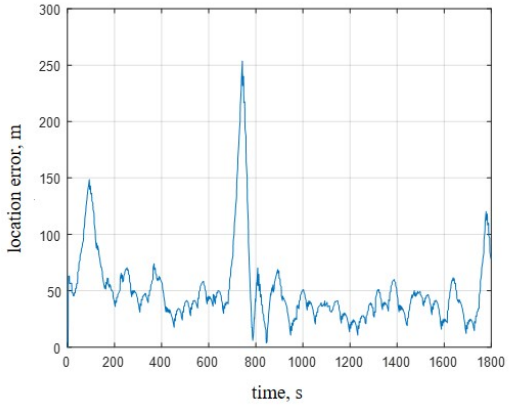


Fig. 2. Location error with filtering

The analysis of the UAS location errors presented in Figs. 1 and 2 shows that the use of the KENS error filtering algorithm allows us to reduce the impact of false hypotheses on the UAS location assessment and increase the location accuracy by approximately two times in certain flight segments. At the same time, for the simulated flight trajectory, the use of the filtering algorithm prevented the divergence of the generated coordinate estimates, and the average location accuracy value was 47.5 m.

Fig. 3 shows the error in estimating the location when simulating a UAS flight at an altitude of 1500 m and using the DEM-1 without applying the filtering algorithm, and Fig. 4 shows the simulation results for the same initial data, but using the proposed optimal filtering algorithm (5).

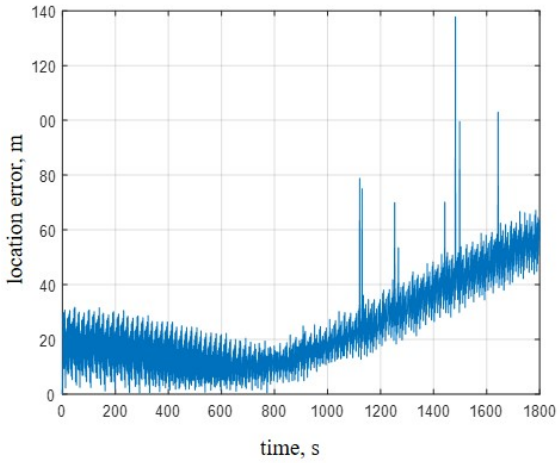


Fig. 3. Location error without filtering

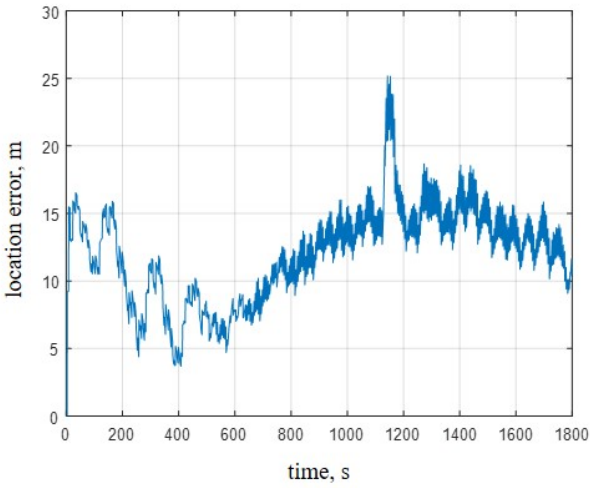


Fig. 4. Location error with filtering

Analysis of the modeling results presented in Figs. 3 and 4 shows that the use of optimal filtering of UAS location estimates generated in CENS reduces their error by an average of three times. At the same time, for the simulated flight trajectory, the average value of the accuracy of determining the location without using the proposed filtering algorithm is 38.6 m, and with the use of filtering it is 13.7 m.

Analysis of Figs. 1 - 4 shows that the use of DEM-1, instead of DEM-3, reduces the error in determining coordinates by an average of 2 times.

V. CONCLUSION

1. A mathematical model of an optimal filter has been developed to improve the accuracy of search correlation-extreme navigation algorithms that use coarse sensors for calculating flight speed, measuring the geometric altitude of an aircraft, and digital matrices of the underlying surface relief.

2. The conducted studies of the influence of the developed optimal filter on the accuracy characteristics of determining the location of an aircraft showed that a two- to three-fold increase in the accuracy of location is ensured when flying over flat and mountainous terrain.

REFERENCES

[1] Avgustov L. I., Babichenko A. V., Orehov M. I., Navigacija letatel'nyh apparatov v okolozemnom prostranstve / Avgustov L. I., Babichenko A.V., Orehov M.I. [i dr.; pod red. Dzhandzhgavy G. I.]. - Moskva : Nauchtehlitizdat, 2015. – 589

[2] Antohina Ju.A., Baburov S.V., Bestugin A.R., Perelomov V.N., Sauta O.I. Razvitie navigacionnyh tehnologij dlja povyshenija bezopasnosti poletov: Monografija / Pod red. Ju.G. Shatrakova. SPb.: SUAI. 2016.

[3] Sauta O.I., Vinogradova E.P Low altitude radio altimeter measurement model for correlation-extreme navigation system // Uspehi sovremennoj radiojelektroniki. 2023. T. 77. № 8. p. 84–90.

[4] Assessment of aircraft coordinates using a terrain model, Sauta O.I., Vinogradova E.P., Zhmurin A.V., Savel'eva I.V. // Metrologicheskoe obespechenie innovacionnyh tehnologij. Sbornik statej V Mezhdunarodnogo foruma SPB, 2025. p. 258-261.

[5] Stepanov O.A. Osnovy teorii ocenivanija s prilozhenijami k zadacham obrabotki navigacionnoj informacii. Chast' 2 Vvedenie v teoriju fil'tracii. – SPb.: Concern CSRI Elektropribor, JSC, 2012 – 417c.

[6] Kuzovkov, N.T., S. V. Karabanov, O. S. Salychev Nepreryvnye i diskretnye sistemy upravlenija i metody identifikacii. — Moscow : Izdatel'stvo MGU, 1978. — 127 p.

[7] Grewal, Mohinder S. Kalman Filtering: Theory and Practice Using MATLAB / Mohinder S. Grewal, Angus P. Andrews. — 4th ed. — Hoboken, NJ : Wiley, 2015. — 432 p

[8] Ogorodnikov, K.O. Position accuracy analysis in terrain-aided navigation systems / K.O. Ogorodnikov // Upravlenie bol'shimi sistemami. Sbornik trudov. - M.: IPU RAN, 2019. №80. - p. 116-134.

.

Adaptive Compensation of Error Estimates in Updatable Strapdown Inertial Navigation Systems Between Observations

A.V. Chernodarov
"NaukaSoft" Research & Production Association, Ltd.,
MAI University, Moscow, Russia,
e-mail: chernod@mail.ru

N.P. Starostin
Ramenskoye Design Company, Ramenskoye,
Moscow Region, Russia,
e-mail: stark201288@gmail.com

***Abstract* — This paper is devoted to the problem of joint estimation and control of parameters of nonlinear dynamic systems under stochastic conditions. These include updated strapdown inertial navigation systems (SINS). The errors of the SINS with respect to the reference phase trajectory, which is determined by external observations, are considered as such parameters. When including the extended Kalman filter in the estimation loop, the control must limit the change in SINS errors to a linear region relative to the reference phase trajectory. The implementation of this approach to damping estimates of errors is associated with the simultaneous solution of nonlinear differential equations of the SINS and linearized error equations. It is proposed to determine the damping coefficients of error estimates based on the method of inverse dynamics problems. Procedures for the formation of such coefficients are synthesized. The results of semi-natural modeling are presented.**

Keywords* — *nonlinear dynamic system, stochastic conditions, joint estimation and control, inverse problem of dynamics, updated strapdown inertial navigation system, extended Kalman filter.

I. INTRODUCTION

The problem of compensation of error estimates in strapdown inertial navigation systems (SINS) between observation sessions is associated with the instability of information external to the SINS. In such a problem, it becomes necessary to matching solutions of nonlinear differential SINS equations and linearized equations of their errors. The matching of solutions is based on controlled compensation of estimates, which ensures stabilization of the SINS parameters relative to the reference phase trajectory determined from observations.

The following schemes for compensating SINS error estimates based on observation results are typical [1]:

• an indicator scheme implementing the write-off of error estimates from the output signals of the SINS;

• a scheme with an updating contour that implements the write-off of error estimates in the dead reckoning equations of the SINS parameters.

When compensating estimates in the indicator mode the necessary conditions for the stable operation of the extended Kalman filter (EKF) [1,2] are not ensured. When implementing the SINS parameter correction contour, estimates are written off from the inertial navigation equations. However, in this case, the error estimates prediction module, which supports the smoothing properties of the EKF, is excluded from the EKF structure. Given the limitations of typical compensation schemes, there is a need to develop procedures for adaptive control of estimates between observations. Such control can be implemented using the damping coefficients of the estimates K_x (see Fig. 1). The purpose of the paper is related to the synthesis of such coefficients based on the inverse dynamics problem.

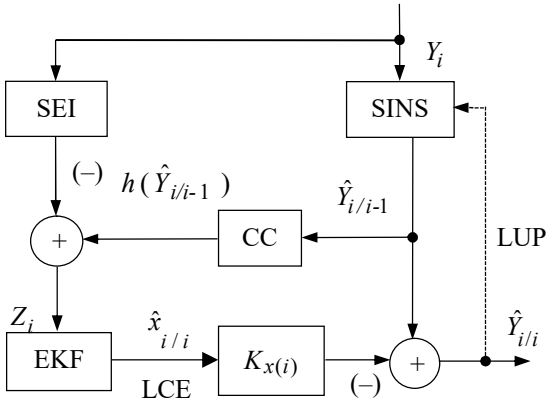


Fig. 1. Block diagram with options for SINS error compensation

Variants of SINS error compensation schemes based on the results of observations are shown in Fig.1, where Y_i is the vector of motion parameters; SEI are sensors of information external to the SINS; CC is a coordinate converter that matches the parameters of the SINS and SEI; $\wedge$ is the symbol for estimate; $\hat{Y}_{i/i-1}$; $\hat{x}_{i/i-1}$ are the predicted values of the estimates of the SINS parameters and errors at time $t = t_i$ after processing the observations Z_{i-1} and solving the equations

$$\dot{\hat{Y}}(t) = F[\hat{Y}(t)]; \quad (1)$$

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t); \quad (2)$$

$\hat{Y}_{i/i}$; $\hat{x}_{i/i}$ are the updated values of estimates after Z_i observations have been processed; LCE is a compensation loop for updated estimates, which is described by the following dynamic equations [1]

$$\hat{x}_{i/i} = \hat{x}_{i/i-1} + K_i(Z_i - \hat{Z}_{i/i-1}); \hat{Y}_{i/i} = \hat{Y}_{i/i-1} - \hat{x}_{i/i}, \quad (3)$$

where $\hat{Z}_{i/i-1} = H_i\hat{x}_{i/i-1}$; $\hat{x}_{i/i-1} = \Phi_i\hat{x}_{i-1/i-1}$; Φ_i is the transition matrix for the error vector, which is determined from the solution of the differential equation

$$\dot{\Phi}(t,t_{i-1}) = A(t)\Phi(t,t_{i-1}) \text{ for } \Phi(t_{i-1},t_{i-1}) = I;$$

I is an identity matrix; LUP is a parameter correction loop by estimates for which the following dynamic equations are valid [1]

$$\hat{x}_{i/i} = K_iZ_i; \hat{Y}_{i/i} = \hat{Y}_{i/i-1} - \hat{x}_{i/i}; \hat{x}_{i/i} := 0, \quad (4)$$

where K_i is the EKF amplification factor.

II. An Inverse Filtering Problem and Identification of the Coefficients of Estimates Damping

Stated classically, the direct and inverse problems of dynamics (IPD) can be interpreted in the following way [3]:

• the direct problem is: for given initial conditions and given input actions, determine the path of system motion in the phase space. For system (1), such a path is described by a matrix exponential;

• the inverse problem is: determine input actions that ensure system motion along the given path in the phase space. If the phase path is given in the form of a matrix exponential, the problem consists in determining the coefficients $A(t)$ in Eq. (1), which provide for the motion along this exponential. For system (1), which has fixed parameters $A(t)$, it is essential to determine additional coefficients K_x, which minimize the deviation of the reckoned phase path $\hat{Y}(t_i)$ from the reference path $Y(t_i)$. To such an interpretation there corresponds the following modification of Eq. (4), i.e.,

$$\hat{Y}_{i/i} = \hat{Y}_{i/i-1} - \hat{\Lambda}_i\hat{x}_{i/i}, \quad (5)$$

where $\hat{\Lambda}_i$ is a matrix for damping coefficients of estimates at the i point in time.

Between observation sessions, equation (1) will have the following numerical solution

$$\hat{Y}_i = F(\hat{Y}_{i-1}, \Delta t_i) - \tilde{\Phi}_i\hat{x}_{i-1}, \quad (6)$$

where $\tilde{\Phi}_i = \tilde{\Phi}(t,t_{i-1}) = e^{-\hat{\Lambda}_i\Delta t_i} \approx I - \hat{\Lambda}_i\Delta t_i$; (7)

$\hat{\Lambda}_i = diag[\lambda_{1(i)} \cdots \lambda_{n(i)}]$; $\Delta t_i = t_i - t_{i-1}$ is the integration step.

The optimality criterion for the damping coefficients was formed based on the η_i residuals between the predicted $\hat{x}_{i/i-1}$ and updated $\hat{x}_{i/i}$ estimates of the SINS errors, i.e.

$$\hat{\bar{\Lambda}}_i = \arg\min_{\bar{\Lambda}_i} 0.5 \sum_{i=i_0}^{i_f} \eta_i^T P_{i/i}^{-1}\eta_i \quad (8)$$

provided that $\bar{\Lambda}_i = \bar{\Lambda}_{i-1} = [\lambda_{1(i-1)} \cdots \lambda_{n(i-1)}]^{\mathrm{T}}$,

where $\eta_i = \tilde{\Phi}_i\hat{x}_{i-1/i-1} - \hat{x}_{i/i} = \hat{x}_{i/i-1} - \hat{x}_{i/i}$; (9)

$P_{i/i} = P_{x(i/i)}$ is the covariance matrix of estimation errors, which is formed by the EKF at the ith instant of time from i observations; i_0, i_f are, respectively, the indices of the initial and final values of observations on the optimization interval $[i_0; i_f]$.

The solution of the problem (8) was obtained using Pontrjagin′s method [4]. The application of this method is based on the formation of the Hamiltonian

$$\mathcal{H} = 0.5\{\|\eta_i\|^2_{P_{i/i}^{-1}} + \|\omega_i\|^2_{N_i^{-1}}\} + \mu_i^{\mathrm{T}}(\bar{\Lambda}_{i-1} + \omega_i), \quad (10)$$

and on the solution of the canonical equations

$$\frac{\partial\mathcal{H}}{\partial\mu_i^{\mathrm{T}}} = \hat{\bar{\Lambda}}_i; \quad \frac{\partial\mathcal{H}}{\partial\bar{\Lambda}_{i-1}} = \hat{\mu}_{i-1}^{\mathrm{T}}; \quad \frac{\partial\mathcal{H}}{\partial\omega_i} = 0 \quad (11)$$

with the boundary conditions

$$\hat{\bar{\Lambda}}_{i_0} = 0; \hat{\bar{\Lambda}}_{i_f} = \text{var}; \hat{\mu}_{i_f} = 0; \hat{\mu}_{i_0} = \text{var}, \quad (12)$$

where μ_i is the vector of Lagrange multipliers; ω_i is the vector of disturbances that affect the compensation coefficients, which is characterized by the covariance matrix $E[\omega(t)\omega^{\mathrm{T}}(t-\tau)] = N(t)\delta(t-\tau)$.

By calculating partial derivatives and performing algebraic transformations, we obtain an expanded notation of the canonical equations

$$\hat{\Lambda}_{i/i-1} = a = \hat{\Lambda}_{i-1/i-1} - N_{i-1}\hat{\mu}_{i-1}; \quad (13)$$

$$\hat{\mu}_i = b = \hat{\mu}_{i-1} + \hat{\bar{x}}_{i-1}^{\mathrm{T}}P_{i/i}^{-1}\eta_i; \quad (14)$$

$$\hat{\omega}_i = -N_{i-1}\hat{\mu}_{i-1}, \quad (15)$$

where $\hat{\bar{x}}_{i-1} = diag\{\hat{x}_{1(i-1)}, ..., \hat{x}_{n(i-1)}\}$.

The estimates of the $\hat{\bar{\Lambda}}_{i/i}$ damping coefficients refined by observations (9) can be found by solving the two-point boundary value problem (TPBP) (12) – (15) using the invariant imbedding method [5]. Following this method, TPBP is included in a more general problem with non-zero boundary conditions

$$\hat{\mu}_{i_f} = c; \hat{\Lambda}_{i_f} = r(c, i_f). \quad (16)$$

For the specified general conditions, the basic equation of invariant imbedding [5] is valid, describing the evolution of the vector $r(c,i_f)$,

$$a[r(c,i_f),c,i_f]-r(c,i_f)= =\left[\frac{\partial r(c,i_f)}{\partial c}+\frac{\partial^2 r(c,i_f)}{\partial c\partial i_f}\Delta i\right]\times \times\{b[r(c,i_f),c,i_f]-c\}+\frac{\partial r(c,i_f)}{\partial i_f}\Delta i. \quad (17)$$

When approximating the solution of equation (17) by a linear function:

$$\bar{\Lambda}_{i_f}(c,i_f): = \Lambda_{i_f}-S_{i_f}c, \quad (18)$$

where $\hat{\Lambda}_{i_f}$ is the TPBP solution at c = 0; S_{i_f} is the covariance matrix of estimation errors, and replacing the final step i_f with the current *i*, the estimation algorithm takes the form

$$\hat{\Lambda}_{i/i}=a[(\hat{\Lambda}_{i/i-1}-S_{i/i-1}c);c;i-1]\big|_{c=0}+ +S_{i/i}b[(\hat{\Lambda}_{i/i-1}-S_{i/i-1}c);c;i-1]\big|_{c=0}; \quad (19)$$

$$S_{i/i}=-\left\{\frac{\partial a[(\hat{\Lambda}_{i/i-1}-S_{i/i-1}c);c;i-1]}{\partial c}\bigg|_{c=0}\right\}\times \times\left\{\frac{\partial b[(\hat{\Lambda}_{i/i-1}-S_{i/i-1}c);c;i-1]}{\partial c}\bigg|_{c=0}\right\}^{-1}. \quad (20)$$

The partial derivatives in equations (19), (20) are determined from relations (13), (14) taking into account (15), (18) and have the form:

$$\frac{\partial a}{\partial c}=-S_{i-1/i-1}-N_{i-1}; \quad (21)$$

$$\frac{\partial b}{\partial c}=I+(\tilde{x}_{i-1}^T P_{i/i}^{-1}\tilde{x}_{i-1})(S_{i-i/i-1}+N_{i-1}), \quad (22)$$

where $S_{i-1/i-1}$ is a matrix of weight coefficients, which was obtained at the (*i*-1)*th* instant of time, having regard to *i* -1 observations (9).

Upon substitution of relations (21), (22) into Eqs. (19), (20) and on performance of algebraic transformations with the use of a lemma on matrix inversion [6], we obtain the following algorithm for damping coefficients estimates.

Prediction: $\hat{\bar{\Lambda}}_{i/i-1}=\hat{\bar{\Lambda}}_{i-1/i-1}$; (23)

$$S_{i/i-1}=S_{i-1/i-1}+N_i. \quad (24)$$

Updating: $\eta_i=\tilde{\Phi}_i\hat{x}_{i-1/i-1}-\hat{x}_{i/i}=\hat{x}_{i/i-1}-\hat{x}_{i/i}$; (25)

$$V_i=\hat{\tilde{x}}_{i/i-1}S_{i/i-1}\hat{\tilde{x}}_{i/i-1}^{\mathrm{T}}+P_{i/i}; \quad (26)$$

$$S_{i/i}=S_{i/i-1}-S_{i/i-1}\hat{\tilde{x}}_{i/i-1}^{\mathrm{T}}V_i^{-1}\hat{\tilde{x}}_{i/i-1}S_{i/i-1}; \quad (27)$$

$$\hat{\bar{\Lambda}}_{i/i}=\hat{\bar{\Lambda}}_{i/i-1}+S_{i/i}\hat{\tilde{x}}_{i/i-1}^{T}P_{i/i}^{-1}\eta_i, \quad (28)$$

where $\hat{\tilde{x}}_{i/i-1}=diag\{\hat{x}_{1(i/i-1)},...,\hat{x}_{n(i/i-1)}\}$; $\hat{\bar{\Lambda}}_{i/i-1}$; $\hat{\bar{\Lambda}}_{i-1/i-1}$ are respectively the predicted and updated estimates of the vector of coefficients; $S_{i/i-1}$ and $S_{i/i}$ are the covariance matrices for the specified estimates.

In the implementation of the algorithm (23) – (28), the need arises for inversion of the matrices V_i and $P_{i/i}$ in Eqs. (27) and (28). These problems can be solved by assuming that the elements of vector $\bar{\Lambda}_i$ are mutually independent, and also considering only the diagonal elements of the matrix $P_{i/i}$. Such elements of the $P_{i/i}$ matrix are variances of errors in estimating the parameters of vector x_i. We will also assume that the dimensions of the vectors $\bar{\Lambda}_i$ and x_i coincide and are equal to $n\times 1$. In this case, a sequential implementation of algorithm (23) – (28) is possible.

Prediction: (23) – (24).

Updating: $\eta_j=\hat{x}_{j(i/i-1)}-\hat{x}_{j(i/i)}$; (31)

$$V_j=\hat{\tilde{x}}_{j(i)}S_{j-1}\hat{\tilde{x}}_{j(i)}^{\mathrm{T}}+P_{jj(i/i)}; \quad (32)$$

$$\tilde{K}_j=S_{j-1}\hat{\tilde{x}}_{j(i)}^{\mathrm{T}}/V_j; \quad (33)$$

$$S_j=(\tilde{K}_j\hat{\tilde{x}}_{j(i)}-I)S_{j-1}(\tilde{K}_j\hat{\tilde{x}}_{j(i)}-I)^{\mathrm{T}}-\tilde{K}_jP_{jj(i)}\tilde{K}_j^{\mathrm{T}}; \quad (34)$$

$$\hat{\tilde{\Lambda}}_j=\hat{\tilde{\Lambda}}_{j-1}+K_j\eta_j;\ j=\overline{1,n}; \quad (35)$$

$$\hat{\tilde{\Lambda}}_{i/i}=\hat{\tilde{\Lambda}}_n;\ S_{i/i}=S_n, \quad (36)$$

where $\hat{\tilde{x}}_i=[\hat{\tilde{x}}_{1(i)}^{\mathrm{T}}\cdots\hat{\tilde{x}}_{j(i)}^{\mathrm{T}}\cdots\hat{\tilde{x}}_{n(i)}^{\mathrm{T}}]^{\mathrm{T}}$; $P_{jj(i/i)}$ is the *j*th diagonal element of the $P_{i/i}$ matrix; $\hat{x}_{j(i/i-1)}$ and $\hat{x}_{j(i/i)}$ are the *j*th elements of vectors $\hat{x}_{i/i-1}$ and $\hat{x}_{i/i}$.

The initial values of the elements of the covariance matrix S_0 and the vector of coefficients $\hat{\tilde{\Lambda}}_0$ are formed taking into account the dynamic properties of the SINS, as well as the intervals and noise of observations.

III. Analysis of the Results of Studies

The presented technology for compensating estimates of SINS errors was tested in a semi-naturalistic experiment. The inertial-satellite navigation system SINS-500NS [7] based on fiber-optic gyros (FOG) developed by the Research & Production Association "NaukaSoft" (Moscow) was considered as an object of studies. In the SINS-500NS system, the components of the relative velocity vector, as

well as the orientation and navigation quaternions, are reckoning by solving 11th-order nonlinear differential equations of the form (1). The integration step of the SINS equations is 2.5 ms. The step of forming observations is ≥ 1 Hz. The vector of SINS errors is comprised of 18 parameters, namely: errors in the reckoning of components of the vector of relative velocity; errors in the reckoning of elements of the navigation quaternion and an attitude one; angular drifts of FOGs; biases of accelerometers signals and also an error in the reckoning of elevation with respect to the Earth ellipsoid. The need to control error estimates is associated with the uneven formation of observations due to the instability of satellite information [8]. The SINS error vector was estimated by processing inertial-satellite observations of the relative velocity $\bar{V} = [V_E V_N V_H]^{\mathrm{T}}$, geodetic latitude φ, longitude λ and altitude h:

$$Z_{C(i)} = [\varphi_i \lambda_i h_i]^{\mathrm{T}}_{\mathrm{SINS}} - [\varphi_i \lambda_i h_i]^{\mathrm{T}}_{\mathrm{GNSS}}; \qquad (35)$$

$$Z_{V(i)} = [V_E V_N V_H]^{\mathrm{T}}_{(i)\mathrm{SINS}} - [V_E V_N V_H]^{\mathrm{T}}_{(i)\mathrm{GNSS}}, \quad (36)$$

Some results of experiments based on the mobile laboratory are presented in the following Figs.: in Fig. 2, the circular positional error ΔS during the implementation of the parameter correction contour (4); in Fig. 3, the error ΔS during the implementation of the adaptive correction contour with the identification of the damping coefficients of estimates, where

$$\Delta S = \sqrt{\delta_\varphi^2 + \delta_\lambda^2}; \quad \delta_\lambda = (\lambda_{\mathrm{SINS}} - \lambda_{\mathrm{GNSS}}) R \cos\varphi_{\mathrm{GNSS}};$$

$\delta_\varphi = (\varphi_{\mathrm{SINS}} - \varphi_{\mathrm{GNSS}}) R$; R is the radius of curvature of the earth's ellipsoid at the SINS location.

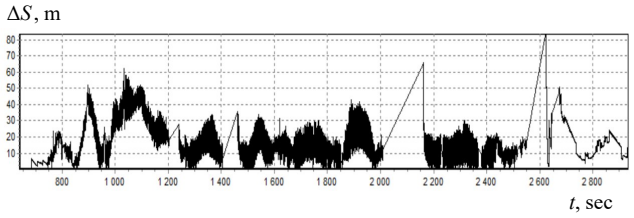


Fig. 2. Circular position error when the loop (4) for the updating of SINS parameters is realized

A comparison of the results shows that, with large time intervals between observations, the implementation of the SINS parameter correction contour (4) can lead to their significant divergence relative to the reference trajectory formed by the global navigation satellite system (GNSS). Under the same conditions, the use of controlled compensation of SINS error estimates makes it possible to increase the reliability of determining inertial parameters.

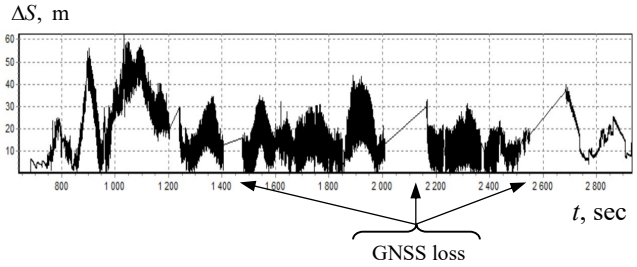


Fig. 3. Circular position error when the loop for adaptive updating of SINS parameters is realized

IV. Conclusions

The paper formulates the problem of adaptive exponential compensation of SINS error estimates between observation sessions. The proposed solution is based on the application of the method of inverse problems of dynamics. The implementation of the method is based on the identification of coefficients for estimates monitoring between observation sessions. Identification is performed based on processing the residuals between the predicted and updated by observation estimates of the error vector. The obtained research results confirm the expediency of controlled compensation of error estimates in the SINS equations.

References

[1] Maybeck, P.S., Stochastic Models, Estimation and Control. N.Y., Academic Press, 1982, vol. 2.

[2] Stepanov, O.A., Osnovyi teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoy informatsii. Ch.1. Vvedenie v teoriyu otsenivaniya. (Fundamentals of Estimation Theory with Applications to Navigation Information Processing Problems. Introduction to the Theory of Estimation. P.1), St. Petersburg, CSRI Electropribor, 2010.

[3] Chernodarov, A.V., Identification and an Inverse Filtering Problem // 18th IFAC Symposium on System Identification (SYSID 2018), IFAC-PapersOnLine, 2018, 51(15), pp. 66–71.

[4] Speedy, C.B., Brown, R.F., Goodwin, G.C., Control Theory: Identification and Optimal Control. Edinburg., Oliver and Boyd, 1970.

[5] Sage, A.P, Melsa, J.L., System Identification. N.Y., Academic Press, 1971.

[6] Sage, A.P., White, C.C., Optimum Systems Control. New Jersey, Prentice-Hall, 1977.

[7] Chernodarov, A.V., Gorshkov, P.S., Patrikeev, A.P., and Starostin, N.P., Investigation of the Emergency Mode of the SINS-500NS Strapdown Inertial-Satellite Navigation System Based on Flight Data, 30st Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2023, St. Petersburg, CSRI Electropribor, 2023. IEEE, 23514304.

[8] Schmidt, G.T., GPS Based Navigation Systems in Difficult Environments, Gyroscopy and Navigation, 2019, vol. 10, no 2, pp. 41 - 53.

Search Algorithm for Localization of a Ground Vehicle using Geometric Primitives for Analyzing Trajectory Fields Obtained by Satellite Navigation

R.R. Bikmaev

Navigation and geodetic systems dept. ANO «Institute of Engineering Physics»
Serpukhov, Russian Federation
e-mail: naviserp5@iifmail.ru

***Abstract* — This paper studies an algorithm for searching the location of a vehicle as a point of intersection of its trajectory with a graph, the vertices of which are global positioning points and the edges are road segments represented as geometric primitives (lines and curves).**

Keywords* — *map localization, map matching, Gaussian Mixture Model

I. Introduction

The solution to a problem of localization using the map allows to compensate navigational error with the help of trajectory fields matching with data mapped on the map. But this requires the use of additional sensors (cameras and lidars) to implement the survey-comparison method of navigation. In general, they use complex algorithms to search for a georeferenced local location, but the error in matching to a map (reference) is usually independent of the geodetic coordinates of the localization and requires the integration of additional data sources into the system. The paper proposes to use a single source of data for map-based localization - vehicle trajectory data.

II. Approaches

The approach [1] solves the problem of correcting inertial system by estimating the mismatch parameters between the digital road map and on-board navigation system of the vehicle. The algorithm for correction point searching aims to minimize the lateral deviations of the driving trajectory with segments whose starts and ends are road axial line points.

In papers [2],[3], a method is implemented that uses a Gaussian mixture model (of the unsupervised learning method) for approximating inference in determining the intersection point of the road graph and the vehicle trajectory obtained by visual odometry. In the proposed search algorithm, in order to reduce computational complexity, solving rules are proposed to simplify the representation of the posterior distribution to ensure the convergence of the Gaussian mixture model.

The algorithm proposed in this paper relies on the second approach, but graphs the road not in the plane of a two-dimensional map, but based on the global positioning track, in which the height defined too. Also in the algorithm, a symmetric matching measure, the Jensen-Shannon divergence [4], is applied when estimating distributions in the Gaussian mixture model by Kulbak-Leibler divergence.

III. The Problem Statement

Assume a map m exists in which roads are presented as a set of consecutive points p', unevenly spaced on the axis line of the path. For certain routes, tracks consisting of time-ordered points p of global positioning by a satellite navigation system are obtained. The measured data represent tuples of latitude, longitude, altitude (B, L, H), and are affected by some noise.

A vehicle is travelling along a certain route, equipped with a inertial system that measures the distance travelled and orientation with accumulated error.

It is required to compose an A search algorithm that determines the intersection points of the vehicle trajectory with the tracks mapped on the map.

IV. Solution

The map m can be represented as a connected graph G whose vertices are the points $p(B, L, H)$ closest to the points p' on the axis line and whose edges are the road segments s. Therefore, the map is a road graph $G(p,s)$.

The vertices of the graph $p(B, L, H)$ are determined by the iterative projection method described in [5], which allows multiple tracks to be combined into a single route track. Road segments s (edges of graph G) are described by a line or curve as in Figure 1, Figure 2. and Figure 3.

Orientation and location of a vehicle on the map is determined by the road segment s on which it is moving. Using parameters of linear and arc segments, the vehicle location can be expressed through the distance from the point $p_1 - D$ and angular shift - ψ. Based on this, the navigation course, location in the linear and arc segment is determined by the following formulae:

$$\hat{\psi} = \psi + \alpha + k \cdot D' \tag{1}$$

$$\hat{p} = \frac{L - D'}{L} \cdot p_1 + \frac{D'}{L} \cdot p_2 \tag{2}$$

$$\hat{p} = O + R \cdot \mathbf{d}\left(\frac{L - D'}{L} \cdot \psi_0 + \frac{D'}{L} \cdot \psi\right) \tag{3}$$

In formulas (1-3) $D` = D * \cos(\beta)$, R is the radius of the circle with centre O, φ_0 is the starting angle of the horizontal grid line and $d(\phi) = [\cos\phi, \sin\phi]^T$.

Based on the representation of the map as G(p,s), where edges are described by formulas (1-3), it is possible to define the state of the vehicle at time t as $x_t = (s_t, D`_t, D`_{t-1}, \psi`_t, \psi`_{t-1})$,

where [D`$_t$ , D`$_{t-1}$, ψ`$_t$, ψ`$_{t-1}$]^T = S_t, which accounts for the distance and angular shift at the previous and current time instance.

Let's denote $\mathbf{y}_t$ as the measurements of linear and angular shift by the inertial system determined at time t. Then the odometry observations are modeled as:

$$p(\mathbf{y}_t|\mathbf{x}_t) = \mathcal{N}(\mathbf{y}_t|\mathbf{M}_{s_t}\mathbf{S}_t, \sum_{s_t}^{y}) \quad (4)$$

where M_s = [$\mathbf{m}_D$, $\mathbf{m}_\psi$]^T, $\mathbf{m}_D$ = (1, -1, 0, 0)^T and $\mathbf{m}_\psi$ = (r_s, -r_s, 1, -1)^T.

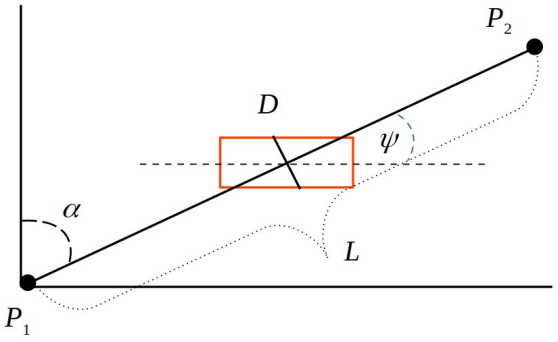


Fig. 1. A linear segment of a road that has a direction, a start and end vertex, a length and the distance travelled by a vehicle.

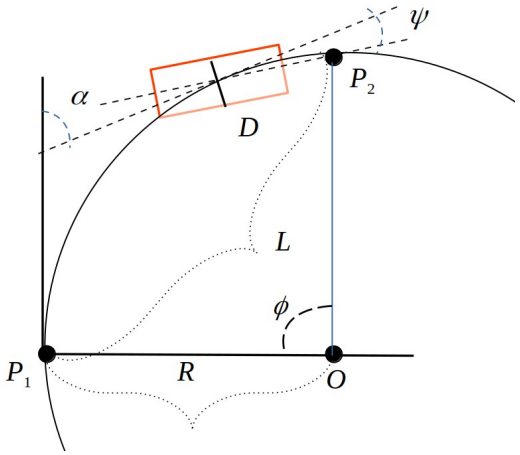


Fig. 2. Arc segment of the road, with curvature *k=1/R*.

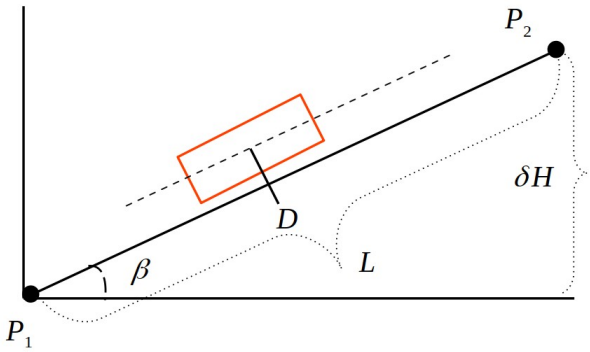


Fig. 3. Road segment in which the height changes.

The state transition model with segment transition probability p(s_t| x_{t-1}) and state transition distribution p (S_t | s_t, x_{t-1}) can be represented as:

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = p(s_t|\mathbf{x}_{t-1})p(\mathbf{S}_t|s_t, \mathbf{x}_{t-1}) \quad (5)$$

The state transition distribution is the result of a linear transformation distorted by Gaussian noise:

$$p(\mathbf{S}_t|s_t, x_{t-1}) = \mathcal{N}(\mathbf{S}_t|\mathbf{A}_{s_t,s_{t-1}}\mathbf{S}_{t-1} + \mathbf{b}_{s_t,s_{t-1}}, \Sigma^{\mathbf{S}}_{s_t}) \quad (6)$$

The segment transition probability p (s_t |x_{t-1}) determines the probability of transitioning to segment s_t given a given previous state x_{t-1}. To determine the segment transition probability, we use the Gaussian state transition model in (6) to obtain

$$p(s_t|\mathbf{x}_{t-1}) \propto \xi_{s_t,s_{t-1}} \int_{L_{s_{t-1}}}^{L_{s_{t-1}}+L_{s_t}} \mathcal{N}(x|\mathbf{a}_D^T\mathbf{S}_{t-1}, \mathbf{a}_D^T\Sigma^{\mathbf{S}}_{s_{t-1}}\mathbf{a}_D)dx \quad (7)$$

where a_D is the first row of the matrix $\mathbf{A}_{st,\ s\{t-1\}}$, which represents predicted position D_t, $\xi_{st,st-1}$ - is the a priori transition probability on s_t from s_{t-1}, and **N**(s) is the set of segments that *s* connects to.

Let's define the inference from the distribution p(x_t | $y_{1:t}$), given the state transition model. For this purpose, we divide the algorithm into an uncertainty propagation (prediction) stage (5) and an uncertainty updating (correction) stage using the observation model (4).

To represent continuous distribution p($\mathbf{S}_t$ | s_t, $y_{1:t}$) the Gaussian mixture model is used.

$$p(\mathbf{S}_t|s_t,\mathbf{y}_{1:t}) = \sum_{i=1}^{N_{s_t}} \pi_{s_t}^{(i)}\mathcal{N}(\mathbf{S}_t|\mu_{s_t}^{(i)}, \Sigma_{s_t}^{(i)}) \quad (8)$$

where N_{st} is the number of mixture components associated with s_t, and M_s^t are the mixture parameters for s_t.

Assuming that the independent observations are defined by states and first-order Markovian state transitions, using the mixture model (8) and state transition dynamics (5), the Bayesian filtering distribution recursively is as follows:

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \sum_{u_t} \frac{P_{s_{t-1}}}{Z_t} p(\mathbf{y}_t|\mathbf{x}_t) \sum_{i=1}^{N_{u_{t-1}}} \pi^{(i)} \int p(s_t|s_{t-1}, \mathbf{s}_{t-1})\mathcal{N}(\mathbf{S}_t|\mathbf{A}\mathbf{s}_{t-1} + \mathbf{b}, \Sigma^s) \times \mathcal{N}(\mathbf{s}_{t-1}|\mu^{(i)}, \Sigma^{(i)})d\mathbf{s}_{t-1} \quad (9)$$

The integral in (9) due to the nonlinear form p (s_t | s_{t-1}, S_{t-1}) cannot be solved analytically. Let's introduce a crucial rule that p(s_t | s_{t-1}, S_{t-1}) is a constant under the condition ||d/(d μ) g (μ, Σ)|| < 10^{-8}, then the integral is approximated analytically as

$$c_{pred}\mathcal{N}(S_t|\mathbf{A}_\mu + \mathbf{b}, \Sigma^S + \mathbf{A}\Sigma\mathbf{A}^T) \quad (10)$$

where c_{pred} = p (s_t | s_{t-1}, μ). When the condition is not satisfied, a Monte Carlo approximation is used, computing from M = 400 samples, sampling $S_{t-1}^{(j)}$ = N(μ, Σ) for j = 1,.., M, and then the integral is approximated by a single component c_{pred} N (S_t | μ_{pred}, Σ_{pred}, где c_{pred} = 1/M $\Sigma_{j=1}^{M}$ $\omega^{(j)}$.

Since the distribution of the observation model p(y_t | x_t) is Gaussian and linear in S_t, this corresponds to the multiplication of two Gaussian functions and can be solved in closed-form. Then the resulting update is calculated as c_{upd} N (S_t | μ_{upd}, Σ_{upd}):

$$c_{upd} = \frac{\omega c_{pred}|\Sigma_{upd}|^{0.5}}{|\Sigma_{upd}|^{0.5}|\Sigma_{s_t}^{\mathbf{y}}|^{0.5}} e^{-\frac{1}{2}||\mathbf{y}_t - \mathbf{M}_{s_t}\mu_{pred}||^2_{\Sigma'_{upd}}} \quad (11)$$

Performing the prediction and correction step for each component and each pair of linked nodes creates a set of mixture model components for each s whose weights are proportional to P_s^t. After normalizing the mixtures for each segment, the probability of finding a vehicle in that segment can be obtained, P_s^t.

Implementing an inference model with prediction and correction step is computationally complex because the number of mixture components grows exponentially with time and the modes are duplicated in two rounds of enumerating s segments. So a Gaussian mixture model simplification procedure is run that removes components at transitions of neighboring segments and updates others while keeping the Kullback-Leibler divergence below a given threshold value.

V. The Search Algorithm

The general algorithm for searching vehicle location as intersection point of trajectory with the graph, considering the map representation model described above and inference model with prediction and correction step, as well as the simplification procedures of the Gaussian mixture model consists of two algorithms:

1) A map representation algorithm in which track obtained by global positioning is transformed into a road graph G(p,s).

Algorithm 1 Map Representation

Input: a map set axial points $id' \mapsto p'_i$
Input: a map set of GPS locations p_j from track id
map Points $\leftarrow \emptyset$
for all p' **do**
  for all p **do**
    if $id == id'$ **then**
      Calculate distance p'_i and p_j
      if distance $<$ threshold **then**
        Points $\leftarrow p$
      end if
    end if
  end for
end for
for all Points **do**
  Calculate parameters $l,r,\alpha,p_1,p_2,\beta,\phi$ Map Graph G
  Compute transition information for each node
  Compute Nearby Roads
end for
Return: graph G(p,s)

2) An algorithm for searching the intersection of a graph representation with a vehicle trajectory as an inference from the distribution (10).

VI. The Experiment

To evaluate the location search algorithm, an experiment was conducted using the digital twin of the survey-comparison method, in which the on-board navigation system in the form of inertial system was corrected by the lidar by recognizing road intersections. Place recognition was based on the histogram method of machine vision by determining the distances of the earth mover's distance between the reference and the projection on the image of the measured lidar scan (inverse matching measure). In the experiment, the search algorithm operated independently of the implementation of the survey-comparison method, using relative coordinates, and adjusted the trajectory taking into account the calculated shifts in coordinates at the points where the inertial system trajectory and the road graph were assumed to intersect.

In the experiment, four series of five drives along a 25.6 km route were set. For each series of drives matching measure was varied from 0.15 to 0.09.

TABLE I. Mean position error by coordinates $X|Y$ in metres

Method	Matching measure in the series			
	0.15	***0.12***	***0.10***	***0.09***
Survey-comparison method	-3.272 3.185	-3.238 3.591	0.072 2.014	-6.147 16.446
Searching algorithm	+2.344 -2.176	-2.789 -2.876	-2.954 +1.567	+2.297 -2.152

Algorithm 2 Search for intersection

Input: Posterior at $t-1$, $\{P_s^{t-1}, \mathcal{M}_s^{t-1}\}$
Input: Observation at t, $\mathbf{y}_t$
Initialize mixtures, $\mathcal{M}_s^t \leftarrow \emptyset$, for all s
for all segments s_{t-1} **do**
  for all segments s_t reachable from s_{t-1} **do**
    $\mathcal{M}' \leftarrow \emptyset$
    for all $(w,\mu,\Sigma \in \mathcal{M}_{s_{t-1}}^{t-1})$ **do**
      Compute $c_{pred}\mathcal{N}(\mu_{pred},\Sigma_{pred})$
      $c_{pred} \leftarrow \frac{1}{M}\sum_{j=1}^{M} w^{(j)}$
      $\mu_{pred} \leftarrow \frac{1}{Mc_{pred}}\sum_{j=1}^{M} w^{(j)}\mathbf{S}_t^{(j)}$
      $\Sigma_{pred} \leftarrow \Sigma^s + \frac{1}{Mc_{pred}}\sum_{j=1}^{M} w^{(j)}(\mathbf{S}_t^{(j)}-\mu_{pred})(\mathbf{S}_t^{(j)}-\mu_{pred})^T$
      Compute $c_{upd}\mathcal{N}(\mu_{upd},\Sigma_{upd})$
      $\Sigma_{upd} \leftarrow (\mathbf{M}_{u_t}^T\Sigma_{u_t}^{\mathbf{y}-1}\mathbf{M}_{u_t} + \Sigma_{pred}^{-1})^{-1}$
      $\mu_{upd} \leftarrow \Sigma_{upd}(\mathbf{M}_{u_t}^T\Sigma_{u_t}^{\mathbf{y}-1}\mathbf{y}_t + \Sigma_{pred}^{-1}\mu_{pred})$
      $c_{upd} \leftarrow \frac{\omega c_{pred}|\Sigma_{upd}|^{0.5}}{|\Sigma_{upd}|^{0.5}|\Sigma_{u_t}^{\mathbf{y}}|^{0.5}} e^{-\frac{1}{2}\|\mathbf{y}_t-\mathbf{M}_{u_t}\mu_{pred}\|^2_{\Sigma'_{upd}}}$
      $\mathcal{M}' \leftarrow \mathcal{M}' \bigcup (c_{upd},\mu_{upd},\Sigma_{upd})$
      if $s_t \neq s_{t-1}$ **then**
        Compute (c,μ,Σ) to approximate $\mathcal{M}'$
        $\mathcal{M}_{s_t}^t \leftarrow \mathcal{M}_{s_t}^t \bigcup (c,\mu,\Sigma)$
      else
        $\mathcal{M}_{s_t}^t \leftarrow \mathcal{M}_{s_t}^t \bigcup \mathcal{M}'$
      end if
    end for
  end for
  for all segments s **do**
    $P_s^t \leftarrow \sum_{(c,\mu,\Sigma)\in\mathcal{M}_s^t} c$
    $\mathcal{M}_s^t \leftarrow \{(\frac{c}{P_s^t},\mu,\Sigma)|(c,\mu,\Sigma) \in \mathcal{M}_s^t\}$
  end for
  Return: Posterior at t, $\{P_s^t,\mathcal{M}_s^t\}$

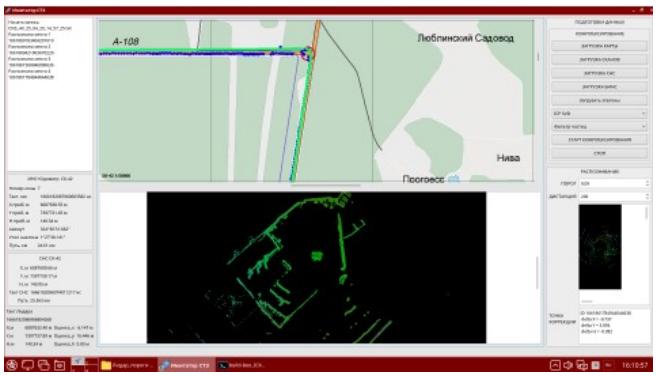


Fig. 4. Example of a run with a matching measure of 0.09. Green line - trajectory with lidar correction, line of blue dots - trajectory with search algorithm, blue line - trajectory of the inertial system



Fig. 5. Example of a run with a matching measure of 0.15. At an intersection, a place recognition error is distinguishable. Green line - trajectory with lidar correction, line of blue dots - trajectory with search algorithm, blue line - trajectory of the inertial system

VII. The Conclusion

Table 1 shows that the mean of the survey-comparison method depends on the matching measure in the recognition algorithm. Its increase (0.15) leads to an increase in correction points due to false positives, thus increasing the error of the inertial system, and excessive reduction (0.09) - increases the error too due to the reduction of correction points. In contrast to the survey-comparison method, the use of the search algorithm in the navigation system does not depend on manual selection of additional parameters and shows a smaller average error in all test series.

VIII. References

[1] Шолохов А.В. «Коррекция наземных навигационных систем по цифровой карте дорог с учетом ее погрешностей», Сборник докладов V конференции молодых ученых «Навигация и управление движением», Концерн "ЦНИИ "Электроприбор", Санкт-Петербург, 2003.

[2] M. A. Brubaker, A. Geiger and R. Urtasun, "Lost! Leveraging the crowd for probabilistic visual self-localization", Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 3057-3064, 2013.

[3] M. A. Brubaker, A. Geiger and R. Urtasun, "Map-Based Probabilistic Visual Self-Localization," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 4, pp. 652-665, 1 April 2016.

[4] Nielsen, F. On a Generalization of the Jensen–Shannon Divergence and the Jensen–Shannon Centroid. Entropy 2020

[5] Adel Javanmard, Maya Haridasan and Li Zhang, "Multi-track Map Matching", in CoRR, 2012

[6] S. M. Oh, S. Tariq, B. N. Walker, and F. Dellaert, "Map-based priors for localization", in ICIRS , 2004.

Technique and System of Contactless Optoelectronic Dual-Beam Time-of-Flight Determination of the Velocity Vector of Micrometeoroids

P.K. Kuznetsov
Samara State Technical University
Samara, Russia
kurnesov@mail

E.A. Schelokov
J.S.C.Rocket and Space Center "Progress"
Samara, Russia
riddick41666@mail.ru

***Abstract*—During long-term operation of Earth remote sensing (ERS) orbital vehicles, one of the factors that significantly affects the service life of the vehicles' optical sensors is space debris, and in particular, micrometeoroids. Meteoroids (MM) have sizes of about 1-100 μm and have velocities of up to 30 km/S. The distribution and density characteristics of large-format space debris flows in the areas of ERS orbital operation are observable and fairly well studied, but the orbital parameters, the density characteristics of MM flows, and their velocities have not been sufficiently studied to date. This article proposes a technique and an optoelectronic system for recording and determining the parameters of MM**

Keywords—— space debris, micrometeoroids, optoelectronic system, registration, determination parameter, velocity vector

I. INTRODUCTION

There are known techniques for measuring the density of micrometeoroid fluxes based on the use of flat metal MM impact recorders, which are delivered to the Earth to analyze the number and speed of MM, which is very inconvenient and uninformative in practical use. The purpose of the work is to ensure safety and extend the active service life of Earth remote sensing devices (ERSDs) by expanding the functionality and reducing the error of the information and measuring system (IMS) of contactless optoelectronic time-of-flight measurement of the micrometeoroid (MM) velocity vector in the vicinity of ERSD operation orbits. The IMS for determining the MM velocity vector parameters (MM VVP) is a system of flat light curtains made on the basis of laser radiation sources and high-speed photodetectors. The MM VVP implements the time-of-flight method for measuring the MM movement velocity.

The report presents a description of the design of the information measuring system (IMS) MM and provides the results of theoretical studies of the error of such IMS, as well as the results of ground tests of their prototypes and the minimum achievable error in measuring the parameters of the MM velocity vector, which is estimated at 12-15%.

II. STRUCTURE OF THE SYSTEM FOR REGISTRATION AND DETERMINATION OF MICROMETEOROID MOTION PARAMETERS

In order to solve the set problem, the structural diagram of the block of optical elements (BOE) IIS for determining the components $V_{мx}, V_{мy}, V_{мz}$of the velocity vector $\vec{V}_м$ of the MM, presented in Fig. 1, was developed.

The structural elements that determine the functional capabilities of the IMS and the measurement error of the MM motion parameters are radiation sources (RS), photodetectors, reflecting planes and a timer in the microprocessor. Remote units of meteoroid registration zones and the remote control and primary data processing unit (RCP) associated with them are located at a distance of up to several kilometers from the spacecraft. The RCP is located next to the micrometeoroid registration zone (MRZ) at a distance of up to several meters. Communication between the RCP and the spacecraft onboard computing system (SOCS) is carried out via a radio channel through a wireless transceiver unit (WTU).

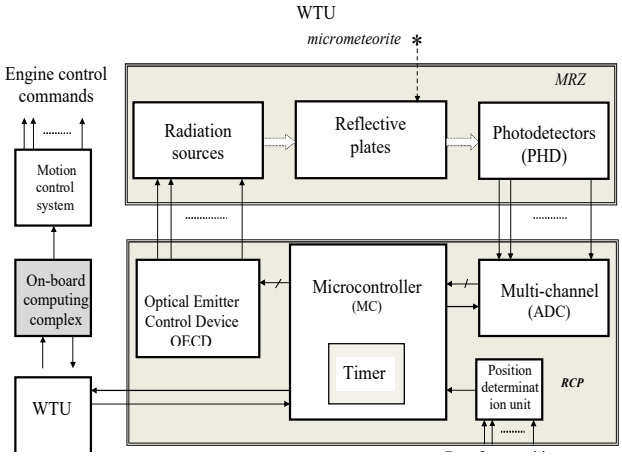


Fig. 1. Structural diagram of the onboard IIS VS MM, MRZ – micrometeorite registration zone, WTU – wireless transceivers, RCP – remote control and primary data processing unit, associated with them, are located at a distance of up to several kilometers from the spacecraft.

The microcontroller (MC) controls the following operating modes: supplying power to the control devices of emitters and photodetectors (PD), inputting signals from the PD into the multichannel ADC, retrieving and processing information received from the ADC, and the collection, systematization and transmission of measurement information and data on the state and spatial position of remote units from the position determination unit (PDU) in the spacecraft UAV to generate tasks for the motion control system (MCS) of the remote units.

Known MRZ types consist of two types of light curtain panels (LCP): LED and laser. In ballistic measuring systems of the FED-7M type, two parallel PSZ A and B are shifted from each other by a distance along the z axis of the object's movement and are normal to this axis. With this method of measurement, only the projection of the velocity vector onto the axis z is measured:

$$z\text{: } V_z = l_z/(TzB — T_z^{A-})V_{MM},$$

where T_z^{A-} and T_z^{B-} are the initial values of the laser beam blocking interval in LCP A and B, respectively.

In an IMS with laser MRZ, the light curtain is formed by multiple beam reflections from two parallel plates with a mirror coating of length l_x, spaced apart by a height высоту l_y, in such a way that the step b_x of the beam grid is less than the minimum permissible size D_x of the object, and the beam after all reflections from the plates in the absence of an object hits the photodetector (FD) along a given trajectory (Fig. 2a). The height of the MRZ depends on the value δt of the timer scale division, the time readings of which exactly coincide with the spatial points of beam reflection from the mirror plates. An object passing through the light curtain interrupts the beam for a time measured by the timer. The advantages of such IMS include high performance and a wide range of measured velocities. The disadvantages include a narrow aperture of the registration zone, determination of only the z coordinate and the V_z component of the $\vec{V}$ velocity vector of the object; lack of consideration and compensation for the influence of N reflections of $\vec{V}$ the laser beam, external and internal background illumination on the measurement error.

Fig. 2 shows the design options for a single-beam MRZ (a) and a double-beam counter-parallel MRZ (b). Blind zones S_T are shown in dark color. The MRZ fill factor $k_{RSZ} = (S_{SZ} - S_T)/S_{SZ}$, where S_{SZ} is the area of the light curtain, S_T is the total area of the blind zones.

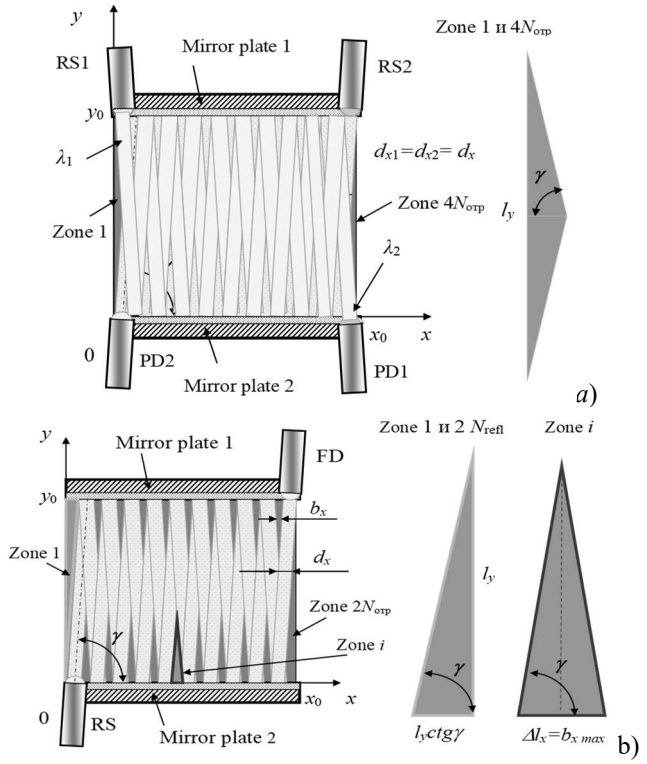


Fig. 2. Single-beam MRZ (a) and double-beam counter-parallel parallel MRZ (b): RS – radiation source; PD – photodetector; λ1 and λ2 – values of central wavelengths RS1 and RS2, respectively (λ1≠λ2); $\Delta l_x = b_{xmax}$ – distance between adjacent beams on the mirror plate

The fill factors of single-beam and double-beam counter-parallel MRZ at $d_{x1} = d_{x2} = d = b_{xmax} = \Delta l_x$ and $\gamma = \gamma_1 = \gamma_2$ are equal, respectively:

$$k_{з1} = 1 - \frac{2(N_{отр}-1)d\,\mathrm{ctg}\gamma}{l_x} - \frac{d^2\,\mathrm{tg}}{2l_y l_x} \quad (1)$$

$$k_{з2} = 1 - \frac{l_y\,\mathrm{ctg}\gamma}{l_x} \quad (2)$$

The paper proposes a modified method of two-beam counter-parallel time-of-flight determination of the MM velocity.

The expansion of the number and ranges of the parameters to be determined is due to the measurement of the beam-break intervals in the MRZ from group I of spatially shifted mutually parallel and orthogonally oriented MRZ with a large value of S_{SZ}. In accordance with the mathematical model of interaction of the MM and MRZ during the MM flight through the MRZ, the coordinates (x_{Mi}, y_{Mi}, z_{Mi}, i=1), dimensions (D_{Mx}, D_{My}, D_{Mz}) and the velocity vector $\vec{V}_M$ of the MM are determined.

Fig. 3 shows an example of a MRZ consisting of two spatially combined x- and y- coordinate LCP A_{xy} and B_{xy}, spaced along the z axis by a distance l_z. The principle of determining the size of the MM is shown based on the results of measuring the time intervals of the beam interruption at the input (point $\mathbf{M^A}$ (x_M=0…l_y; y_M=0…l_y; z_M=0)) and at the output (point $\mathbf{M^B}$ (x_M=0…l_x; y_M=0…l_y; z_M=l_z)) MM from MRZ.

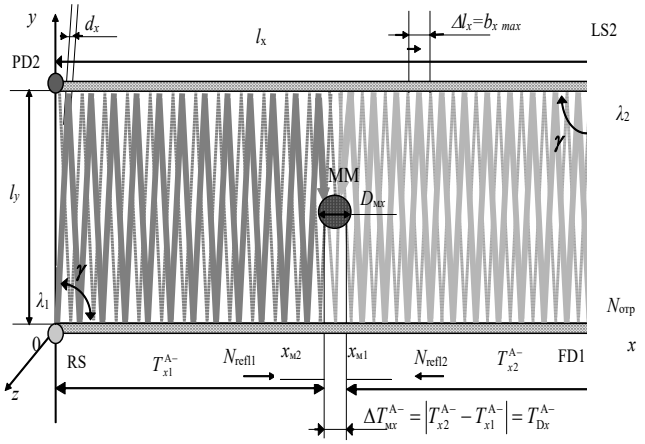


Fig. 3. Projection onto the coordinate plane 0yz of the MM flight trajectory from MM A to the MRZ B being parallel to it, shifted along the z axis by the value lz . Where N_{refl} i , i ∈ {1, 2} - number of reflections in zones 1 and 2.

Figure 4 illustrates the model of the relationship between time and the coordinate of the beam interruption in the x - coordinate LCP A_x.

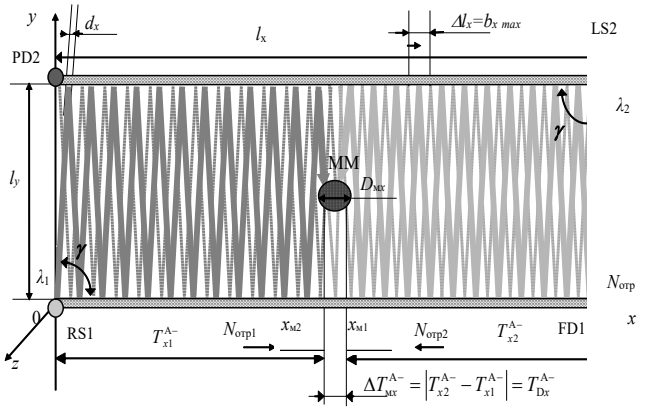


Fig. 4. Relationship between the start time of the interruption of the rays RS1 (T_{x1}^{A-}) and RS2 (T_{x2}^{A-}) with the position of the MM in the x-coordinate LCP A_x

A number of mathematical models have been developed that link the parameters of the MM and modified MM

registration zones, taking into account the geometric dimensions, MM velocity and the main design parameters of the MRZ. These relationships form the basis of the mathematical model of the IMS. Fig. 5 shows the algorithm of operation of the measuring channel of the system built on the basis of the developed mathematical model.

A number of mathematical models have been developed that link the parameters of the MM and modified MM registration zones, taking into account the geometric dimensions, MM velocity and the main design parameters of the MRZ.

These relationships (Fig. 4) form the basis of the mathematical model of the IMS. Fig. 5 shows the algorithm of operation of the measuring channel of the system built on the basis of the developed mathematical model. When the IIS is putting into action, the values of the constants determined by the geometric and energy parameters of the MRZ and the time scale T_c of the measurement cycle oriented to the lower limit of the range of measured MM velocity are set. Then the system forms the curtains, monitor the synchronicity of their operation and puts the MM into standby mode. When the MM flies into the MRZ at the moment of the beginning of the interruption interval of the first beam, the cycle and scale of measuring of the beam interruption intervals in all curtains are launched (see Figs. 3 and 4). The coordinates of the MM in the MRZ, their geometric parameters and the components of the velocity vector are determined by these values. If during the movement of one MM in the ZRM no other objects appear, then the system determines the size of the MM and its velocity. After the MM leaves the MRZ, the system returns to the standby mode. For example, when the MM flies into a composite curtain A , the values $T_{x1}^{\text{A}-}$ and $T_{x2}^{\text{A}-}$ of the beginning of the interval of interruption of the rays RS1, RS2 in the curtain Ax and the values $T_{y1}^{\text{A}-}$and $T_{y2}^{\text{A}-}$ of the beginning of the interval of interruption of the rays in RS and RS4 in the curtain A_y combined with it are first measured (Fig. 5). For spatio-temporal and energy coordination, curtains A_x, A_y, B_x, B_y are made with identical geometric and energy parameters with different wavelengths RS.

Let the timer scale starts when the beam from RS1 appears in the curtain Ax. Then $T_{x1}^{\text{A}-}= 0$ is the initial count of the timescale, and in further calculations the value $D_{\text{м}x}$ obtained from the results of measuring the beam interruption time in the curtain B_x. is used.

Due to the identity of the curtain sizes ($l_x = l_y$) the test time $T_{\text{сз}+}$ of the formation and $T_{\text{сз}-}$ of the disconnection of the curtains without MM is equal to

$$T_{\text{сз}} = T_{\text{сз}-} = T_{\text{сз}+} = \frac{N_{\text{отр}} l_y}{c\sin(\gamma)} = \frac{N_{\text{отр}} l_x}{c\sin(\gamma)} \qquad (3)$$

Then, using the value of $T_{x2}^{\text{A}-}$, we can obtain the start time of the interval of switching off the RS1 beam relative to the conditionally backward shifted start time of the timer scale as

$$T_{x1}^{\text{A}-} = T_{\text{сз}} - T_{x2}^{\text{A}-} - T_{D\text{м}x}^{\text{A}} = T_{x1}^{\text{B}-} + T_{x2}^{\text{B}-} - T_{x2}^{\text{A}-}. \qquad (4)$$

To determine the geometric dimensions $D_{\text{м}x}^{B}$ (see Fig. 4) and the coordinate $x_{\text{м}0}^{\text{B}}$ of the center of MM along the *x* axis in the curtain B_x, the start time of the beam interruption intervals from RS1 and RS2 is used (see Fig. 6):

$$T_{Dx}^{\text{B}} = T_{\text{сз}} - (T_{x1}^{\text{B}-} + T_{x2}^{\text{B}-}), \qquad (5)$$

$$T_{\text{м}0x}^{\text{B}} = T_{x1}^{\text{B}-} - \frac{T_{\text{сз}} - (T_{x1}^{B-} + T_{x2}^{B-})}{2} = \frac{T_{\text{сз}} + T_{x1}^{\text{B}-} - T_{x2}^{\text{B}-}}{2}. \qquad (6)$$

where T_{Dx}^{B} – time proportional to the size $D_{\text{м}x}^{\text{B}}$, T_{x0}^{B} – time proportional to the coordinate время, $x_{\text{м}0}^{B}$.

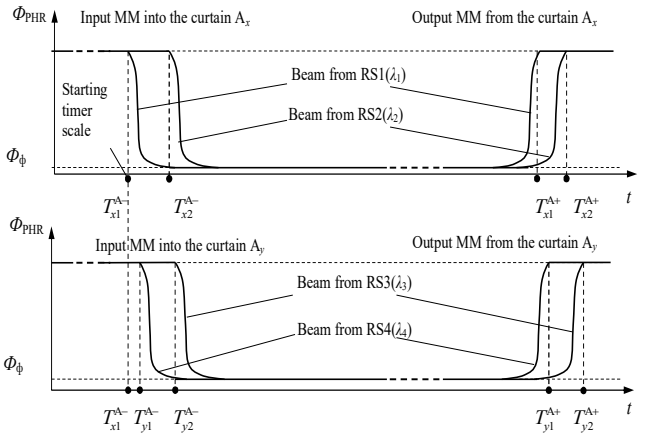


Fig. 5. Measuring the RS1-RS4 beam-break intervals in PS Ax and Ay to determine the coordinates and geometric dimensions of the MM

Taking into account the trajectory of the multiple reflected beams, the coordinate axis *x* is set, which, with precise spatial-temporal coordination of the scale of the T_c timer operation cycle with the spatial scales l_x and l_y of the curtains, gives the exact value of the coordinate of the center of the MM and the size $D_{\text{м}x}^{\text{B}}$ of the MM along the x axis: $x_{\text{м}0}^{\text{B}}$:

$$x_{\text{м}0}^{\text{B}} = cT_{x0}^{\text{B}}\cos(\gamma), \qquad D_{\text{м}x}^{\text{B}} = cT_{D\text{м}x}^{\text{B}}\cos(\gamma).$$

In a similar way, in accordance with (3-7), the values of the coordinates $x_{\text{м}}^{A}$, $y_{\text{м}}^{A}$, $y_{\text{м}}^{B}$ and the geometric dimensions $D_{\text{м}x}^{\text{A}}$, $D_{\text{м}y}^{\text{A}}$, $D_{\text{м}y}^{\text{B}}$ are determined. The coordinates of the movement of the center MM along the z axis are known a priori: $z_{\text{м}0}^{\text{A}} = 0$, $z_{\text{м}0}^{\text{B}} = l_z$.

The size $D_{\text{м}z}$ is determined after determining the component $V_{\text{м}z}$ of the velocity vector MM by measuring the interruption interval of any beam in any curtain, for example, by measuring the interruption interval of beam RS2 in curtain A_x (see Figs. 3 and 4):

If we assume that the acceleration $d\vec{V}_{\text{м}}/dt = 0$, then based on the obtained values of the coordinates $x_м$, $y_м$ and the intervals of the beam interruption in the curtain A ($T_{\text{м}z}^{\text{A}-} = 0$) and in the curtain B ($T_{\text{м}z}^{\text{B}-}$) the time interval

$$(T_{\text{м}z}^{\text{B}-} - T_{\text{м}z}^{\text{A}-} = T_{\text{м}z}^{\text{B}-})$$

between the beginning of the intersection of related coordinate-sensitive curtains of planes A and B is meagered and then, in accordance with the model (3) – (8) of the interaction of MM and MRZ after obtaining the values of the beam interruption and dimensions of MM, the modulus of the velocity vector is determined as:

$$|\vec{V}_{\text{мм}}| = \sqrt{\left(\frac{x_{\text{м}}^{\text{B}-} - x_{\text{м}}^{\text{A}-}}{T_{\text{м}x}^{\text{B}-} - T_{\text{м}x}^{\text{A}-}}\right)^2 + \left(\frac{y_{\text{м}}^{\text{B}-} - y_{\text{м}}^{\text{A}-}}{T_{\text{м}y}^{\text{B}-} - T_{\text{м}y}^{\text{A}-}}\right)^2 + \left(\frac{l_z}{T_{\text{м}z}^{\text{B}-}}\right)^2} \qquad (7)$$

and the angles of direction of its movement in space:

$$\alpha_x = \arccos\left(\frac{x_{\text{м}}^{\text{B}} - x_{\text{м}}^{\text{A}}}{|\vec{V}_{\text{мм}}|(T_{\text{м}x}^{B} - T_{\text{м}x}^{A})}\right);\ \alpha_{\text{y}} = \arccos\left(\frac{y_{\text{м}}^{\text{B}} - y_{\text{м}}^{\text{A}}}{|\vec{V}_{\text{мм}}|(T_{\text{м}y}^{\text{B}} - T_{\text{м}y}^{\text{A}})}\right);(8)$$

$$\alpha_z = \arccos\left(\frac{l_z}{|\vec{V}_{\text{мм}}|T_{\text{м}z}^{\text{B}}}\right).$$

When solving the problem of matching the spatio-temporal and energy parameters of the MRZ, an expression was obtained for the maximum possible number of reflections N_{ref}, taking into account the a priori known energy parameters of the optoelectronic devices and the reflection coefficient of the mirror plates:

$$N_{\text{отр}max} = log k_{\text{отр}}(\Phi_{threh}/\Phi_0), \quad (9)$$

where Φ_0 - the luminous flux at the RS output; Φ_{thrs} - the sensitivity threshold of the FP, k_{rfl} - the reflectivity of the mirror surface.

The resulting error of a parallel-type measuring system, to which the developed system belongs, consists of the methodical measurement error, digital conversion of time into spatial indicators, and the instrumental error of the MRZ. When determining the total error of the BOE IIS, the identity of all single-coordinate CCDs is taken into account.

The errors of the elements and cannels of the measuring part of the system were considered independently of each other, then the total instrumental error of the measuring part of the IMS is equal to:

$$E_0 = \sqrt{\sum_{j=1}^{J} E_{1.j}^2 + \sum_{j=1}^{J} E_{2.j}^2 + \sum_{j=1}^{J} E_{3.j}^2 + \sum_{j=1}^{J} E_{4.j}^2 + E_5^2}, \quad (10)$$

where $E_{1.1} - E_{1.J}$ – errors of the IR; $E_{2.1} - E_{2J}$ E; 2.1 – E2J – errors of single-beam light curtains; E3.1 – E3.J – errors of PHR; E4.1 – E4.J – errors of amplifier-comparators, ADC; E5 – timer error; J=8 – number of measuring channels (the presented errors have a normal distribution law).

An absolute error of the method when measuring the velocity vector by the proposed method will be no more than one discrete measurement of the time of the fact of the particle's flight through each light curtain, which is no more than τ_m=l/s. Considering that in the worst case $|\vec{V}_{\text{мм}}| = 30\ \text{km/c}$) the error of the measured time by the proposed method is less than the ratios $(T_{\text{м}x}^{B} - T_{\text{м}x}^{A})$, $(T_{\text{м}y}^{B} - T_{\text{м}y}^{A})$ and values $T_{\text{м}z}^{\text{B}}$ by six orders of magnitude, it is permissible not to take the error of the method into account in the calculations.

When a beam is reflected from the MM, the background light flux Φ_{Φ} can partially enter the PM location zone and exceed the threshold level of its sensitivity ($\Phi_{\Phi} > \Phi_{thrs}$). A method is proposed for compensating for the effect of background illumination on the transformation error and eliminating the omission of the MM by placing an additional PHD in the blind zone of the light curtain next to the main one, the signal from the output of which, in accordance with the algorithm (see Fig.5), is used to make a decision on the possibility of determining the MM parameters by registering only the fact of the MM being in the MRZ, as well as to reduce the measurement error taking into account the sources and nature of the error components.

The main operational factors influencing the measurement results were identified and the dependence of the relative current transmittance coefficient was obtained:

$$\frac{\Delta K_I}{K_I} = \frac{1}{K_h}\frac{dK_h}{d\Theta} + \frac{1}{K_{\text{изг},\Sigma}}\frac{dK_{\text{изг},\Sigma}}{d\Theta} + \frac{(1+\tau_0 K_{y,n}\Phi_n)^{-m}}{K_{I0}}\left(\frac{dK_I}{d\Phi_\gamma} + \frac{dK_I}{d\Phi_n}\right), \quad (11)$$

where K_i - current transfer coefficients (with bending of mirror plates and radiation background, respectively); Θ – temperature; Φ_γ, Φ_n – integral fluxes of γ-radiation and protons, respectively, τ_0 – lifetime of charge carriers during radiation background.

The coordination of spatial and energy parameters of measuring channels and tolerance control of dominant errors aimed at reducing the error in determining the coordinates of the beam interruption are proposed.

The introduction of self-monitoring functions for the IMS performance during the inter-verification time interval $\Delta t_i = t_{i+1} - t_i$, $i = 0,1,2,...$ is justified, which allows for determining whether the error goes beyond the acceptable value E_{acs} and making decisions on conducting tests and calibrations using the finite distribution of the error:

$$F_{Ci}(t) = \varkappa\{F(t_{i+1}) - F(t_i)\}, \quad (12)$$

where $\varkappa$ is the normalizing factor determined by the boundary conditions $E_{0\text{ecs}}$, the achievement of which means the metrological failure of the IMS. *F(t)* corresponds to the Bernstein distribution law.

A laboratory model of the proposed IIS has been developed. The model contains two mirror plates with a silver coating; a frame with a 4-position adjustment system using a laser level. An imitation of particles commensurate in linear size with a laser beam has been produced on the mirror surface.

The obtained results of the study of the model of IMS with two MRZ are comparable with the theoretical ones. Silver was chosen for the production of reflective surfaces, which are most critical to operating conditions (k_{refl} =0.991÷0.992 at λ=0.5 MCM and Θ=-120÷150 ^{0}C).

The possibility of creating an industrial prototype of the measuring system discussed in the article has been demonstrated.

REFERENCES

[1] Gladyshev, A.I., Telegin, E.A., and Shchelokov E.A., Mathematical model of an optical system for measuring the parameters of high-speed microparticles, in: Tsifrovaya obrabotka signalov (Digital signal processing), 2023, no. 1. pp. 39-42

[2] Shchelokov, E.A. Method and device for estimating the micrometeoroid motion parameters based on optical systems, Vestnik of RGRTU, 2016, no. 56, pp. 131-135.

Objectives and Certain Results of Pointing ISS Scientific Instruments Towards their Targets in Earth Science Experiments

M.Yu. Belyaev
S.P.Korolev Rocket and Space Corporation Energia
Korolev, Moscow region, Russia
Mytishchi branch of Bauman Moscow State Technical University
mikhail.belyaev@rsce.ru

P.A.Borovikhin
S.P.Korolev Rocket and Space Corporation Energia
Korolev, Moscow region, Russia
pavel.borovikhin@rsce.ru

A.M. Esakov
S.P.Korolev Rocket and Space Corporation Energia
Korolev, Moscow region, Russia
aleksey.esakov2@rsce.ru

***Abstract* – The paper discusses some aspects and methods of pointing scientific instruments (SI) of the space experiment (SE) Uragan from the ISS towards their targets. It describes methods for manual and automatic pointing of the equipment. The methods enable the crew to point the SI towards targets on the ground with high degree of reliability. For automatic pointing of the SI, methods were proposed which make it possible the probability of successful imaging and draw up optimal target observation programs. It provides examples of using the methods in space experiment Uragan onboard the ISS.**

Keywords — International Space Station ISS, scientific equipment, pointing, probability of capturing an image, optimization

I. Introduction

The program of research onboard the ISS pays considerable attention to development of scientific instrumentation (SI) and technologies for Earth studies in space experiments (SE) Uragan (Hurricane) and Stsenari (Scenario) [1]. Important objectives in refining Earth studies technologies are to improve reliability and precision of pointing SI towards their targets and to increase the amount and quality of obtained scientific data under constraints imposed by station resources [2]. The task of pointing scientific instruments towards their targets, in the case of the ISS, is made difficult due to the large size and mass of that station and low angular momentum available from the control moment gyros (19000 Nms) installed onboard the US orbital segment of the ISS and used for station attitude control. For this reason, SE Uragan uses manual methods of pointing scientific instruments towards their targets, as well as specially designed pointing platforms (PP) and special pan-and-tilt devices [1,2].

In order to improve reliability of the SI pointing at their targets, cosmonauts onboard the space station employ an approach based on preliminary math simulation of the observed Earth surface. Simulated images of surface areas to be observed by the crew that include target designations are uplinked to the ISS prior to the imaging session. Implementation of this approach to running SE prevented errors in manual pointing of scientific instruments at their targets [3].

In order to improve reliability of target capture using an automatic PP, a series of images needs to be taken [4]. Thus, in the course of a mission, a problem emerges: What is the optimal number of images, in what directions, and at what points in time they need to be taken in order to make sure that the probability of successful capture is not below a specified value? In order to solve this problem, experiments onboard the ISS that involve the use of SOVE PP apply the approach that was first proposed by academician A.N. Kolmogorov in his theory of gunnery practice [5]. The obtained solution of the problem was implemented onboard the ISS in the Sigma system.

In order to increase the amount and quality of data obtained in a SE in Earth science, optimal planning methods for observation sessions are employed [6]. In particular, methods implemented onboard the ISS made it possible to improve efficiency of research programs in SE Uragan and Stsenari using our country's first Hyperspectrometer instrument that operates in visible and IR regions of electromagnetic spectrum, which was delivered to the ISS in June of 2024 [2,7].

II. Improving Reliability of Crew Pointing Scientific Instruments Toward their Targets

The lack of capability to establish and maintain the ISS attitude required for conducting some of the experiments made it necessary to develop special approaches to solving this problem. During the early phase of the ISS mission, the SI (photographic cameras, spectrometers) used in the SE Uragan were pointed towards their targets by the crew. In this case, the reliability of capturing a target depends on the ability of the crew to visually identify the target in due time and point the portable equipment towards it. In some instances, especially when one needs to photograph a relatively small target with high resolution (long-focal-length lenses for such photography have a narrow field of view), the crew may fail to identify the target and take the picture. Such a situation is especially critical when there are synchronous subsatellite observations scheduled to be performed simultaneously with photography from space, which require organizing special expeditions to the target areas.

To solve this problem, a method described in [3] was developed and is used. In accordance with this method, a digital model of the target area is used in order to prepare in advance its pictures representing the actual view the crew will see from the porthole when overflying the target at the time scheduled for imaging. Pictures for the crew are generated to various scales. Each picture shows the field of view through the porthole depending on possible available pointing angles of the equipment that is used and the ISS flight path, as well as highlights the outlines of the target (or a group of targets). These pictures are provided to the crew along with other instructions for carrying out the imaging.

Thus, the crew, having familiarized themselves with the task and the simulated images, performs the imaging reliably. The realism of the images simulated ahead of time makes it possible for the crew to identify the targets, to point the equipment and take pictures of the target in a timely manner.

III. USING ARTIFICIAL SCATTER DURING IMAGING

Probability of successfully capturing the target even after multiple attempts may become significantly reduced if its position with respect to the imaging equipment is known with a significant error due to insufficient accuracy of such current data as the position of the ISS center of mass, SI attitude, time reading. The knowledge of geographical coordinates of the target may also lack accuracy, for example during search and rescue missions.

In cases like this, the probability of successful capture with the same number of attempts is achieved through the use of artificial scattering, that is, by purposefully deflecting the SI line of sight away from the most probable target location point calculated from current data. Used in order to determine quantitative characteristics for such a scattering are results obtained by academician A.N.Kolmogorov in his mathematical theory of artillery firing a series of shots at a point target. For instance, as applied to capturing photos of ground targets from the ISS, hitting a target with a missile corresponds to the target getting into the frame of the photo, the size of the gun fire target corresponds to the size of the frame, while the gun azimuth and elevation settings correspond to the two angles of pointing the optical axis of the lens towards the target [4].

Effective application of artificial scattering is possible when it is sufficient to capture the picture of a target at least once in the course of several attempts, and the factors that generate uncertainty in the target position data simultaneously affect the entire imaging session, that is, they actually result in systematic errors, the exact values of which are not known a priori.

Of special practical importance is the analogy with a particular case of shooting, when the gun aiming is characterized by a single adjustable parameter [8] (and one exact hit is sufficient for the kill). Similar conditions, in certain cases, also occur during imaging of ground targets from space. [4]. For example, due to inaccuracy in trajectory data or time reading, the direction of maximum uncertainty in ground target position with respect to the ISS may coincide with the direction of its orbital motion. Then the need to introduce artificial scattering and its parameters are determined by the balance between the planned number of images n and parameters k and c:

$$k = \sigma/\sigma_O, \;\; c = d/\sigma_O,$$

where σ is the root-mean-square deviation (along the ISS orbit) of the actual direction of the camera lens axis from the preset direction;

σ_O is the root mean square deviation of the ground target along the ISS orbit;

d is the size of the area on the Earth surface that gets within the frame measured along the ISS orbit.

For example, if the uncertainty of the apparent position of a ground target along the orbit is characterized by the root mean square value σ_O = 1°, while the accuracy of pointing along the orbit is estimated to have the root mean square value of σ= 0.2°, estimates show that a Nikon D5 camera with 1200 mm lens focal distance (the angular size of the frame of no less than 1.14°) has the more than 0.90 probability of successfully capturing the target when taking multiple shots in three directions. The central direction coincides with the direction towards the most likely position of the target, while the two other are deflected in different directions from the central direction along the ISS flight course through the same angular distance of 1.35°. Furthermore, no less than two shots must be taken in each of the three directions. Had all of those six shots been taken exclusively in the direction of the most likely position of the target, the probability of the target getting into the frame would have been less than 0.59.

If σ_O = 5°, σ = 1°, and the frame size is 10°, in order to exceed the 0.96 probability of successful shot it is sufficient to take shots in two directions that do not coincide with the direction towards the most likely position of the target, but rather are symmetrically directed to the sides, along the flight course. The angle between those two directions is about 10°, no less than three shots are taken in each direction. But if one were to take all these six shots in the direction of the most probable position of the target, the probability of successful capture would be less than 0.79.

If σ = 0° (that is, the lens is pointed towards the preset direction virtually without any random deflections), while $\sigma_O > 0°$, then in one-dimensional case the pointing is done symmetrically towards both sides away from the most likely position of the target in increments of d.

If the area of possible positions of the target is two-dimensional (for example, circumscribed by an ellipse), the imaging is similar to areal target imaging and its algorithm may be determined, for example, by the optimal use of the SI pointing platform mechanisms.

IV. OPTIMIZING TARGET POINTING PROGRAM

Observation program optimization is used when planning experiment sessions that involve Earth surface observations in order to improve their informative value (the quantitative estimate of the value of scientific information obtained from the study of the target). The developed methods take into account the characteristics and special aspects of using SI of the SE Uragan that impose constraints on observation sessions onboard the ISS [2,7].

When implementing a scientific program that uses SI Hyperspectrometer for a specified planning interval lasting from several days to weeks or months, several linked target observation sequences selected from a specified general targets list can be put together. Such sequences are traditionally called zones of observations (experiments). When planning and conducting observations with the use of SI, the informative value of such sequences is calculated as follows [2]:

$$I_j = \sum_{j=1}^{N} p_i o_i l_i s_i, \quad j = 1, \dots, M, \qquad (1)$$

where M and j are the number of zones and their indices;

N and i are the number of all the targets in a zone and their indices;

p_i is the priority of the i-th target объекта ($1 \le p_i \le 10$)

o_i is the cloud cover factor during observations of the i-th target ($1 \le o_i \le 10$);

l_i is the Sun lighting factor during observations of the i-th target ($l_i = 1 - \left|h_i^{opt} - h_i\right|/h_i^{opt}$), where h_i is the expected actual angle of Sun elevation above local horizon of the i-th target, and h_i^{opt} is the Sun elevation that is optimal for observing that target);

s_i is the scanned surface area ratio for the target ($s_i = S_i'/S_i$, where S_i' is the captured surface area of the target and S_i is the total surface area of the target);

Various constraints are taken into account in planning of the experiments. Conditions and constraints taken into account when planning observations with SI Hyperspectrometer are listed in [2].

The method for optimizing observation program under real flight conditions was used in planning imaging conducted in the course of test sessions of SE Uragan with SI Hyperspectrometer in September 2024. The search for solution was conducted using a genetic algorithm. The optimization task was to maximize the informative value under constraints on the crew time and the use of the not always readily available workstation near the porthole. The targets were fires detected by unmanned spacecraft during the two days preceding the day when the observation program was drawn up, agricultural fields, as well as areas stained with man-made environmental pollution.

The program for searching the optimal solution was written in Python. In particular, it employed a genetic algorithm involving tournament selection, one-point crossover, mutations (0.9 and 0.1 probabilities, respectively) and elitism. The fitness function used the informative value formula (1). In addition to this the evaluation took into account the hardware parameters (fields of view, angles and pointing speed) that affect inter-compatibility of the targets: genome of individuals with incompatible targets was corrected.

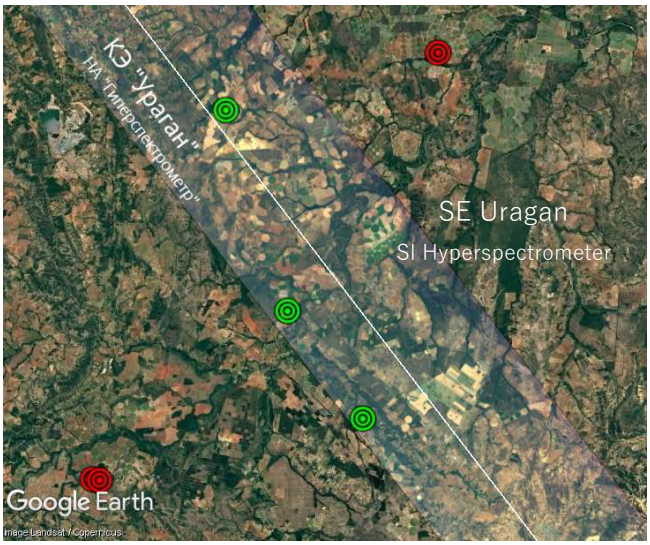


Fig. 1. Composite RGB image (52nd, 59th and 80th channels of the visible band camera of SI Hyperspectrometer), images captured on September 3, 2024 from the Russian Segment of the ISS (SE Uragan).

The maximum fitness (informative value) was achieved in the 28th generation. All the targets selected for imaging were successfully captured. Fig. 1 and 2, respectively, provide examples of data acquired in visible and near infrared bands, overlaid on an orthophotomap (from Google Earth): generated from the visible band imaging data was a composite RGB image (52nd, 59th and 80th channels of visible band camera of SI Hyperspectrometer); generated from the near-infrared band imaging data was a black an white image (38th channel of near infrared band camera). In both Figures, shown in white is the ISS ground track, bull's-eye symbols mark the targets available for observation in the zone, with those that were selected using the genetic algorithm being highlighted in green, while others are painted in red.

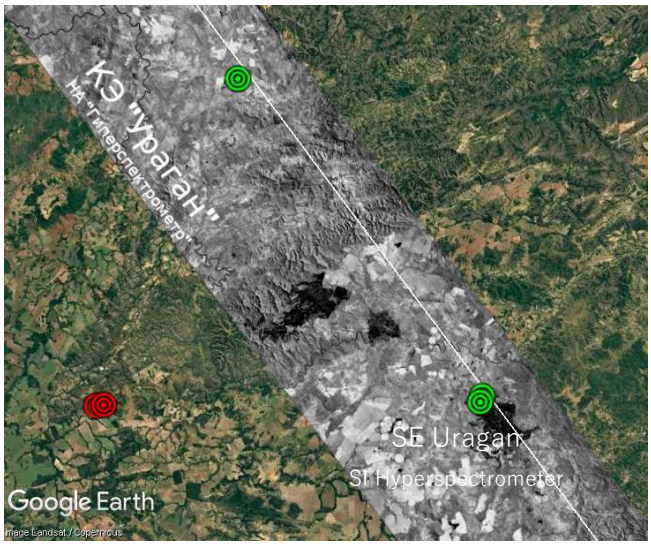


Fig. 2. Black and white image (38th channel of the near infrared band camera of SI Hyperspectrometer), images captured on September 3, 2024 from the Russian Segment of the ISS (SE Uragan).

The use of the above approach made it possible to achieve the required informative value of observation sessions, while saving hardware and station resources, as well as to automate the process of putting together an observation program. Automation of that process not only makes it possible to reduce the labor input required to put an observation program together, but also permits to update the observation program in real time depending on changing weather conditions in the area of the targets and priority (urgency) of their imaging, which is especially important when planning to capture various catastrophic events: wild fires, floods, etc.

V. SUMMARY

Some special aspects of the ISS made it necessary to develop and create new, special methods for effectively carrying out observations of targets on the Earth surface. Methods that were developed and introduced in the course of conducting experiments Uragan, Stsenari, Dubrava enable the crew to reliably point research instruments towards their targets and increase the probability of successfully pointing SI towards their targets using pointing platform SOVE in automatic mode. Moreover, the developed optimal planning methods for target observation programs that are used in experiments with SI Hyperspectrometer onboard the ISS make it possible to improve effectiveness (informative value) of the programs.

REFERENCE

[1] Belyaev M.Y. Scientific equipment and methods for studying Earth in the Uragan space experiment onboard the International Space Station. // Current problems in Earth remote sensing. 2021. vol. 18, No3. p. 92–107.

[2] Optimizing pointing of scientific instruments towards their targets in the *Uragan* experiment onboard the ISS / M.Yu. Belyaev , P.A. Borovikhin, A.M. Esakov [et al.] // Gyroscopy and navigation. 2023. v. 31. № 4(123). p. 118-137.

[3] A method of pointing portable equipment in a spacecraft towards its target. patent 2821223 RF / A.M. Esakov, M.Yu. Belyaev, D.Yu. Karavaev, P.A. Borovikhin. Application filed November 20, 2023; published June 18, 2024. Bulletin No.17.

[4] Borovikhin P.A. Methods of jointly solving the problems of pointing scientific equipment, navigation processing of images, and controlling the mission of spacecraft: Engineering Science Candidate's thesis. Moscow. 2022. 158 p.

[5] Kolmogorov A.N. Number of hits as a result of several shots and general principles of evaluating effectiveness of a shooting system. //

Proceedings of V.A. Steklov Mathematical Institute. 1945. vol. 12. p. 7–25.

[6] Belyaev M.Y. Scientific Experiments Onboard Spacecraft and Orbital Stations. Moscow: Mashinostroeniye, 1984. 264 p.

[7] Esakov A.M. Planning observation sessions for targets on the Earth surface from onboard the ISS RS. // Forestry Bulletin, 2020. vol. 24. No.5. p.109–115.

[8] Kolmogorov A.N. Artificial scattering in the case of one hit kill and scattering in one dimension. // Proceedings of V.A. Steklov Mathematical Institute. 1945. vol. 12. p. 26-45.

Determination of Stable Equilibrium Positions of the SamSat-Ionosphere Nanosatellite under the Influence of Gravitational and Aerodynamic Moments

E.V. Barinova
Inter-University Department of Space Research
Samara National Research University
Samara, Russia
L5545@yandex.ru

I.A. Timbai
Department of Higher Mathematics
Samara National Research University
Samara, Russia
timbai@mail.ru

Y. Mironov
Inter-University Department of Space Research
Samara National Research University
Samara, Russia
mironovevgeniy01@gmail.com

N.A. Elisov
NIL-102
Samara National Research University
Samara, Russia
mr07th@gmail.com

I.A. Lomaka
NIL-102
Samara National Research University
Samara, Russia
igorlomaka63@gmail.com

***Abstract*— The required spacecraft attitude is important for a many mission. Therefore, determination of the spacecraft equilibrium positions relative to its center of mass is a topical problem. Estimation of the aerodynamic characteristics is a complex problem, especially for spacecraft with deployable structures. Thus, this paper proposes the algorithm for estimation of aerodynamic characteristics for such spacecraft. Also, the paper suggests the algorithm for numerical estimation of stable equilibrium positions of the SamSat-Ionosphere under the influence of gravitational and aerodynamic moments.**

***Keywords*— nanosatelite, stable equilibrium position, panel method.**

I. INTRODUCTION

SamSat-Ionosphere nanosatellite (NS) was created at Samara University and was launched from the Vostochny cosmodrome in November 2024 [1]. The main objectives of NS are study of the Earth's ionosphere and magnetosphere. The payload consists of the magnetometric system, plasma parameters sensor (PPS), and receiver of navigation signals. To conduct scientific measurements, two attitude requirements must be satisfied. Firstly, the NS longitudinal axis must be close to the local vertical vector for receiving signals from the GLONASS system. Secondly, the plane between the sensitive elements of PPS must be perpendicular to the upstream flow. Thus, the NS attitude must be stabilized.

The average orbit altitude is 500 km. At this altitude, the gravitational moment dominates, however, the aerodynamic moment is also significant. Therefore, the three-axis gravitational-aerodynamic stabilization was chosen [2]. B-dot algorithm is used for decreasing the angular velocity after separation process. Knowledge of all stable equilibrium positions may be useful for an attitude determination and control system.

The research was supported by the Russian Science Foundation grant No. 23-67-10007, https://rscf.ru/project/23-67-10007/

Determination of equilibrium positions relative to the center of mass is an important problem. For example, Sarychev and Gutnik [3-6] estimated the equilibrium positions under influence of the gravitational and aerodynamic moments for spacecraft close to spherical form. The authors [7,8] obtained the analytical expressions for determining equilibrium positions in the case of rectangular NS, wherein we consider the two special cases. The first one is displacement of the center of pressure along three coordinates for dynamically symmetric NS. The second case is displacement of the center of pressure along one of the principal axes of inertia. Also, the numerical algorithm for equilibrium positions was developed in case of a NS with non-diagonal inertia tensor when the center of mass is displaced from the geometric center [9]. SamSat-Ionosphere has three deployable on-board structures: turnstile antenna, the deployable magnetometer, and the PPS deployable resonator. Thus, its aerodynamic characteristics significantly differ from those of NS without external structures.

This work is focused on determination of the stable equilibrium positions of SamSat-Ionosphere under the influence of gravitational and aerodynamic moments. To achieve this, two algorithms were developed. The first is designed to determine the aerodynamic characteristics (AC) of SamSat-Ionosphere, while the second is for the numerical search of all stable equilibrium positions relative to the center of mass.

In this paper, the equilibrium positions of SamSat-Ionosphere are determined for different atmospheric density models, considering the day/night atmospheric variations and solar flux index.

II. MATHEMATICAL MODEL

To describe attitude motion of NS, we used two reference frames. The trajectory reference frame *OXYZ* coincides with the orbital one in the case of circular orbit. The body-fixed *Oxyz* reference frame with the origin in the NS center

of mass. The body-fixed *Ox* axis is directed along the longitudinal axis of the NS, the *Oy* and *Oz* axes are perpendicular to its lateral sides. To define the relation between the trajectory and the body-fixed reference frames, we used three Euler angles. The first angle α is the spatial angle of attack (the angle between the longitudinal axis and the velocity vector of the center of mass). The second angle ψ is the angle of precession (the angle between the planes of flight and spatial angle of attack). The third angle φ is the angle of proper rotation (the angle between the plane of the angle of attack and the transverse axis of the NS) [10].

NS attitude motion in the body-fixed reference frame is described by the following vector equation:

$$J\frac{d\vec{\omega}}{dt}+\vec{\omega}\times J\vec{\omega}=\overrightarrow{M^e},$$

$$\vec{\omega}=\overrightarrow{\dot{\alpha}_n}+\vec{\dot{\varphi}}+\vec{\dot{\psi}}+\overrightarrow{\omega_{orb}},$$

where J is the inertia tensor of NS, $\vec{\omega}$ is the absolute angular velocity of NS, $\overrightarrow{\omega_{orb}}$ is the orbital angular velocity of NS, $\overrightarrow{M^e}=\overrightarrow{M_a}+\overrightarrow{M_g}$ is the main moment of external forces, consisting of the gravitational [10] and aerodynamic moments.

The drag and aerodynamic moment dependencies on the attitude can be described analytically for CubeSats that have the shape of a rectangular parallelepiped [9]. Since the SamSat-Ionosphere has deployable external structures (Fig. 1), these characteristics are determined numerically.

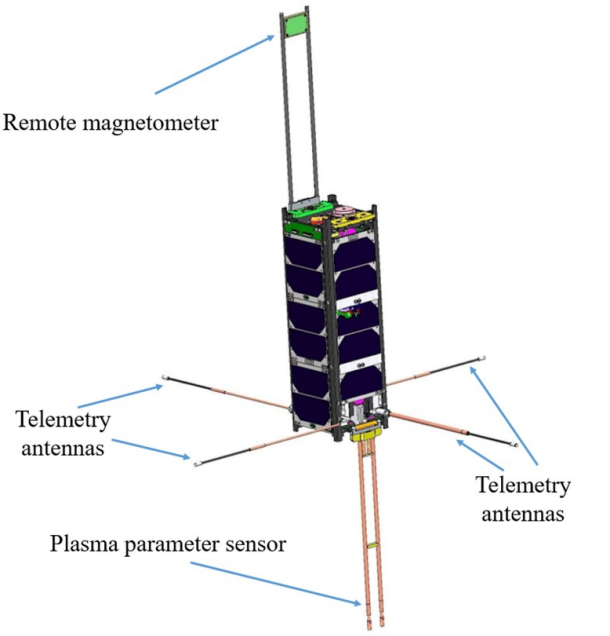


Fig. 1. SamSat-Ionosphere in flight configuration

III. Algorithm for Aerodynamic Characteristics Determination for NS with External Structures

The algorithm for AC determination was developed. It allows us to estimate AC for NS of arbitrary configuration. NS configuration is described by a triangular grid obtained from the computer-aided design system. The algorithm takes into account the free-molecular flow features such as particle thermal velocity and diffusive gas-surface interaction (GSI). The collision of particles is calculated according to [11]. The most probable thermal velocity of colliding particles is modeled according to the Maxwell distribution [12]. GSI model was adopted in accordance with the Cercignani-Lampis-Lord model [13].

The aerodynamic forces and moments are determined by the following relations:

$$\overrightarrow{F_j}=m_j\frac{\overrightarrow{V_{ij}}-\overrightarrow{V_{rj}}}{\Delta t},\quad \overrightarrow{M_{aj}}=\vec{r_j}\times\overrightarrow{F_j},$$

$$\vec{Q}=\sum_j\overrightarrow{F_j},\quad \overrightarrow{M_a}=\sum_j\overrightarrow{M_{aj}}.$$

Here, $\overrightarrow{V_{ij}}$, $\overrightarrow{V_{rj}}$ are the velocity vectors of the incident and reflected *j-th* particle, respectively; $\overrightarrow{F_j}$, $\overrightarrow{M_{aj}}$ are the force and moment created by the *j-th* particle, respectively; Δt is the time-step; m_j is the mass of the *j-th* particle; $\vec{r_j}$ is the radius-vector from the center of mass of NS to the point of particle collision.

As a result, the aerodynamic drag force and the aerodynamic moment were determined. The angles of attack α and slip angle β were used in the calculations. These angles are related to the Euler angles by the following relations:

$$\sin(\alpha)=\sin(\alpha_n)\cos(\varphi),$$

$$\cos(\beta)=\frac{\cos(\alpha_n)}{\cos(\alpha)},\quad \sin(\beta)=\frac{\sin(\alpha_n)\sin(\varphi)}{\cos(\alpha)}.$$

As an example, Fig. 2 shows the dependence of drag coefficient C_D on the angle of attack and slip angle, which is used in modelling of the NS center of mass motion.

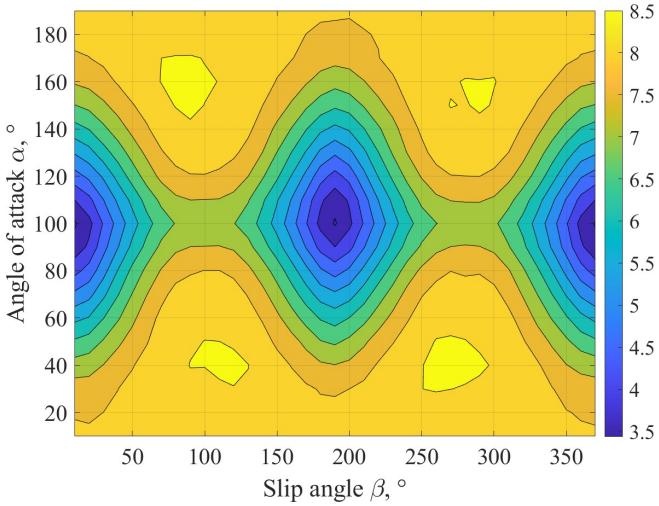


Fig. 2. Dependence of drag coefficient C_D on the angle of attack and slip angle

IV. Algorithm for Determination of Stable Equilibrium Positions for NS with External Structures

The algorithm for determination of the stable equilibrium positions can be divided into the following steps:

1. We create a set of initial attitudes that divides the domains of Euler angles $\alpha_n\in[0;\pi]$, $\varphi\in[0;2\pi]$, $\psi\in[0;2\pi]$ into equal local regions of search.

2. The search for stable equilibrium positions is performed by the trial attitude modelling in the local region with the selected initial position. Then, oscillations amplitudes are averaged over the entire modelling region for each angle.

3. If the averaged attitude does not belong to the considered local area, this array element is rejected and the algorithm proceeds to the next initial attitude.

4. If the averaged attitude belongs to the considered local region, it is used as the initial condition for the next modelling. Next, the stable equilibrium position is refined by successive averaging of the oscillation amplitudes until the desired accuracy is reached.

5. Since the adjacent regions overlap, it is possible to obtain the repeated results that are rejected.

As a result, all stable angular equilibrium positions of the NS are determined.

V. Determining Stable Equilibrium Positions of the SamSat-Ionosphere NS

As an example, using the developed algorithms, we determined the stable equilibrium positions for SamSat-Ionosphere which has the following characteristics:

$$m = 3.55\ kg;\ S = 0.017\ m^2;$$

$$J_x = 0.0099\ kg \cdot m^2;\ J_y = 0.0537\ kg \cdot m^2;\ J_z = 0.0522\ kg \cdot m^2;$$

$$J_{xy} = -0.0001\ kg \cdot m^2;\ J_{xz} = -0.0028\ kg \cdot m^2;\ J_{yz} = 0.00008\ kg \cdot m^2.$$

For the study we chose a range of altitudes from 300 to 700 km, since the aerodynamic and gravitational moments have a significant influence on the NS attitude dynamics.

We find only two stable equilibrium positions within the given altitude range using model of standard atmosphere [14]. Stable positions differ only by the angle ψ at $180°$. The solutions have an identical behaviour depending on the altitude or density function. It was also found that at altitudes above 650 km, there may be two additional stable equilibrium positions that differ from those given above by the angle φ at $\approx 180°$.

It is important to note that the navigation antenna will be pointing upward at $\psi \approx 0°$ and downward at $\psi \approx 180°$. Therefore, to provide the required orientation for the navigation antenna, the NS must be oriented relative to the stable equilibrium position at $\psi \approx 0°$, $\varphi \approx 6°$. For the considered equilibrium position, Fig. 3 shows the dependence of stable equilibrium position on the flight altitude.

Figure 3 shows that the stable equilibrium positions change within the following ranges: $\alpha_n \in [91.0°;\ 93.6°]$, $\varphi \in [4.8°;\ 6.6°]$ and $\psi \in [0.18°;\ 0.28°]$. It indicates a weak influence of altitude and atmospheric density on the stable equilibrium positions of the NS. This is due to the fact that the NS was designed to ensure passive three-axis gravitational-aerodynamic stabilization.

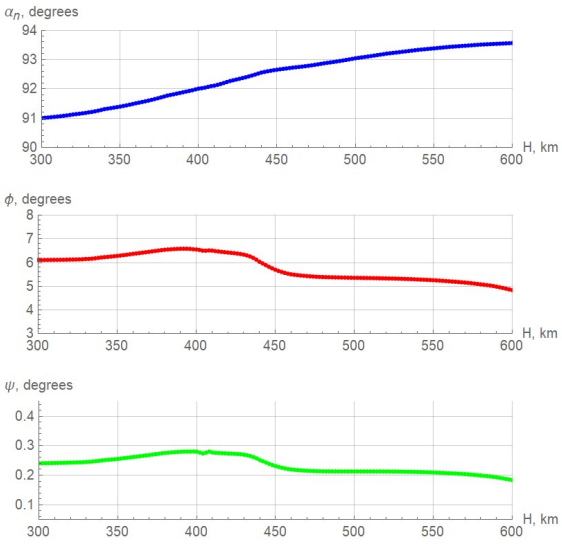


Fig. 3. Dependence of stable equilibrium position on the altitude

Table 1. Stable Equilibrium Positions of SamSat-Ionosphere for Different Atmospheric Models (Altitude is 500 km)

Model	ψ,°	φ,°	α_n,°
Without atmospheric influence	0.1	3.5	93.8
	180.1	3.5	93.8
	359.9	183.5	86.2
	179.9	183.5	86.2
Min. night	0.2	4.4	93.7
	180.2	4.4	93.7
	0.2	178.6	86.2
	180.2	178.6	86.2
Min. day	0.2	5.3	93.2
	180.2	5.3	93.2
Max. night	0.1	6.0	92.2
	180.1	6.0	92.2
Max. day	0.3	6.3	91.4
	180.3	6.3	91.4

Table 1 presents the stable equilibrium positions of SamSat-Ionosphere at 500 km altitude for different atmospheric models taking into consideration the day and night variations as well as solar flux index [15]. It can be seen that for all atmospheric models there are two stable equilibrium positions differing from each other by the angle ψ by $180°$. In case of minimum solar activity at night and without atmospheric influence, two additional stable equilibrium positions may exist. These positions differ from the two given above by the angle φ at $\approx 180°$. The existence of two additional stable equilibrium positions of NS shows that the influence of gravitational moment is dominant.

Thus, in this paper we developed the algorithms for determination of AC and stable equilibrium positions for NS with external structures. The algorithm search of stable positions can be useful in the development of stabilization and attitude control systems. The stable equilibrium positions determined in this work show that the aerodynamic and gravitational moments are restoring for the required NS attitude. This allows us to conclude that the passive three-axis gravitational-aerodynamic stabilization of SamSat-Ionosphere spacecraft has been successfully implemented.

REFERENCES

[1] State Corporation "Roscosmos" URL: http://www.roscosmos.ru/39352/ (accessed 05.11.2024).

[2] E.V. Barinova, I.V. Belokonov, and I.A. Timbai, "Technology for Designing the Angular Motion of CubeSat Nanosatellites with a Passive Stabilization System", Giroskopiya i Navigatsiya, 2024, vol.32, no. 4, pp. 346-355.

[3] V.A. Sarychev, and M.Y. Ovchinnikov, "Dynamics of a satellite with a passive aerodynamic orientation system", Cosmic Research, 1994, vol. 32, no. 6, pp. 561–575.

[4] V.A. Sarychev, "Investigation of equilibria of a satellite subjected to gravitational and aerodynamic torques", Celest. Mech. Dyn. Astron., 2007, vol, 97, no 4. pp. 267–287.

[5] V.A. Sarychev, and S.A. Gutnik, "Dynamics of a satellite subject to gravitational and aerodynamic torques. Investigation of equilibrium positions", Cosmic Research, 2015, vol. 53, no.65, pp. 449–457.

[6] V.A. Sarychev, and S.A. Gutnik, "Satellite dynamics under the influence of gravitational and aerodynamic torques. A study of stability of equilibrium positions", Cosmic Research, 2016, vol. 54, no 5, pp. 388–398.

[7] S.A. Barinova, and I.A. Timbai, "Study of Relative Equilibrium Positions of a Dynamically Symmetric Cubesat Nanosatellite under Aerodynamic and Gravitational Moments" in 26th Saint Petersburg International Conference on Integrated Navigation Systems. ICINS 2019 Proceedings, St. Petersburg. Concern CSRI Elektropribor, 2019.

[8] E.V. Barinova, and I.A. Timbai, "Determining of Equilibrium Positions of CubeSat Nanosatellite under the Influence of Aerodynamic and Gravitational Moments", in 27th Saint Petersburg International Conference on Integrated Navigation Systems. ICINS 2020 Proceedings, St. Petersburg. Concern CSRI Elektropribor, 2020.

[9] E.V. Barinova, I.A. Timbai, and Y. Mironov, "Numerical-analytical determining of equilibrium positions of CubeSat nanosatellite under aerodynamic and gravitational moments. the Influence of Aerodynamic and Gravitational Moments", in 31th Saint Petersburg International Conference on Integrated Navigation Systems. ICINS 2024 Proceedings, St. Petersburg. Concern CSRI Elektropribor, 2024, pp. 403-406.

[10] E.V. Barinova, I.V. Belokonov, and I.A. Timbai, "Preventing Resonant Motion Modes for Low-Altitude CubeSat Nanosatellites", Gyroscopy and Navigation, 2021, vol. 12, no. 4, pp. 350–362.

[11] Vas.V. Sazonov, "Construction of interactive geometric model of outside surface of a spacecraft", Matematicheskoe modelirovanie, 2020, vol. 32, no.6, pp. 37-52.

[12] G.A. Bird, Molecular Gas Dynamics and the Direct Simulations of Gas Flows. Oxford: Clarendon Press, 1994.

[13] R.G. Lord, "Application of the Cercignani-Lampis Scattering Kernel to Direct Simulation Monte Carlo Calculations. 17th International Symposium om Rarefied Gas Dynamics", Aachen, Rheinisch-Westfaelische Technische Hochschule Aachen, 1995, pp. 1427-1433.

[14] GOST (State Standard) 4401-81, Atmosfera standartnaya. Parametry (Standard Atmosphere. Parameters), Introd. 1981-02-27, Moscow: Izd. standartov, 1981.

[15] GOST (State Standard) 25645.101-83, Verkhnyaya atmosfera Zemli. Model' plotnosti dlya proektnykh ballisticheskikh raschetov iskusstvennykh sputnikov Zemli (Earth's upper atmosphere. Density model for design ballistic calculations of artificial Earth satellites), Moscow: Izd. standartov, 1983.

Ballistic Design of Spacecraft Inspection Missions in Circular Orbits in a Normal Gravitational Field

D.P. Avariaskin
Inter-University Space Research Department
Samara National Research University
Samara, Russia
avaryaskind@gmail.com

I.V. Belokonov
Inter-University Space Research Department
Samara National Research University
Samara, Russia
belokonov.iv@ssau.ru

M.S. Shcherbakov
Inter-University Space Research Department
Samara National Research University
Samara, Russia
sherbakov.m.s@mail.ru

Abstract – This study investigates the inspection motion of an inspector spacecraft (hereafter "inspector") relative to an object of inspection (hereafter "OI") in a non-central gravitational field, accounting for the second zonal harmonic of Earth's gravitational potential (hereafter "normal gravity field"). The inspection trajectory is defined as the inspector's fly-around trajectory in the Orbital Coordinate System (OCS) centered at the OI's mass center. Initial motion parameters of the inspector are selected based on the condition of equal orbital energies between the inspector and OI, which is derived under the perturbing influence of Earth's second zonal harmonic (J_2). The deformation of the inspection trajectory in the normal gravity field is investigated. Unlike previous studies, this work explores the applicability of previously derived expressions for selecting the inspector's initial motion parameters during inspection in the normal gravity field and provides recommendations for ballistic design of inspection missions.

Keywords: spacecraft formation flying, passive inspection motion, fly-around trajectory, orbital energy equality, normal gravity field, ballistic mission design.

I. Introduction

Since the early 21st century, the number of small and ultra-small spacecraft operating in near-Earth orbits has increased very much. From 2020 to 2024 alone, over 1,300 CubeSat-class spacecraft were launched into low Earth orbit [1]. A popular application of small satellites is formation flying, where coordinated operation of multiple spacecraft achieves mission objectives [2]. The type of formation depends on the mission's goals. This work focuses on inspection motion, where an inspector performs periodic fly-arounds of an OI in the same orbital plane. In addition to the spacecraft that are part of the orbital group, a space debris can also be an OI; its examination allows us to obtain the necessary information about its orbital motion, geometric parameters, etc.

Circular orbits at 500–1000 km altitudes are among the most common for small satellites. At these altitudes, inspection trajectory deformation is driven by differential gravitational acceleration caused by Earth's non-spherical gravity. Studies [3–5] demonstrate that it is possible to significantly reduce the influence of differential gravitational acceleration by selecting the initial motion parameters at the moment of formation of the inspection trajectory in conditions of the inspector and the OI motion in a non-central gravitational field taking into account the second zonal harmonic J_2 (normal gravity field). This feature underscores the importance of ballistic mission design — selecting initial inspector velocity to achieve desired inspection trajectories.

II. Problem Statement

The inspector's motion relative to the OI is analyzed in the Orbital Coordinate System (OCS). The OCS origin coincides with the OI's center of mass; the *Oy*-axis aligns with the OI's radius vector; the *Ox*-axis points along the orbital velocity's normal component; and the *Oz*-axis completes the right-handed system perpendicular to the orbital plane. The study assumes coplanar motion and a circular initial orbit for the OI. The task of an inspector flying into the vicinity of the OI is not considered.

For successful inspection, the inspector's relative trajectory must be closed. In an unperturbed (central gravity field) motion, this requires equal orbital energies — i.e., identical semi-major axes and orbital periods for both spacecraft. When perturbations (e.g., higher-order gravity harmonics) are introduced, the relative trajectory evolves and loses closure.

The primary perturbation considered arises from the normal gravity field. Figure 1 shows the nominal (red) and perturbed (blue) inspection trajectories.

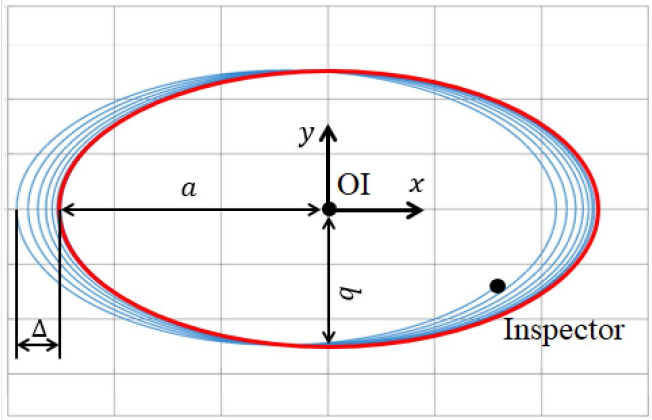


Fig. 1. . Nominal (red) and perturbed (blue) inspection trajectories

The evaluation of the inspection trajectory deformation relative to the nominal one is determined by the formula:

$$\bar{\Delta} = \frac{\Delta}{a} \cdot 100\% \qquad (1)$$

In the general formulation, the problem of ballistic design of inspection trajectories consists of choosing the projections of the initial velocity of the inspector in the OCS ($\dot{x}_0$, $\dot{y}_0$),providing the required type of relative trajectory for given initial positions of the inspector in the OCS (x_0, y_0) and the IO in its orbit (determined by the latitude argument θ_0^{oi}).

To find the projections of the initial velocity of the inspector in the OSC ($\dot{x}_0$, $\dot{y}_0$) in the plane of the orbit of the IO, a system of equations is used [6]:

$$\begin{cases} \left(V_0^i\right)^2 = \left(\dot{x}_0 + \dot{\theta}_0^{oi} y_0 + Vn_0^{oi}\right)^2 + \left(\dot{y}_0 - \dot{\theta}_0^{oi} x_0 + Vr_0^{oi}\right)^2 \\ \dot{y}_0 = \dfrac{x_0 \dot{\theta}_0^{oi}}{2} \end{cases} \quad (2)$$

where $\left(V_0^i\right)^2$ – initial inspector velocity in the Earth-centered inertial (ECI) coordinate system; $\dot{\theta}_0^{oi}$ – initial OI orbital angular velocity; Vn_0^{oi}, Vr_0^{oi} – normal and radial projections of the OI initial velocity in ECI.

To ensure a closed trajectory of the inspector relative to the OI, it is necessary to select the desired initial velocity V_0^i and substitute it into (2). In this case, if we substitute into (2) such an orbital velocity of the inspector V_0^i, that will ensure a closed trajectory and equality of orbital energies, then this closed trajectory will have a shape close to an ellipse, whose center is at the beginning of the OCS and the major semi-axis is twice as large as the minor one. In the case of unperturbed motion, such a trajectory remains closed and it was chosen as the nominal one.

We write the equality of the orbital energies of the OI and the inspector in the following form:

$$\begin{aligned} &\frac{\left(V_0^{oi}\right)^2}{2} + \frac{\mu}{R_0^{oi}} - \frac{\mu J_2 R_\text{Э}^2}{2\left(R_0^{oi}\right)^3} \cdot \left(3\sin^2 i_0^{oi} \cdot 3\sin^2 \theta_0^{oi} - 1\right) = \\ &= \frac{\left(V_0^i\right)^2}{2} + \frac{\mu}{R_0^i} - \frac{\mu J_2 R_\text{Э}^2}{2\left(R_0^i\right)^3} \cdot \left(\frac{3\left(Z_0^i\right)^2}{\left(R_0^i\right)^2} - 1\right) \end{aligned} \quad (3)$$

where μ – gravitational parameter of the Earth; J_2 – coefficient of the second zonal harmonic; $R_\text{Э}$ – radius of the Earth's equator; V_0^i – initial orbital velocity of the OI; R_0^{oi} – initial radius vector of the OI; i_0^{oi} – inclination of the OI orbit; R_0^i, Z_0^i – value of the initial radius vector and Z coordinate of the inspector in ECI.

From the obtained expression, the orbital velocity of the inspector is found:

$$\begin{aligned} \left(V_0^i\right)^2 &= \left(V_0^{oi}\right)^2 + \frac{2\mu}{R_0^i} - \frac{2\mu}{R_0^{oi}} + \\ &+ \mu J_2 R_\text{Э}^2 \left(\frac{3\sin^2 i_0^{oi} \cdot 3\sin^2 \theta_0^{oi}}{\left(R_0^{oi}\right)^3} - \frac{1}{\left(R_0^{oi}\right)^3} - \frac{1}{\left(R_0^i\right)^3} - \frac{3\left(Z_0^i\right)^2}{\left(R_0^i\right)^5}\right) \end{aligned} \quad (3)$$

By solving (2) and (3) together, we determine the projections of the inspector's velocity in the OSC ($\dot{x}_0$, $\dot{y}_0$), at which the condition of equality of orbital energies will be satisfied.

However, even when choosing the initial conditions from (2) and (3) when modeling the motion taking into account the normal gravitational field, the relative trajectory is not closed and has a small deformation along the *Ox* axis in OCS over a long time interval. The magnitude of the deformation depends on the orbital inclination, the initial argument of the latitude of the OI and the initial position of the inspector in the OSC.

In connection with this, the solution to the problem of ballistic design of inspection motion is the choice of such initial parameters of the inspector relative motion at which the influence of disturbances will be minimal. This provides the long-term inspection motion. In this paper, we consider the time interval of the inspection mission equal to 120 days, after which the percentage of deformation of the inspection trajectory relative to the nominal one is determined.

III. NUMERICAL SIMULATION

For numerical simulation, the orbits of the OI with an altitude of 500 km and inclinations of 51.6° and 90° were selected as an example. Figure 2 shows the dependences of the maximum deformation of the disturbed relative trajectory (the disturbance described by the second zonal harmonic) on the OI argument of latitude for different initial positions of the inspector.

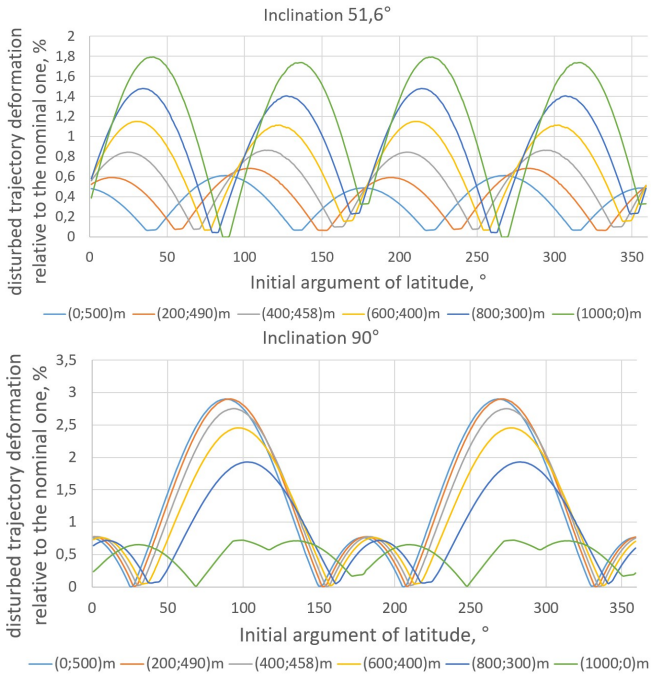


Fig. 2. Dependence of the disturbed trajectory deformation relative to the nominal one on the OI argument of latitude in normal gravitational field

In Figure 2, the colors of the lines correspond to different positions of the inspector at the beginning of the inspection motion formation in the OSC. The initial positions of the inspector were selected on the nominal trajectory with a major semi-axis equal to 1000 m, and a minor semi-axis equal to 500 m. From the graphs in Figure 2, it is clear that the maximum deformation of the disturbed trajectory from the nominal one for these initial conditions will not exceed 3% in 120 days.

To analyze the sensitivity of the obtained results to the adopted model of the gravitational field, a study was conducted of the fourth zonal harmonic of the Earth's gravitational field influence on the inspection trajectory as the next most important. In this case, the initial conditions of the motion were selected similarly to the previous example. The results are shown in Figure 3.

As can be seen from Figure 3, the effect of taking into account the fourth zonal harmonic increases the maximum deformation of the relative trajectory (Figure 2). Analysis of the obtained results allows us to conclude that it is possible to select such initial conditions for the motion of the pair "OI + inspector" that will provide inspection motion with al zero deviation.

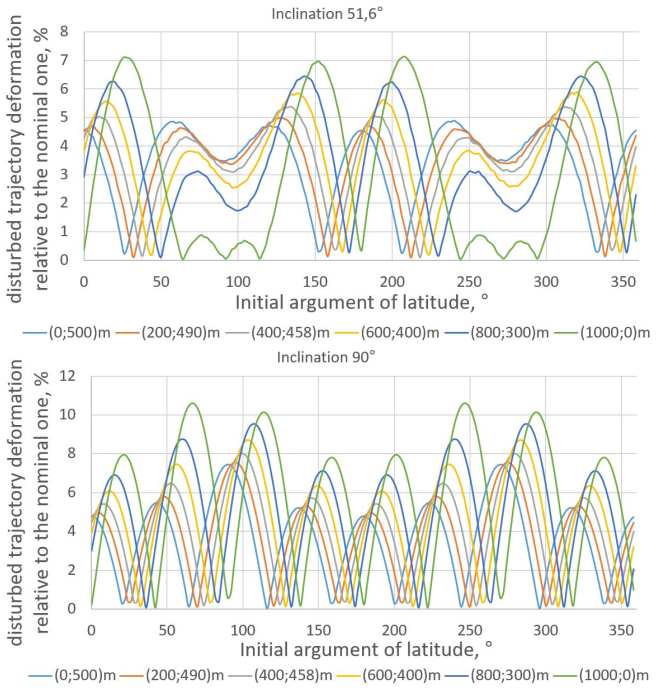


Fig. 3. Dependence of the disturbed trajectory deformation relative to the nominal one on the OI argument of latitude (with J_4)

Based on the results of the study, the following sequence of ballistic design stages of inspection missions in circular orbits is proposed. In this case, the initial parameters are the inclination and altitude of the OI orbit, as well as the parameters of the nominal inspection trajectory.

At the first stage, for the specified dimensions of the nominal inspection orbit, the initial conditions of the OI motion are selected in accordance with relations (2) and (3) and the dependences of the disturbed trajectory deformation on the nominal one are calculated taking into account the second and fourth zonal harmonics.

At the second stage, based on the obtained dependencies, the values of the arguments of latitude and position of the inspector are determined, at which the deformation will be minimal for a given duration of the inspection mission.

An inspection mission around the Russian Orbital Service Station (ROSS) is considered as an example. It is assumed that the ROSS will have a near-circular orbit with an altitude of 500 km and an inclination of 97°. An ellipse with a major semi-axis equal to 1000 m is adopted as the inspection trajectory, and, accordingly, the minor semi-axis will be 500 m.

Figure 4 shows the dependences of the deformation of the disturbed trajectory relative to the nominal one.

Next, for different positions of the inspector, the initial arguments of the ROSS latitude are selected, on which the inspection trajectory will be formed, which will subsequently lead to a deformation of no more than 0.5% (5 m along the *Ox* axis in OCS) relative to the nominal trajectory over a period of 120 days. For example, if to start motion on such an inspection trajectory, when the inspector is on the same radius vector with the ROSS (in this case, at a point with coordinates (0; 500) m), then it is necessary that ROSS should have argument of latitude equal to 22°, 64°, 116°, 158°, 206°, 244°, 296°, 338°. Moreover, at the moment of forming the inspection trajectory, with an ROSS' argument of latitude of 22°, the inspector relative to the ROSS must have a speed with components $\dot{x}_0 = -1.11$ м/с, $\dot{y}_0 = 0$ м/с.

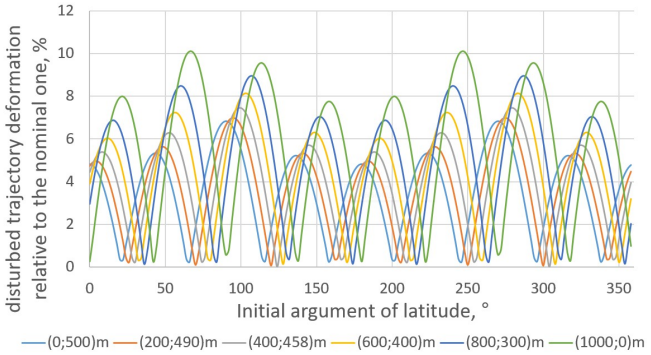


Fig. 4. Dependence of the disturbed trajectory deformation relative to the nominal one on the ROSS argument of latitude (with J_4)

IV. Conclusion

The work demonstrates the possibility of ensuring long-term passive inspection motion in circular orbits by selecting the initial motion conditions of the pair "OI+inspector" taking into account the second and fourth zonal harmonics of the gravitational field.

ACKNOWLEDGMENT

The work is a continuation of the research carried out within the framework of the completed project 0777-2020-0018, financed from the state assignment to the winners of the competition of scientific laboratories of higher education organizations subordinate to the Ministry of Education and Science of Russia.

REFERENCES

[1] Nanosat datebase https://www.nanosats.eu/#figures

[2] D. V. Naumochkin, A. I. Petukhov, M .M. Poluyan "Analysis of trends in the development of ultra-small spacecraft", Armament and economy, 2019, #4 (50), pp. 37-43.

[3] Belokonov I.V., Shcherbakov M.S., Avaryaskin D.P. "Investigation of a Single-Axis Control Algorithm for the Inspection Motion of a Gravitationally Stabilized Nanosatellite", 29th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2022.

[4] Belokonov I.V., Shcherbakov M.S., Medvedev S.A. "Selection of Conditions for Ensuring Long-Term Passive Inspection in the Close Vicinity of a Spacecraft in a Highly Elliptical Orbit", 30th Anniversary Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2023.

[5] Shcherbakov M.S., Medvedev S.A. Study of the possibility of using osculating ellipses of relative motion in the task of inspection of space objects // Materials of FGUP "NPTsAP". Sistemy I pribory upravleniya. — 2023. — # 2 (64). — P. 42-50.

[6] Shcherbakov M.S., Avariaskin D.P. "Studying problems on choosing stable orbits of nanosatellites to provide passive and periodic relative trajectories", J. Phys.: Conf. Ser. 1536 012001, 2020.

SESSION II – INERTIAL SYSTEMS AND SENSORS

Solid-State Laser Gyroscope Based on YAG:Cr^{4+}, Features of Mode-Locking with KTP Phase Modulator Frequency Biasing

Yu.Yu. Broslavets
MIPT
JSC «Lasex»
Dolgoprudny, Russia
laseruu@mail.ru

E.A. Polukeev
MIPT
JSC «Lasex»
Dolgoprudny, Russia
epolukeev2105@yandex.ru

Yu.N. Travina
MIPT
Dolgoprudny, Russia
travina.iun@phystech.edu

A.A. Fomichev
MIPT
JSC «Lasex»
Dolgoprudny, Russia
laser@mail.mipt.ru

V.G. Semenov
MIPT
JSC «Lasex»
Dolgoprudny, Russia
valerii.semenov.g@gmail.com

***Abstract* — This paper presents the results of a study on mode locking in a YAG:Cr^{4+} solid-state laser gyroscope with frequency biasing achieved by alternating out-of-phase shifts of the counter-propagating pulses using a KTP-based electro-optic phase modulator. Synchronization of longitudinal modes reduces noise in each mode and improves the beat signal of the counter-propagating waves, enhancing accuracy. Implementing frequency biasing with the phase modulator required optimization of resonator parameters to simultaneously minimize group delay dispersion, reduce mode distortion and achieve the required frequency splitting of the counter-propagating waves. As a result, a resonator configuration was found, enabling both Kerr-induced mode-locking and weak radiation response to external perturbations. The influence of the KTP phase modulator induced frequency bias during mode-locking was simulated by using equation for the complex amplitudes of the counter-propagating waves, each represented by four longitudinal modes. It was found, that the bias amplitude is required to be smaller than the longitudinal mode coupling coefficient. The feasibility of generating a frequency bias of several hundred kilohertz was demonstrated.**

Keywords — solid-state laser gyroscope, ring laser, mode locking, frequency lock-in.

I. INTRODUCTION

He-Ne laser gyroscopes [1] are widely used in strapdown inertial navigation systems (SINS). However, their accuracy is approaching fundamental limits. Emerging technologies and novel laser crystals [2-8, 12, 17] with numerous advantages have intensified the pursuit of developing laser gyroscopes [3, 9, 10, 12-14] based on these new media. Longitudinal mode synchronization in solid-state laser gyroscopes operating in a multi-frequency lasing regime can significantly enhance accuracy and mitigate the fundamental precision limit of He-Ne laser gyroscopes, which is associated with quantum noise [16]. In a YAG:Cr^{4+} laser gyroscope with a broadband gain medium the number of synchronized modes can reach ~$2.7*10^5$ [17]. Studies indicate that noise in each mode – both fundamental (intrinsic to the laser) and technical (related to external perturbations) – can be substantially reduced proportionally to the number of synchronized modes [11, 12, 15-17]. Additionally, advancements in mirror fabrication, ultra-fine surface polishing, novel gain media and high-speed electronics have made the development of high-precision solid-state laser gyroscopes both feasible and relevant [10, 13]. The accuracy of such gyroscope is determined by several factors: lasing parameters, frequency lock-in range (characterizing the degree to which adjacent modes are captured by the modulation sidebands), lasing dynamics, scattering radiation affecting counter-propagating waves, resonator stability and other parameters. To achieve high accuracy, this work employs several solutions: unique properties of YAG:Cr^{4+} active centers [6-8, 17], which reduce competition between counter-propagating waves; mode synchronization to suppress noise; cavity length stabilization system; frequency biasing via a phase modulator. In the YAG:Cr^{4+} crystal, tetravalent chromium ions (Fig. 21) replace aluminium ions in the cubic lattice. These ions are spatially surrounded by oxygen ions and divalent magnesium or calcium ions, which compensate for the charge imbalance during the transition of trivalent chromium ions to tetravalent states. This configuration creates distinct amplification conditions for counter-propagating waves by different groups of active centers. The broad gain linewidth (~350 nm), attributed to strong electron-phonon interaction, enables emission across a wide

spectrum with a large number of longitudinal modes. Synchronizing these modes suppresses competition and coupling between counter-propagating waves, as well as frequency fluctuations, thereby narrowing the lasing spectrum of each mode. The short lifetime of the upper working level (3.6 μs) [7] significantly weakens the formation of inversion grating that induce interaction between counter-propagating waves. These characteristics distinguish YAG:Cr^{4+} from other gain media, ensuring stable bidirectional lasing in laser gyroscopes [12, 17].

This paper presents the results of developing a laser gyroscope based on a YAG:Cr^{4+} gain medium. The solid-state laser gyroscope was built as a ring laser comprising a four-mirror resonator, a YAG:Cr^{4+} gain medium and a dispersive prism made of infrared (IR) quartz with ultra-low impurity content [17]. A KTP crystal-based phase modulator was utilized to implement frequency biasing during mode locking (Fig. 14). This modulator exhibits significantly lower noise caused by piezoelectric "ringing" compared to lithium niobate modulators. The modulator consisted of a Brewster-angle-mounted plate with an optical path length of 1.5 mm (excluding the refractive index). Amplifying, delaying, filtering and further amplifying pulses initially detected at the ring laser output by a high-speed InGaAs photodetector generated an alternating voltage applied to plate. This voltage induced refractive index changes in the modulator with opposite signs during the out-of-phase passage of counter-propagating radiation pulses, thereby creating frequency splitting between the counter-propagating waves. The out-of-phase propagating radiation pulses were generated in the counter-propagating waves.

Additionally, the work analyses the influence of the physical principles used for frequency biasing and the frequency stabilization method on the noise and stability of the beat frequency between counter-propagating waves during Kerr-induced mode locking. Optimization of the optical scheme parameters was performed to achieve the required phase modulation for frequency biasing. This included determining the orientation of the KTP crystal, accounting for its Brewster-angle alignment for incoming and outgoing radiation pulses, and defining the modulator's position in the resonator relative to the gain medium. An acetylene-filled absorption cell was also utilized to stabilize the lasing frequency during mode locking.

Numerical modelling of mode synchronization, based on equations for the complex amplitudes of counter-propagating waves generated in the multi-frequency regime, enabled the determination of optimal longitudinal mode coupling parameters. These parameters are defined by the Kerr-induced synchronization mechanism and ensure maximal narrowing of each mode linewidth during synchronization, accounting for the influence of mode-specific noise. As a result, the noise in the beat signal of the counter-propagating waves was significantly reduced.

II. Determining lasing parameters for maximal noise suppression in each mode during multi-frequency generation with mode locking

Mode locking in a laser is known to supress noise and fluctuations [11, 12, 15]. In a solid-state laser gyroscope, the ring resonator facilitates the lasing of a large number of longitudinal modes in two counter-propagating directions, resulting in highly complex coupling between all waves and lasing dynamics influenced by the amplification characteristics of the Cr^{4+} active centers in the YAG:Cr^{4+} gain medium. In such a gyroscope, wave coupling via scattering on optical elements, inversion gratings in the gain medium, and mode locking mechanisms can involve varying phase shifts. This significantly complicates the overall wave interaction and necessitates laser parameter optimization to minimize fluctuations and spectral linewidth of the emitted radiation of each mode. This is critical for achieving the highest accuracy in the laser gyroscope. Frequency and amplitude fluctuations in the modes lead to fluctuations in the beat signal of counter-propagating waves and degrade gyroscope precision.

For modelling and calculating lasing parameters, a system of equations (4) was derived by solving the wave equation and density matrix equations [2, 9, 17]. The following assumptions were used: radiation was represented as plane waves; spatial homogeneity in transverse field distribution was assumed; diffraction nonreciprocity was neglected; linear polarization and uncorrelated noise in longitudinal modes were considered [15, 16]. The total field in the ring laser was modelled as a superposition of counter-propagating wave fields, each expressed as an expansion over the resonator's longitudinal modes:

$$\vec{E}(z,t) = \\ = \mathrm{Re}\sum_j \left(e_{j1}\tilde{E}_{j1}(t)\exp\left(i\left(\omega_{j1}t + k_{j1}z\right)\right) + e_{j2}\tilde{E}_{j2}(t)\exp\left(i\left(\omega_{j2}t - k_{j2}z\right)\right)\right), \quad (1)$$

The complex amplitudes of the counter-propagating waves in (1) can be represented through their amplitude and phase:

$$\tilde{E}_{j1}(t) = E_{j1}\exp\left(i\varphi_{j1}\right), \quad \tilde{E}_{j2}(t) = E_{j2}\exp\left(i\varphi_{j2}\right). \quad (2)$$

In this case, the frequency difference between counter-propagating waves for each longitudinal mode j can be derived from the time-varying phases of the fields:

$$\Delta\omega_j = \omega_{j1} - \omega_{j2} = \frac{d\varphi_{j1}(t)}{dt} - \frac{d\varphi_{j2}(t)}{dt}. \quad (3)$$

Nonreciprocity for counter-propagating waves, caused by the Sagnac effect and other phenomena, is introduced via Ω without accounting for frequency splitting differences between counter-propagating waves for distinct longitudinal modes. Numerical simulations based on system (4) were performed for four modes in each direction j=1..4. In equations (4), the intermode frequency interval was defined as Ω_{ml}, resulting in curves depicting the frequency locking range of longitudinal modes during mode locking. The simulations assumed $\Omega_{ml} << \omega_{ij}$. Additionally, inversion gratings were modelled to account for interaction only in one pair of counter-propagating waves, corresponding to frequency split of a single longitudinal mode. Mode locking was implemented via loss modulation at the intermode frequency. The resulting sidebands at adjacent mode frequencies induced mode coupling and phase locking. Coupling between modes in the same direction was introduced through coefficients $\tilde{m}l_{jm} = K_{jm}\delta$, determined by the modulation depth δ, where K_{jm} are complex coefficients. Mode amplitude modulation followed $U=U_0(1 - \delta\cos(\omega_m t))$, where ω_m is the modulation frequency.

The magnitude of this coupling $\tilde{m}l_{jm}$ determines the phase locking range of adjacent modes, the stability of mode locking and its robustness to phase perturbations. Mode locking generates femtosecond pulses, but the longitudinal mode phase alignment process is relatively slow; thus, femtosecond timescales were neglected in the equations. Inversion gratings do not form when counter-propagating pulses traverse the gain medium at different times due to the absence of counter-propagating radiation overlap. Therefore, in this regime, the amplitudes of spatial inversion harmonics were considered negligible.

In summary, the truncated equations for the complex amplitudes [17, 8, 16] of counter-propagating waves and population inversion under multi-frequency lasing and mode locking in a solid-state laser gyroscope can be expressed as follows:

$$\frac{d\tilde{E}_{11}}{dt}=-\frac{\omega_{11}}{2Q_{11}}\tilde{E}_{11}+i\frac{\Omega}{2}\tilde{E}_{11}+i\frac{\Omega_{ml}}{2}\tilde{E}_{11}+\frac{i}{2}\tilde{m}_{11}\tilde{E}_{11}+\frac{i}{2}\tilde{m}l_{11}\tilde{E}_{21}+i\Delta\omega_r\tilde{E}_{11}+\frac{\sigma_1 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{11}+N_-\tilde{E}_{11}\right)$$

$$\frac{d\tilde{E}_{12}}{dt}=-\frac{\omega_{12}}{2Q_{12}}\tilde{E}_{12}-i\frac{\Omega}{2}\tilde{E}_{12}+i\frac{\Omega_{ml}}{2}\tilde{E}_{12}+\frac{i}{2}\tilde{m}_{12}\tilde{E}_{12}+\frac{i}{2}\tilde{m}l_{12}\tilde{E}_{22}+i\Delta\omega_r\tilde{E}_{12}+\frac{\sigma_1 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{12}+N_+\tilde{E}_{12}\right)$$

$$T_1\frac{dN_{1+}}{dt}=-N_{1+}\left(1+a_1\left(|E_{11}|^2+|E_{12}|^2\right)\right)-N_0a_1E_{11}^*E_{12}$$

$$T_1\frac{dN_0}{dt}=N_{th}(1+\eta)-N_0\left(1+a_1\left(|E_{11}|^2+|E_{12}|^2\right)+a_2\left(|E_{21}|^2+|E_{22}|^2\right)+\ldots+a_n\left(|E_{n1}|^2+|E_{n2}|^2\right)\right)$$

$$\frac{d\tilde{E}_{21}}{dt}=-\frac{\omega_{21}}{2Q_{21}}\tilde{E}_{21}+i\frac{\Omega}{2}\tilde{E}_{21}-i\frac{\Omega_{ml}}{2}\tilde{E}_{21}+\frac{i}{2}\tilde{m}_{21}\tilde{E}_{21}+\frac{i}{2}\tilde{m}l_{21}\left(\tilde{E}_{11}+\tilde{E}_{31}\right)+i\Delta\omega_r\tilde{E}_{21}+\frac{\sigma_2 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{21}+N_-\tilde{E}_{21}\right)$$

$$\frac{d\tilde{E}_{22}}{dt}=-\frac{\omega_{22}}{2Q_{22}}\tilde{E}_{22}-i\frac{\Omega}{2}\tilde{E}_{22}-i\frac{\Omega_{ml}}{2}\tilde{E}_{22}+\frac{i}{2}\tilde{m}_{22}\tilde{E}_{22}+\frac{i}{2}\tilde{m}l_{22}\left(\tilde{E}_{12}+\tilde{E}_{32}\right)+i\Delta\omega_r\tilde{E}_{22}+\frac{\sigma_2 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{22}+N_+\tilde{E}_{22}\right)$$

$$T_1\frac{dN_{2+}}{dt}=-N_{2+}\left(1+a_2\left(|E_{21}|^2+|E_{22}|^2\right)\right)-N_0a_2E_{21}^*E_{22}$$

...

$$\frac{d\tilde{E}_{n1}}{dt}=-\frac{\omega_{n1}}{2Q_{n1}}\tilde{E}_{n1}+i\frac{\Omega}{2}\tilde{E}_{n1}-i\frac{\Omega_{ml}}{2}\tilde{E}_{n1}+\frac{i}{2}\tilde{m}_{n1}\tilde{E}_{n1}+\frac{i}{2}\tilde{m}l_{n1}\tilde{E}_{(n-1)1}+i\Delta\omega_r\tilde{E}_{n1}+\frac{\sigma_n l}{2T}(1-i\delta)\left(N_0\tilde{E}_{n1}+N_-\tilde{E}_{n1}\right)$$

$$\frac{d\tilde{E}_{n2}}{dt}=-\frac{\omega_{n2}}{2Q_{n2}}\tilde{E}_{n2}-i\frac{\Omega}{2}\tilde{E}_{n2}-i\frac{\Omega_{ml}}{2}\tilde{E}_{n2}+\frac{i}{2}\tilde{m}_{n2}\tilde{E}_{n2}+\frac{i}{2}\tilde{m}l_{n2}\tilde{E}_{(n-1)2}+i\Delta\omega_r\tilde{E}_{n2}+\frac{\sigma_n l}{2T}(1-i\delta)\left(N_0\tilde{E}_{n2}+N_+\tilde{E}_{n2}\right)$$

$$T_1\frac{dN_{n+}}{dt}=-N_{n+}\left(1+a_n\left(|E_{n1}|^2+|E_{n2}|^2\right)\right)-N_0a_nE_{n1}^*E_{n2}$$

(4)

Here j denotes the longitudinal mode index, $\tilde{m}_{jm}=m_{jm}\exp\left(\pm i\Theta_{jm}\right)$ are complex coupling coefficients for counter-propagating longitudinal modes, $\tilde{m}l_{jm}=ml_{jm}\exp\left(\pm i\Psi_{jm}\right)$ are complex coupling coefficients for longitudinal modes, T_1 is the population inversion relaxation time, $\Delta\omega_r$ is a complex coefficient representing random phase and amplitude fluctuations, N_{th} is the threshold population inversion, L is the cavity length, l is the crystal length, T is the resonator round-trip time, $\delta_j=(\omega_j-\omega_0)/\Delta\omega_g$ is relative detuning of the j mode frequency from the gain line center, $\sigma_j=\sigma/(1+\delta_j^2)$ is the transition cross-section, which varies for different longitudinal modes, $a_j=(\sigma_j cT_1)/(4h\omega)$ is the saturation parameter, $\Omega=\omega_{j1}-\omega_{j2}$ is the frequency difference between counter-propagating waves, $\Omega_{ml}=\omega_{j+1\,m}-\omega_{jm}$ is the frequency difference between longitudinal modes, Q_j is the resonator quality factor, $k_j=2\pi n_j/L$ is the wave number, n_j is the number of wavelengths along the resonator length, $N_0=\frac{1}{l}\int_0^l N dz$, $N_{j\pm}=\frac{1}{l}\int_0^l N\exp(\pm i2kz)dz$ are spatial harmonics of the population inversion [2, 9, 17].

To achieve mode locking, the coupling between co-propagating waves was chosen to exceed the coupling between counter-propagating waves caused by scattering. This creates frequency locking regions (Fig. 1) for longitudinal modes, ensuring their phase alignment. The coupling magnitude is determined by the entry point into the lock-in region for mode synchronization (Fig. 1, $6.5*10^5$ rad/s) and for counter-propagating wave frequency locking (Fig. 4, lock-in $2.5*10^3$ rad/s; Fig. 5 lock-in $1.7*10^3$ rad/s). Consequently, the frequency bias amplitude Ω needs to be smaller than the entry threshold for longitudinal mode locking and larger than that for counter-propagating wave locking.

Numerical simulations based on system (4) revealed that the noise spectra of individual modes (Fig. 2) narrow during mode locking (Fig. 3). Noise for each mode was modelled using random complex coefficients $\Delta\omega_r$ with a normal distribution.

Notably, reducing counter-propagating wave coupling and narrowing the lock-in range during mode synchronization critically depend on the optical scheme design. The scheme needs to ensure alternating passage of counter-propagating pulses through scattering elements in the resonator. If this condition is unmet, the coupling reduction between counter-propagating waves remains minimal (Figs. 4, 5). Simulations showed that without mode locking, amplitude and phase modulation occur in each wave due to scattering-induced coupling (Fig. 6).

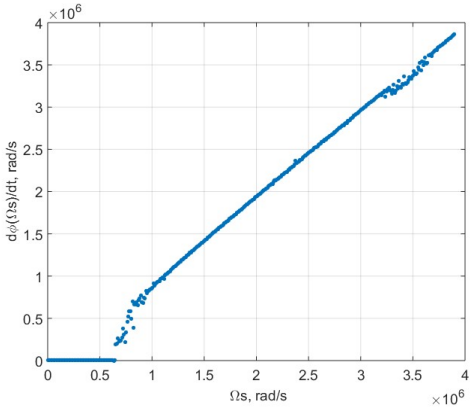


Fig. 1. Frequency difference between two adjacent modes versus resonator detuning (offset of the intermode interval from the loss modulation frequency). The locking region for co-propagating modes is observed. A larger locking range improves suppression of perturbations disrupting mode synchronization.

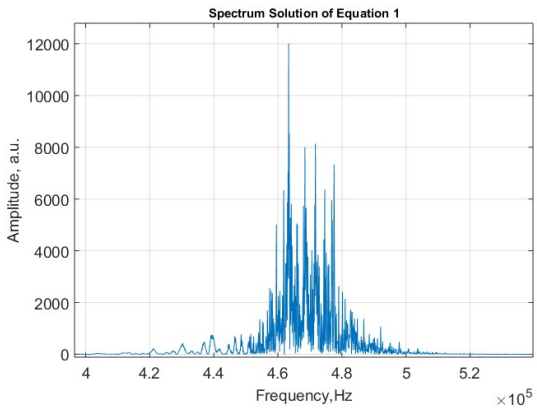


Fig. 2. Noise spectrum (amplitude and phase) of one wave, shifted to the low-frequency region, for **unsynchronized** longitudinal modes.

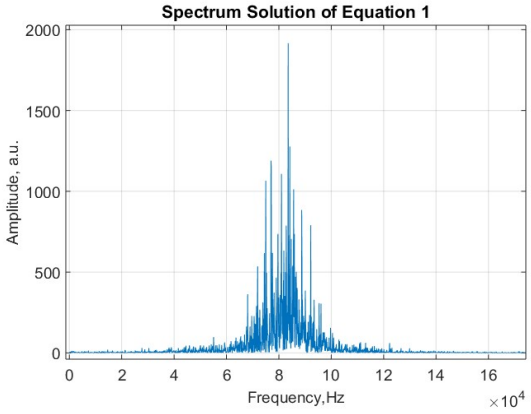


Fig. 3. Noise spectrum (amplitude and phase) of one wave, shifted to the low-frequency region, for phase-aligned longitudinal modes **during mode locking**.

In the YAG:Cr^{4+} solid-state laser gyroscope, frequency biasing was implemented using a KTP-based electro-optic phase modulator in the resonator. The modulator applied a phase shift of one sign synchronously with pulses traveling in one direction and of an opposite sign for pulses in the reverse direction. To create an alternating frequency bias, the polarity of the applied voltage relative to the counter-propagating pulses was switched at the bias reversal frequency. This was achieved by shifting the modulator's control signal phase by π. In simulations of frequency biasing using system (4), the frequency shift Ω for counter-propagating waves was induced by opposite-sign phase shifts in the electro-optic modulator.

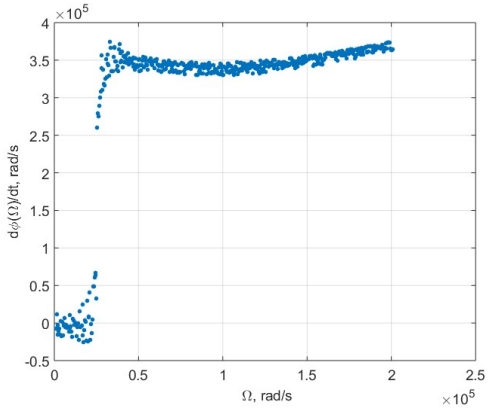


Fig. 4. Frequency response near the lock-in region for two counter-propagating waves with four **unsynchronized longitudinal modes**.

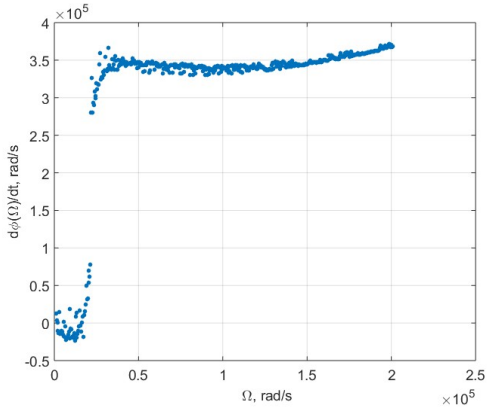


Fig. 5. Frequency response near the lock-in region for two counter-propagating waves with four **phase-aligned longitudinal modes**.

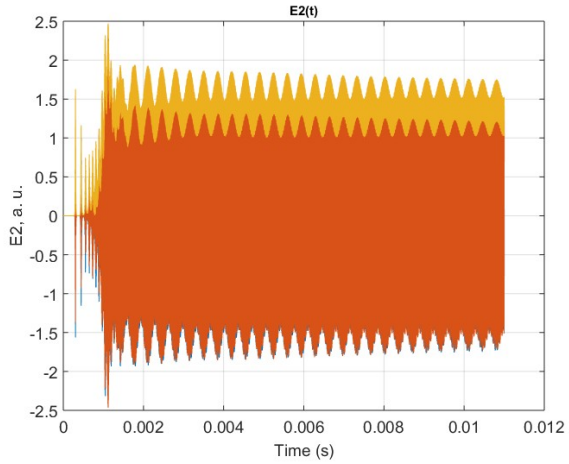


Fig. 6. Evolution of the complex amplitude E_2 (real, imaginary, and magnitude components of one wave) after pump activation. Transient oscillations in amplitude due to mirror scattering and inversion grating effects are observed. **Longitudinal modes are unsynchronized**.

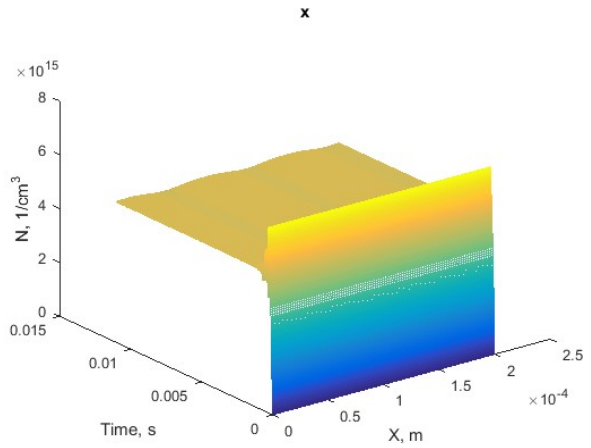


Fig. 7. Population inversion dynamics after pump activation (time and spatial distribution along the active medium) **without mode locking**. Low-amplitude inversion gratings emerge.

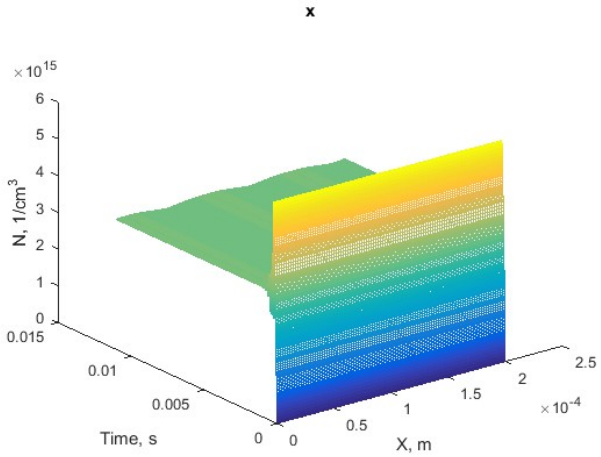


Fig. 8. Population inversion dynamics after pump activation (time and spatial distribution) **during mode locking**. The stationary inversion level decreases, matching the threshold at steady-state generation, indicating lower losses compared to the unsynchronized mode regime.

Studies revealed that Kerr-induced mode locking – where pulses traverse the Kerr medium, undergo self-focusing, and pass through a limiting aperture – reduces pulse duration and induces amplitude modulation. This generates sidebands that

phase-lock adjacent modes, significantly suppressing noise and narrowing its spectral width in each mode (Figs. 2, 3). The lock-in range (Fig. 1) of adjacent modes via sideband coupling determines the robustness of phase alignment to perturbations. This range is defined by the co-propagating wave coupling coefficient $\tilde{m}l_{jm}$, set by the Kerr modulation depth δ. Simulations confirmed that inversion gratings, due to the short upper-level lifetime of YAG:Cr^{4+}, exhibit negligible amplitude (Figs.7, 8), resulting in minimal counter-propagating wave coupling.

III. Astigmatic Ring Resonator with KTP Phase Modulator for Kerr-Induced Mode Locking

Integrating a KTP phase modulator into an astigmatic ring resonator of a ring laser with Kerr-induced mode locking demands careful selection of resonator configuration and parameters. Equations describing the propagation of an astigmatic Hermite-Gaussian beam in the ring resonator were derived using a matrix approach and transformations of two complex parameters in orthogonal planes. Solving there equations yielded eigenmode parameters and identified the optimal resonator configuration with the KTP phase modulator.

Frequency biasing in the solid-state laser gyroscope with mode locking was achieved by periodically applying opposite-sign phase shifts to counter-propagating pulses traversing the modulator. The KTP electro-optic crystal, grown using high-resistivity technology, exhibited reduced piezoelectric "ringing" compared to conventional lithium niobate.

Kerr-induced mode locking requires high power density in the gain medium and tight beam focusing to form a Kerr lens, as the Kerr nonlinearity in YAG:Cr^{4+} is weak (n_2=6.2*10^{-16} cm^2/W). Gyroscope noise and accuracy constraints demand stable Kerr-induced mode locking and precise control of the Kerr-induced amplitude modulation depth. This imposes resonator design requirements to ensure first-order mode parameter invariance under external perturbations – i.e., dynamic stability. The ring laser's optical scheme (Fig. 9) employs a Z-shaped configuration with strong focusing in the YAG:Cr^{4+} gain medium (Brewster-cut, 2 cm long) and a collimated beam in the dispersive arm. Spherical mirrors M3 and M4 (5 cm focal length) create a tight beam waist in the gain medium to enhance Kerr lensing. A thermal lens with greater optical power also forms in the medium. Due to oblique beam incidence on the gain medium and other elements, the Hermite-Gaussian beam becomes astigmatic. Optimizing the gain medium's position relative to focusing mirrors maximizes Kerr modulation depth and mode-locking stability. Additionally, the resonator configuration minimizes pulse front and back broadening via aperture clipping and reduces group delay dispersion (GDD) to suppress pulse spreading. A quartz prism compensates for positive dispersion in the YAG:Cr^{4+} medium and KTP modulator (Figs. 15-17). Notably, KTP exhibits lower GDD per mm (Fig. 16) than lithium niobate, enabling effective dispersion compensation with quartz.

Resonator parameter calculations and mode structure simulations were done through the "ABCD" ray matrix method for astigmatic Gaussian beam propagation. For each astigmatic resonator element (Fig. 9), separate ray matrices were defined for p- and s-planes. Total resonator round-trip matrices were computed for both planes, yielding beam waist sizes in the gain medium. Mirror spacing and incidence angles were symmetrically adjusted to optimize beam alignment. The prism's influence on mode formation was deemed negligible due to the large beam size and wavefront curvature at its location.

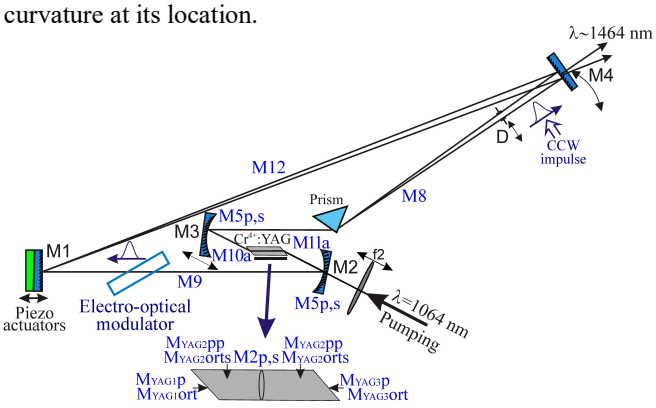


Fig. 9. Optical scheme of the astigmatic ring resonator enabling Kerr-induced mode locking.

For the sagittal s-plane, the resonator round-trip matrix at the gain medium center is derived as [17]:

MMMLs=M2s*MYAG2orts*MYAG1ort*M10a*M5s*M8 *M12*__M9s__*M5s*M11a*MYAG3ort*MYAG2orts, (5)

where the ray matrices for resonator elements are defined as follows [17]: M2s is thermal or Kerr lens, MYAG2orts is segment of the gain medium, MYAG1ort is interface boundary, M10a, M11a are segments between the spherical mirror and gain medium, M5s is spherical mirror, M8, M12 are free-space propagation matrices, MYAG3ort is interface boundary.

For the KTP phase modulator (Fig. 14), the matrix **M9s** (modulator segment for the s-plane) can be written as:

M9s= M9L2*M9edge2s*M9Lktps*M9edge1s* M9L1, (6)

where $\mathrm{M9Lktps}=\begin{pmatrix}1 & \mathrm{Lktp}/nk\\ 0 & 1\end{pmatrix}$ is segment of the KTP crystal, $\mathrm{M9edge1s}=\begin{pmatrix}nk & 0\\ 0 & nv\end{pmatrix}$, $\mathrm{M9edge2s}=\begin{pmatrix}nv & 0\\ 0 & nk\end{pmatrix}$ are interface boundaries of the KTP crystal, $\mathrm{M9L1}=\begin{pmatrix}1 & L_1\\ 0 & 1\end{pmatrix}$, $\mathrm{M9L2}=\begin{pmatrix}1 & L_2\\ 0 & 1\end{pmatrix}$ are free-space propagation matrices, Lktp is optical path length traversed by radiation in the KTP modulator, *nv*, *nk* are refractive indices of air and KTP, accounting for birefringence, L1, L2 are distances before and after the KTP modulator.

For the incidence plane p-, the resonator round-trip matrix at the gain medium center is [17]:

MMMLp=M2p*MYAG2pp*MYAG1p*M10a*M5p*M8* M12*__M9p__*M5p*M11a*MYAG3p*MYAG2pp, (7)

where M2p is thermal or Kerr lens, MYAG2pp is gain medium segment, MYAG1p is interface boundary, M10a, M11a are segments between spherical mirror and gain medium, M5p is spherical mirror, M8, M12 are free-space propagation matrices, MYAG3p is interface boundary.

For the KTP modulator in the p-plane (Fig. 14), the matrix **M9p** is:

M9p= M9L2*M9edge2p*M9Lktpp*M9edge1p* M9L1, (8)

where $\mathrm{M9Lktpp}=\begin{pmatrix}1 & \mathrm{Lktp}/nk\\ 0 & 1\end{pmatrix}$ is segment of the KTP crystal, $\mathrm{M9edge1s}=\begin{pmatrix}\mathrm{A9} & 0\\ 0 & \mathrm{D9}\end{pmatrix}$, $\mathrm{M9edge2s}=\begin{pmatrix}\mathrm{D9} & 0\\ 0 & \mathrm{A9}\end{pmatrix}$ are interface boundaries of the KTP crystal,

The coefficients for the KTP boundary matrices are determined by the angles of incidence and refraction:

$\mathrm{A9}=nk*\cos(\varphi_{B2})/\cos(\varphi_{B1})$, $\mathrm{D9}=nv*\cos(\varphi_{B1})/\cos(\varphi_{B2})$,

where angle of incidence (Brewster angle) and refraction angle:

$\varphi_{B1}=\mathrm{arctg}(nk/nv)$, $\varphi_{B2}=\arcsin(nv/nk*\sin(\varphi_{B1}))$.

From the resonator round-trip matrices at the gain medium center the elements A, B, C, D can be found from the following: $\mathrm{MMMLp,s}=\begin{pmatrix}\mathrm{A_{p,s}} & \mathrm{B_{p,s}}\\ \mathrm{C_{p,s}} & D_{p,s}\end{pmatrix}$.

Beam waist radius (Figs. 10-12) and wavefront curvature radii are derived by solving the quadratic equation:

$$\omega_{\mathrm{p,s}}=\sqrt{\frac{2\cdot\lambda\cdot B_{\mathrm{p,s}}}{\pi\sqrt{4-(\mathrm{A_{p,s}}+D_{\mathrm{p,s}})^2}}} \quad (9),$$

$$\mathrm{R_{p,s}}=\frac{2B_{p,s}}{D_{p,s}-A_{p,s}} \quad (10)$$

According to mode structure calculations (Figs. 10-12), the KTP modulator's position was optimized to ensure proper mode structure both in the gain medium and aperture while maximizing frequency bias. Counter-propagating pulses traversed the modulator out of phase. Mirror spacing was adjusted to minimize beam size sensitivity (first derivatives near zero for p- and s- planes), enhancing Kerr mode locking stability against external perturbations.

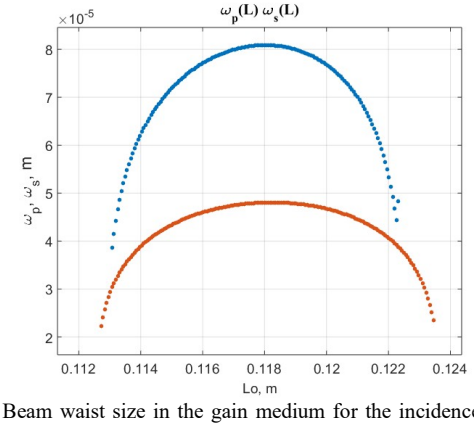


Fig. 10. Beam waist size in the gain medium for the incidence plane (ω_p, top) and orthogonal plane (ω_s, bottom) versus spherical mirror spacing L_o.

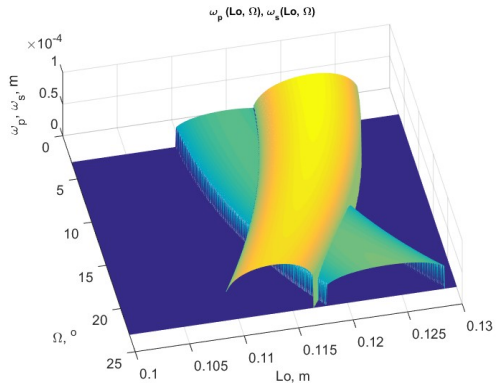


Fig. 11. Beam waist size in the active medium (ω_p, ω_s) versus spherical mirror spacing L_o and incidence angles θ.

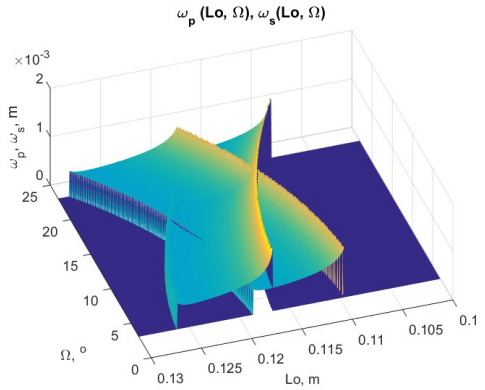


Fig. 12. Beam waist size at the output mirror (ω_p, ω_s) versus L_o and θ.

Beam waist analysis (Fig. 10, 11) for p- and s- planes depending on mirror angles and spacing identified optimal alignment and minimal waist separation.

IV. KTP-BASED ELECTRO-OPTIC PHASE MODULATOR

In a solid-state laser gyroscope with Brewster-angle-mounted optical elements, linear polarization is generated. Implementing frequency biasing via magneto-optical methods is challenging, and leveraging the Sagnac effect with periodic rotation introduces additional complexities. Thus, this work employs alternating phase modulation with opposite signs for counter-propagating pulses generated in the ring laser. In a mode locked ring resonator, counter-propagating pulses circulate and traverse the phase modulator out of phase under specific conditions. This enables frequency shifting of the counter-propagating waves in opposite directions, creating frequency splitting.

Electro-optic crystals driven by out-of-phase modulating voltages are ideal for phase modulation. Among commonly used crystals ($LiNbO_3$, KTP, BBO), KTP is optimal for a ~1.5 μm wavelength laser gyroscope (Table. 1).

Applying a control voltage to the modulator during pulse transit alters the refractive index, shifting the lasing frequency relative to the counter-propagating wave (11, 12).

$$\Delta\varphi(U)=\frac{\pi n_e^3 r_{33} l U}{\lambda D},\quad \mathrm{r_{ij}}=\begin{pmatrix}0 & -r_{22} & r_{13}\\ 0 & r_{22} & r_{13}\\ 0 & 0 & r_{33}\\ 0 & r_{51} & 0\\ r_{51} & 0 & 0\\ -r_{22} & 0 & 0\end{pmatrix}, \quad (11)$$

where n_e is the extraordinary refractive index, r_{ij} is the electro-optic coefficients tensor, r_{33}=35 pm/V is the electro-optic coefficients tensor for the chosen KTP orientation, U is the voltage applied to crystal, l is the crystal length, λ is the wavelength, D is the crystal thickness. The KTP crystal is oriented such that radiation propagates at the Brewster angle (Fig. 14), with the electric field aligned along the z-axis (polarization direction) and radiation propagating along the x-axis. For frequency modulation in the ring laser (Fig. 13), the frequency shift for one wave direction under phase modulation is:

$$\Delta\nu(U) = \frac{c}{2\pi L}\Delta\varphi(U)\,. \qquad (12)$$

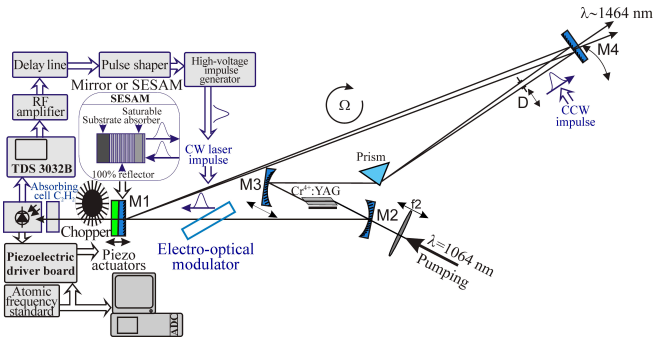


Fig. 13. Control system for the KTP electro-optic phase modulator and cavity length stabilization via a C_2H_2 absorption cell.

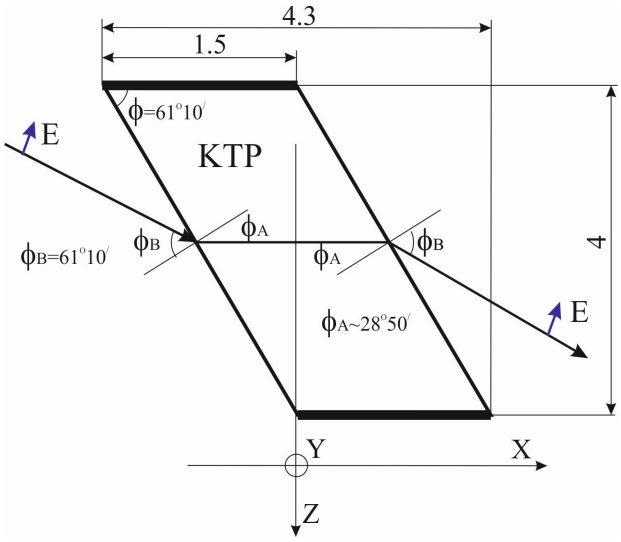


Fig. 14. Optical scheme of the KTP electro-optic phase modulator.

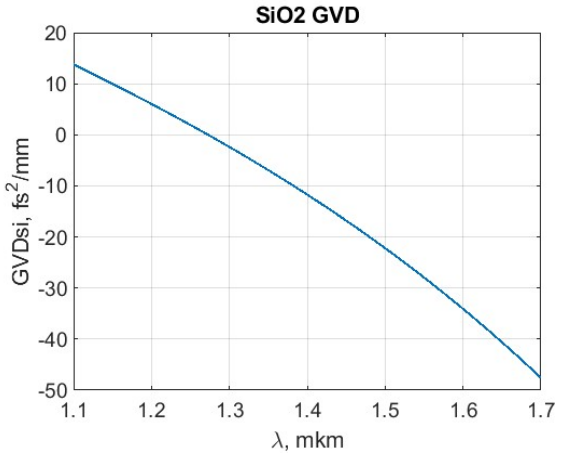


Fig. 15. Group velocity dispersion (GVD) versus wavelength for SiO_2.

TABLE I. OPTICAL PARAMETERS OF KTP ($KTiOPO_4$)

Transmitting Range[2]:	350 nm ~ 4500 nm			
Refractive Indices[1]:	**λ 1500 nm**	1.7290 (n_x)	1.7358 (n_y)	**1.8169 (n_z)**
Sellmeier Equations[2]: (λ in μm)	$N_x^2 = 3.0065 + 0.03901 / (\lambda^2 - 0.04251) - 0.01327\lambda^2$			
	$N_y^2 = 3.0333 + 0.04154 / (\lambda^2 - 0.04547) - 0.01408\lambda^2$			
	$N_z^2 = 3.3134 + 0.05694 / (\lambda^2 - 0.05658) - 0.01682\lambda^2$			
Sellm. Eq.[1]	**$N_z^2 = 4.59423+0.06206/(\lambda^2-0.04763)+110.80672/(\lambda^2-86.12171)$**			
Therm-Optic Coefficient[2]: (10^{-5}/°C)	dn_x/dT=1.1	dn_y/dT=1.3	dn_z/dT=1.6	
Absorption Coefficient[2]:	a<1%/cm @1064 nm and 532 nm			
Electro-optic coefficients[2]:	Low Frequency (pm/V)	High Frequency (pm/V)		
r_{13}	9.5	8.8		
r_{23}	15.7	13.8		
r_{33}	36.3	**35.0**		
r_{51}	7.3	6.9		
r_{42}	9.3	8.8		
Dielectric constant[2]:	ε_{eff}=13			

[a] 1: https://refractiveindex.info/?shelf=main&book=KTiOPO4&page=Kato-%CE%B3
[b] 2: https://www.unitedcrystals.com/KTPProp.html

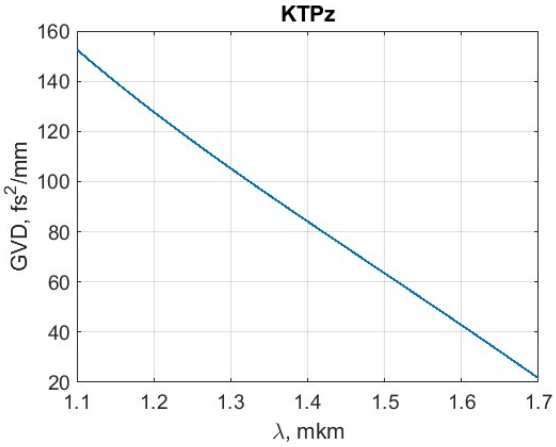


Fig. 16. GVD versus wavelength for KTP (polarization along the z-axis).

This induces a frequency difference between counter-propagating beams, forming the bias (Fig. 18). Pulse detection at the ring laser output utilized a high-speed InGaAs PIN photodiode with an amplifier (Fig. 13).

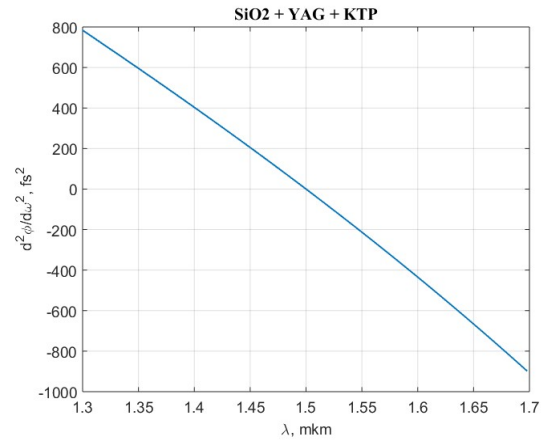


Fig. 17. Group delay dispersion versus wavelength for a **2 cm YAG:Cr^{4+}** crystal, **13.052 mm SiO_2** prism, and **1.5 mm KTP** modulator. **Zero dispersion** is achieved at 1.5 μm.

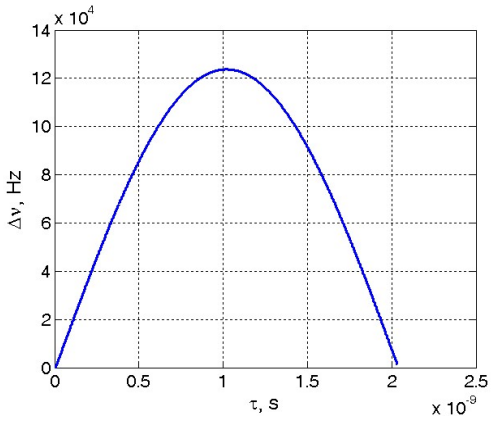


Fig. 18. Frequency bias (counter-propagating wave splitting) versus control voltage phase delay relative to modulator pulses (U = 10 V).

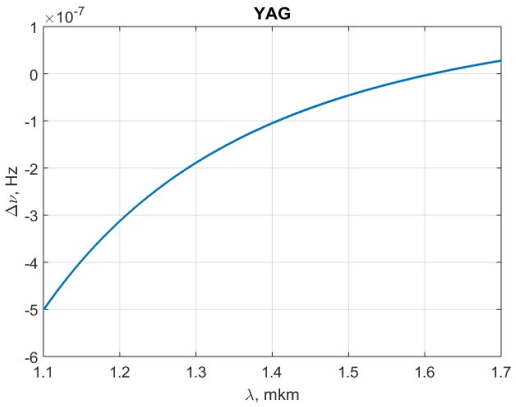


Fig. 19. Zero-drift (frequency splitting) caused by YAG:Cr^{4+} dispersion during 1 MHz laser frequency tuning. Temperature-induced wavelength instability in the C_2H_2 absorption cell is <0.01 pm/°C (1.33 MHz/°C). At ~1.5 µm, **zero-drift is ~5×10⁻⁸ Hz/°C**.

V. Experimental Setup

The solid-state laser gyroscope employs a YAG:Cr^{4+} gain medium (Fig. 20) [17] with a custom-designed KTP phase modulator operating at the Brewster angle.

This modulator enables phase modulation for frequency biasing during mode locking. Compared to $LiNbO_3$, KTP offers lower group delay dispersion (GDD), reduced piezoelectric coefficients (minimizing "ringing"), higher electro-optic coefficients, and other advantages. A critical design compromise involved balancing the optical path length (determining frequency bias magnitude at ~220 MHz driving voltage) and the need to compensate for GDD introduced by the crystal (Fig. 17). The YAG:Cr^{4+} ring laser operates across a broad wavelength range (1.3-1.6 µm), enabling femtosecond pulse generation via Kerr-induced mode locking. In the YAG:Cr^{4+} crystal, Cr^{4+} ions occupy sites with varied oxygen and Mg^{2+} coordination (Fig. 21), creating an inhomogeneous intracrystalline field. This leads to inhomogeneous broadening of the gain linewidth and suppresses competition between counter-propagating waves, allowing stable bidirectional lasing – both with and without mode locking. Total output power in both directions reached ~ 200 mW. A fiber-coupled ytterbium laser was used for pumping. Dispersion compensation of resonator elements was achieved by utilizing a Brewster-cut prism (Suprasil 300 IR quartz, $OH^- < 1$), providing negative GDD at 1.5 µm ppm (Fig. 17). A single prism turned to be sufficient for the task. Wavelength tuning and emission spectrum adjustment were performed through an aperture D (Fig. 20), while mirror angles (M3, M4) optimized resonator alignment.

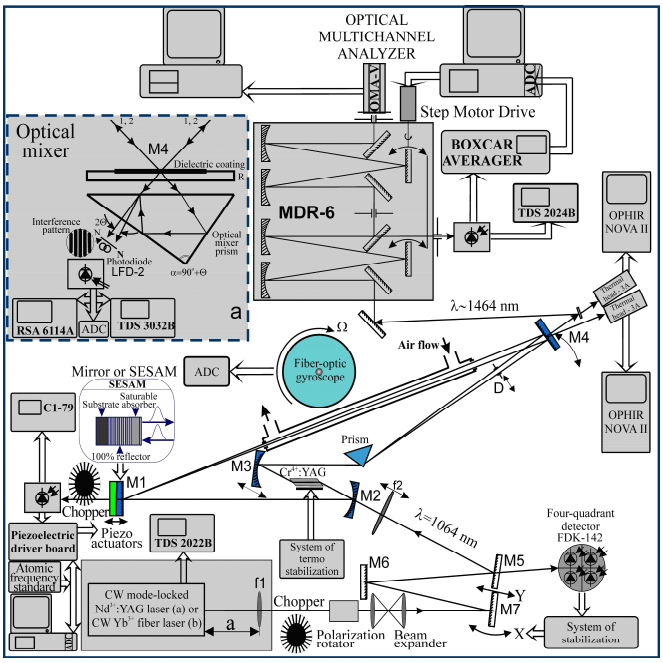


Fig. 20. Scheme of the YAG:Cr^{4+} laser gyroscope experimental setup.

A YAG:Cr^{4+} of 2 cm length and 6 mm diameter with Brewster-polished faces served as gain medium. Cr^{4+} concentration was 5×10^{17} cm^{-3}. An alternative 1.5 cm × 5 mm crystal with higher Cr^{4+} concentration was also tested, showing reduced GDD, thermal lensing and improved performance. Parameters of the crystal are following [6, 12]: upper state lifetime of 3.6 µs, gain cross-section at 1.42 µm is $(7\text{-}8)\times10^{-19}$ cm^2. Absorption peak at 1 µm corresponding the transition of $^3B_1\left(^3A_2\right)\rightarrow{}^3A_2\left(^3T_1\right)$. Absorption coefficient at 1.06 µm is ~2.5 cm^{-1} (2 cm crystal) and ~1.5 cm^{-1} (1.5 cm crystal). Pump absorption at crystal length is ~ 90%. Operating scheme is a four-level system between the broad $^3T_2$ and $^3A_2$ vibrational bands.

Experimental results have shown a reduced wave competition in a YAG:Cr^{4+} ring laser attributed to femtosecond pulse dynamics and Cr^{4+} - polarized light interactions (Fig. 21). Perimeter stabilization via an acetylene absorption cell limited thermal zero drift to ~5×10^{-8} Hz/°C (Fig. 19). Mode locking narrowed mode spectra and suppressed noise (Fig.22).

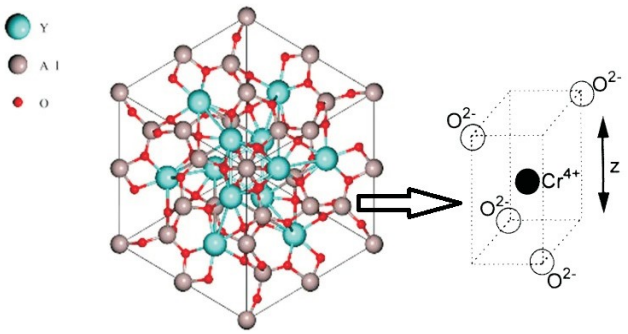
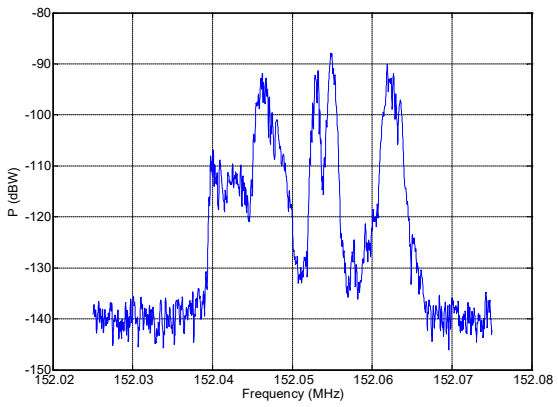
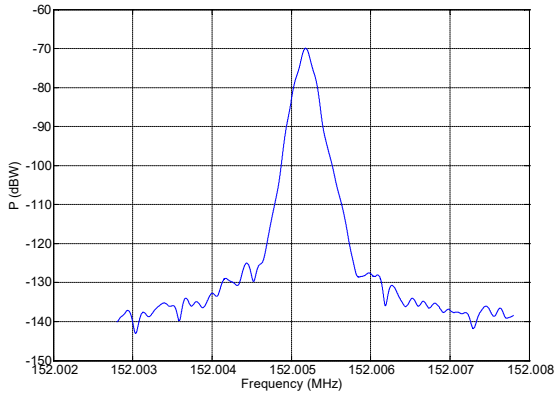


Fig. 21. YAG:Cr^{4+} crystal lattice: Cr^{4+} ion placement and coordination environment, which alters lasing properties and reduces counter-propagating wave competition.



a.



b.

Fig. 22. . Beat spectrum of longitudinal modes: (a) without mode locking (25 kHz linewidth); (b) with mode locking (700 Hz linewidth).

VI. Conclusion

The conducted studies demonstrate that an electro-optic phase modulator, particularly using a KTP crystal, is optimal for implementing frequency biasing in a solid-state laser gyroscope operating in the mode locking regime. The KTP modulator achieves a frequency bias of 240-480 kHz at feasible RF voltage levels (10-20 V). For fixed-frequency operation, the required RF voltage can be generated using a resonant circuit. Frequency biasing was achieved by synchronizing phase shifts of opposite signs with the alternating passage of counter-propagating pulses through the modulator. The geometry, orientation, and optimal parameters of the KTP phase modulator were determined. A balance was established between the group delay dispersion introduced by KTP, radiation perturbations, and the achievable frequency bias magnitude, which is limited by practical RF voltage constraints.

Numerical simulations of lasing of the solid-state laser gyroscope — both in mode-locked and free-running regimes — using equations for the complex amplitudes of four modes per direction (including frequency biasing effects) revealed that phase alignment via Kerr-induced coupling of adjacent modes suppresses noise when the co-propagating mode coupling significantly exceeds both the counter-propagating wave coupling (due to scattering) and the frequency bias amplitude. Enhancing mode coupling and Kerr modulation stability amplitude improves robustness to perturbations. The significantly shorter upper-level lifetime of YAG:Cr^{4+} (compared to YAG:Nd^{3+}) reduces inversion grating effects, further weakening counter-propagating wave coupling relative to conventional solid-state laser gyroscopes.

Through astigmatic Gaussian beam modelling, a ring resonator configuration was identified that ensures radiation stability under external perturbations. This configuration nullifies the first derivatives of radiation parameter dependencies on perturbations when the KTP phase modulator is integrated.

An acetylene-filled absorption cell was used. At ~1.5 μm, the gyroscope's zero drift—caused by YAG:Cr^{4+} dispersion and thermal instability of the acetylene absorption line—reaches ~5×10^{-8} Hz/°C.

In general, the conducted studies have shown the possibility of stable operation of a solid-state YAG:Cr^{4+} laser gyroscope during mode locking with a cavity length control system and a frequency bias created by a KTP-based electro-optic phase modulator.

REFERENCES

[1] Aronowitz, F., Laser gyros, In: Lasers Applications. Russian edition, M: «Mir», 1974. p. 182-269.

[2] Khanin, Ya.I., Osnovy Dinamiki Lazerov, Foundations of Laser Dynamics. M.: Nauka, «Fizmatlit», 1999.

[3] M.J. Bohn, The TI:Sapphire Ring Laser Gyroscope, dissertation, University of New Mexico, 1998.

[4] Sorokin, E. Solid-State Materials for Few-Cycle Pulse Generation and Amplification. In: Kärtner, F.X. (eds) Few-Cycle Laser Pulse Generation and Its Applications. Topics in Applied Physics, vol 95. Springer, Berlin, Heidelberg.. https://doi.org/10.1007/978-3-540-39849-3_1

[5] H. Eilers et al., "Performance of a Cr:YAG laser", in IEEE Journal of Quantum Electronics, vol. 29, no. 9, pp. 2508-2512, Sept. 1993.

[6] A.G. Okhrimchuk and A.V. Shestakov, "Absorption saturation mechanism for YAG:Cr4+crystals", Phys. Rev. B 61, 988, 2000.

[7] G.M. Zverev and A.V. Shestakov, "Tunable Near-Infrared Oxide Crystal Lasers," in Advanced Solid State Lasers, M. Shand and H. Jenssen, eds., Vol. 5 of OSA Proceedings Series (Optical Society of America, 1989), paper BB5.

[8] I.T. Sorokina, S. Naumov, E. Sorokin, E. Wintner, and A.V. Shestakov, Directly diode-pumped tunable continuous-wave room-temperature Cr4+:YAG laser, Opt. Lett. 24, 1578-1580 (1999).

[9] N.V. Kravtsov, E.G. Lariontsev "Self-modulation oscillations and relaxation processes in solid-state ring lasers" Quantum Electronics 24(10) 841-856 (1994).

[10] Schwartz, S., Lauprêtre, T., Goldfarb, F., Bretenaker, F., Ghosh, R., Carusotto, I., Feugnet, G., Gutty, F., Pocholle, J.-P., The solid-state ring laser gyro: current and future trends, Proc. SPIE 8636, Advances in Slow and Fast Light VI, 86360M (March 6, 2013).

[11] Malakhov, A.N. and Mal'tsev, A.A., Width of the spectral lines of an N-mode system which has an equidistant spectrum of natural frequencies and operates in a self-synchronous mode, Radiofizika, 1973. Vol. 16, No. 8, p. 1211-1220.

[12] Y.Y. Broslavets, E.A. Polukeev, V.P. Surovtseva, A.A. Fomichev and V. G. Semenov, "Multifrequency YAG:Cr4+ Solid-State Laser Gyroscope, Perimeter Control and Dither Creation System, Operating Regimes", 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-4, doi: 10.23919/ICINS51784.2022.9815385.

[13] Schwartz, S., Gutty, F., Feugnet, G., Loil, É., and Pocholle, J.-P., Solid-state ring laser gyro behaving like its helium–neon counterpart at low rotation rates, OPTICS LETTERS. 2009 .Vol. 34, № 24, p. 3884-3886.

[14] N.V. Kravtsov, B.V. Lvov, V.Yu. Petrun'kin, K.B. Samusev, A.N. Shelaev, V.I. Shokalo "Synchronization of longitudinal modes in a solid-state ring laser with semiconductor pumping", - Avtometriya, 4, Leningrad, 1987.

[15] Ya.Z. Virnik, A.S. Kovalev and E.G. Lariontsev, Natural frequency fluctuations in a laser having locked modes, Radiophysics and Quantum Electronics, Volume 13, pages 1363-1367, (1970).

[16] V. Arzt, H. Haken, H. Risken, H. Sauermann, Ch. Schmid, and W. Weidlich " Quantum Theory of Noise in Gas and Solid State Lasers with an Inhomogeneously Broadened Line" Zeitschrift fur Physik 197, 207-227 (1966).

[17] Y.Y. Broslavets, E.A. Polukeev, D.S. Redichkina, A.A. Fomichev and V. G. Semenov, A.R. Pokrovskaya, "Noise suppression in the output signal of a YAG:Cr4+ solid-state laser gyroscope with cavity length control and mode-locking", 31st Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2024, pp. 1-9.

Vibration-Resistant Quartz Pendulum Accelerometer with Digital Feedback

D. M. Kalikhman, E. A. Deputatova, N. V. Tarakanov

Yuri Gagarin State Technical University of Saratov,
Branch of Academician Pilyugin Center –
Production Association Korpus,
Saratov, Russia
lidkalihman@yandex.ru, deputatova@bk.ru
nikolaytarakanov87@gmail.com

D. S. Gnusarev, L. Ya. Kalikhman

Branch of Academician Pilyugin Center –
Production Association Korpus,
Saratov, Russia
men4eg@gmail.com, lidkalihman@yandex.ru

Abstract — The paper discusses the results of the development of a quartz pendulum accelerometer with a digital feedback amplifier, in which a variable structure controller is implemented to reduce vibration error.

Key words — quartz pendulum accelerometer, digital feedback amplifier, vibration error, variable controller, algorithmic compensation.

I. INTRODUCTION

Technologies of strapdown inertial navigation systems (SINSs) are becoming widely used in all areas of navigation development: aviation, space, marine, pedestrian navigation systems, unmanned aerial vehicles and highly maneuverable objects [1, 2]. In turn, inertial sensing elements – primary SINS sensors, in which apparent acceleration sensors play a key role, – are also advancing [3–9].

Vibration errors in primary sensors have always been a serious hindrance for SINS performance in any navigation system, especially in highly maneuverable objects [3–6].

Quartz pendulum accelerometers are widely used in modern SINSs for rocket, space and aviation application, as well as in other control systems [5, 10, 11]. The Saratov branch of Academician Pilyugin Center – Production Association Korpus has been developing and manufacturing quartz pendulum accelerometers (QPAs) for more than 20 years and has repeatedly reported on the development of QPAs with digital feedback at international conferences on integrated navigation systems. Since 2019, the inertial measurement unit featuring non-orthogonal orientation of the measuring axes and a digital feedback amplifier in each measuring channel (also known under its label name “BILU”) has been successfully operated as part of the control system of the Soyuz MS and Progress MS spacecraft [4, 12]. Digital feedback allows for algorithmic compensation of QPA temperature errors in each measuring channel, which was discussed in [4, 13–17].

However, in manned aviation and space systems, external vibration disturbances are not very large for the understandable reason of crew presence. Highly maneuverable objects are subject to more serious disturbances that will inevitably have a negative impact on QPA operation, significantly increasing its vibration error. If analog feedback is used, it is impossible to combat this error by means of algorithmic compensation, so other methods, described in research reports [18], are to be implemented. The presence of digital feedback allows for applying the methods of automatic control theory and synthesizing a variable structure controller, thus providing a significant reduction of vibration error with simultaneous compensation of QPA temperature errors and ensuring the dynamic stability of the navigation instrument over the entire range of apparent acceleration measurement [19–23].

II. SIMULATION RESULTS

Analysis of the results of mathematical modeling of QPA operation performed in the Comsol Multiphysics simulation environment showed that the quartz plate, when subjected to alternating acceleration in the direction perpendicular to the plate plane (i.e. in the direction of the sensing axis), performs not only bending oscillations, but also oscillations around the transverse axes [5, 16]. A number of proper resonance frequencies in the range of 0-2000 Hz, characteristic of the studied accelerometer design, were identified. It was determined experimentally that, despite the actual presence of these resonance frequencies, only one in the range of 450 Hz to 600 Hz is of interest, since the others do not affect the instrument operation. The vibration errors of measurement of the constant component of the effective acceleration are increased at this frequency. In the case of analog feedback using amplifiers built into the mentioned types of accelerometers, it is not possible to reduce these errors at this frequency. The problem can be solved by building a QPA automatic control system (ACS) with a variable structure digital controller [20, 21, 24–27].

The digital feedback algorithm is defined by a difference equation in the following form [4, 13, 14]:

$$\begin{aligned} I_{MS}(nT_0) = {} & 3.078714 \cdot U_{AS}(nT_0) - \\ & - 4.191684 \cdot U_{AS}((n-1)T_0) + \\ & + 1.281341 \cdot U_{AS}((n-2)T_0) - \\ & - (-1.358005 \cdot Y_{AS}((n-1)T_0) + \\ & + 0.3580045 \cdot Y_{AS}((n-2)T_0)), \end{aligned} \quad (1)$$

where I_{MS} is the current of the pulling force moment sensor; U_{AS} is the voltage of the angle sensor mismatch error expressed in

analog-to-digital converter (ADC) units; Y_{AS} is the control value at the previous feedback cycles, corresponding to the discrete transfer function [19, 27]:

$$W_{reg}(z) = \frac{2.837292 \cdot z^2 - 3.862987 \cdot z^1 + 1.180863}{z^2 - 1.357829 \cdot z^1 + 3.579417}; \quad (2)$$

obtained as the result of digital controller synthesis [4, 13, 14].

As was shown in [4, 12–14], it is possible to provide temperature stabilization for a QPA under thermal perturbations with a conventional structure of a digital controller. Let us consider the functional and kinematic diagram of the instrument, which enables a rather effective compensation of QPA vibration errors.

III. Functional and Kinematic Diagram of QPA with Variable Structure Digital Controller

Fig. 1 shows the functional and kinematic diagram of the QPA with a variable structure digital controller. The proposed circuit solution differs from the previously developed ones in that it includes an adapter, which is responsible for changing the structure of the digital controller under the action of external disturbing influences, mainly vibration.

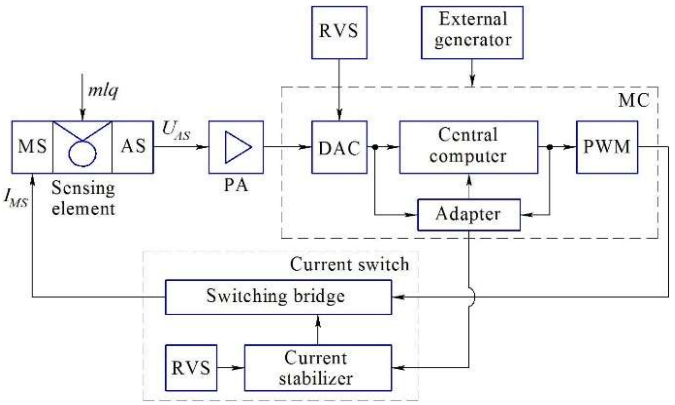


Fig. 1. Functional and kinematic diagram of QPA with variable structure digital controller

The following designations are introduced in Fig. 1: *ml* – pendulosity; *q* – input acceleration; MS – pulling force moment sensor; AS – angle sensor; PA – preliminary amplifier; RVS – reference voltage source; ADC – analog-to-digital converter; PWM – pulse width modulator; MC – microcontroller.

The adapter does not have to be a separate physical device, it can be understood as a portion of computational resources of the processor included in the QPA feedback.

The instrument operation is carried out as follows. The QPA measures the apparent acceleration, working as part of the SINS, a signal in the form of the variable voltage U_{AS} is then sent from its angle sensor to the preliminary amplifier of the accelerometer, where it is amplified and fed to the ADC of the microcontroller in the instrument's feedback. The central computer implements digital control in the form of a finite-difference equation (1) based on the synthesized discrete transfer function (2).

The laws of variation of the difference equation coefficients, which are obtained in the Comsol Multiphysics simulation environment, are flashed into the adapter. The character of these laws can be diverse depending on the operating conditions of the control object, which are specified in the technical specification (TS) for the SINS development. As soon as the operating conditions of the instrument approach the regions close to the resonance frequencies, the adapter is switched on and changes the corresponding coefficients of the difference equation. As soon as the object leaves the region of resonance frequencies, the coefficients of the difference equation return to their original values. Other than that, the circuit solution shown in Fig. 1 operates according to the feedback principle with pulse width control of the current of the pulling force moment sensor as described in [4, 12–14].

IV. Results of Experimental Research

A measuring channel of the BILU device with digital control of the pulling force moment sensor's current was taken as an experiment. The QPA resonance frequencies were determined by means of mathematical modeling, and the exponential law of the form shown in Fig. 2 was taken as the law of variation of the difference equation coefficients.

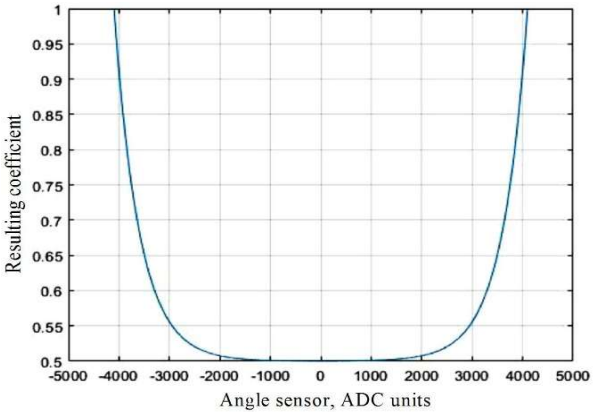


Fig. 2. The law of variation of the coefficients of the difference equation of the digital controller

The coefficient of the difference equation standing at the U_{AS} was varied according to the exponential law in the range from 0.5 to 1. A coefficient greater than 0.5 was applied only when the angle sensor mismatch error was greater than 2500 ADC units, which ensured its application when approaching resonance frequencies. The discrete transfer function for the 0.5 coefficient is represented by expression (3), and for the coefficient equal to 1 – by expression (2).

$$W(z) = \frac{1.418646 \cdot z^2 - 1.9314935 \cdot z^1 + 0.5904315}{z^2 - 1.357829 \cdot z^1 + 3.579417}. \quad (3)$$

The phase stability reserves amounted to 40° for coefficient 0.5 and 30° for coefficient 1, the amplitude reserves were 20 dB for coefficient 0.5 and 10 dB for coefficient 1, which guarantees stable operation of the accelerometer under perturbations of any kind, as well as ensures good robustness of the automatic control system. In view of the introduced changes, the considered

measuring channel was tested on the vibration test bench where it was experimentally established that the control performed in such a way allows for the reduction of the vibration errors in the region of resonance frequencies by a factor of 3 or 4. The absolute values of vibration error for a pair of accelerometers in consideration of the described methodology are presented in Fig. 3.

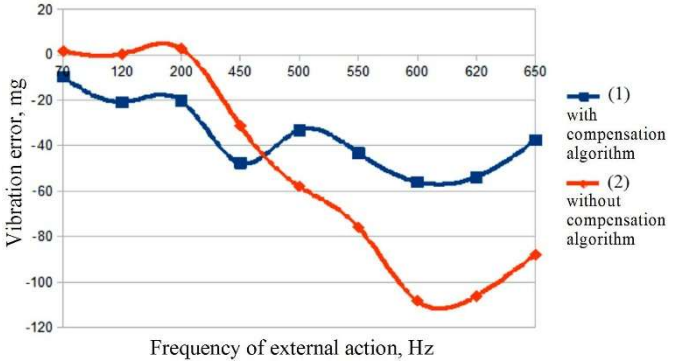


Fig. 3. Graph of the dependence of the vibration error of measuring the constant component of the accelerometer with the compensation algorithm (1) and without the algorithm (2)

V. Conclusions

The following conclusions can be drawn from the results of presented theoretical and experimental research.

1. The proposed QPA circuit solution with digital feedback, pulse width control of the pulling force moment sensor current, and variable structure of the controller has shown its effectiveness under the conditions of vibration disturbances.

2. The selected exponential law of variation of the coefficients of the difference equation allowed to reduce the influence of vibration error of the QPA by a factor of 3 or 4 while maintaining the quality indicators of the accelerometer's ACS, providing the required robust properties of the system.

3. The character of change in the coefficients of the control law in finite-difference form, programmed into the processor of the QPA digital feedback amplifier, does not necessarily have to obey the exponential law, but can change according to other laws depending on the technical task given by the developers of the control system of the object on which the SINS is installed.

4. The proposed approach makes it possible in the future to build an adaptive QPA control system for any highly maneuverable object of rocket, space or any other area of engineering where pendulum gas-filled accelerometers are applied.

References

[1] Izmailov, E.A., Modern Trends in the Development of Technologies of Inertial Sensing Elements and Aircraft Systems, Trudy FGUP NPTsAP, 2010, № 1, pp. 27–35.

[2] Rivkin, B.S., Analiticheskii obzor sostoyaniya issledovanii i razrabotok v oblasti navigatsii za rubezhom, Issues 1–8 (Analytical Review of the State of Research and Development in the Field of Navigation in Foreign Countries), St. Petersburg, CSRI Elektropribor, 2017–2023.

[3] Andreichenko, K.P., Dinamika poplavkovykh giroskopov i akselerometrov (Dynamics of Float Gyroscopes and Accelerometers), Moscow, Mashinostroenie, 1987.

[4] Nakhov, S.F., Kalikhman, D.M. et al., Opyt proektirovaniya i izgotovleniya blokov izmeritelei lineinogo uskoreniya na kvartsevykh mayatnikovykh akselerometrakh s analogovoi i tsifrovoi sistemami upravleniya (Experience in Design and Manufacture of Linear Acceleration Sensors Units on Quartz Pendulum Accelerometers with Analog and Digital Control Systems), Saratov, Yuri Gagarin State Technical University of Saratov, 2021.

[5] Konovalov, S.F., Navigatsionnye akeselerometry. Uchebnoe posobie (Navigation Grade Accelerometers: Study Guide), Moscow, MGTU, 2023.

[6] Konovalov, S.F., Teoriya vibroustoichivosti akselerometrov (Theory of Vibration Resistance in Accelerometers), Moscow, Mashinostroenie, 1991.

[7] Konovalov, S.F., Nikitin, E.A., Pel'por, D.S., Ed., Giroskopicheskie sistemy. Proektirovanie giroskopicheskikh sistem. Akselerometry, datchiki uglovoi skorosti, integriruyushchie giroskopy i girointegratory: Uchebnoe posobie (Gyroscopic Systems: Design of Gyroscopic Systems: Accelerometers, Angular Rate Sensors, Integrating Gyroscopes: Study Guide), Moscow, Vysshaya shkola, 1980.

[8] Raspopov, V. Ya., Mikromekhanicheskie pribory: uchebnoe posobie (Micromechanical Instruments: Study Guide), Moscow, Mashinostroenie, 2007.

[9] Sinel'nikov, A.E., Nizkochastotnye lineinye akselerometry. Metody i sredstva proverki i graduirovki (Low Frequency Linear Accelerometers: Methods and Tools for Inspection and Calibration), Moscow, Izdatel'stvo standartov, 1979.

[10] Mel'nikov, V.E., Elektromekhanicheskie preobrazovateli na baze kvartsevogo stekla (Quartz Glass Electromechanical Converters), Moscow, Mashinostroenie, 1984.

[11] Seo, Jae-Beom, Parameter Optimization and Modeling of Operating Modes in Q-flex and Si-flex Compensating Accelerometers, Cand. Sci. dissertation, Moscow, 2012.

[12] Kalikhman, D.M., Samitov, R.M. et al., Flight Tests of the Six-Axis Unit of Specific Force Meters Based on Precision Quartz Pendulum Accelerometers with Digital Feedback as Part of the Soyuz MS-14 Spacecraft, 27th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2020, pp. 184–187.

[13] Kalikhman, D.M., Skorobogatov, V.V. et al., Quartz Pendulum Accelerometer with Digital Feedback as an Automatic Control System: Synthesis of Controllers and Software: Further Prospects, Navigatsiya i upravlenie letatel'nymi apparatami, no. 26, 2019, pp. 2–37.

[14] Grebennikov, V.I., Deputatova, E.A. et al., Pendulum Accelerometer with Digital Control and New Functional Capabilities, Izvestiya RAN. Teoriya i sistemy upravleniya, no 2, 2021, pp. 73–95.

[15] Gnusarev, D.S., Deputatova, E.A., Kalikhman, D.M., Analysis of Noise Components in a Quartz Pendulum Accelerometer with Digital Feedback Amplifier, Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologii, mekhaniki, optiki, St. Petersburg, vol. 18, no. 6, 2018, pp. 1091 – 1098.

[16] Gnusarev, D.S., Skorobogatov, V.V., Deputatova, E.A., 3D Modeling of Quartz Glass Sensitive Elements in a Hemispherical Resonator Gyroscope and Pendulum Accelerometer, Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologii, mekhaniki, optiki, St. Petersburg, vol. 19, no. 3, 2019, pp. 531–537.

[17] Kalikhman, D.M. et al., Method to Provide Vibration Stability in a Pendulum Linear Accelerometer with a Digital Feedback and a Vibration-Stable Pendulum Accelerometer Thereof, Patent RU no. 2615221, 2017.

[18] Il'inskii, V.S., Zashchita apparatov ot dinamicheskikh vozdeistvii (Protection of Apparatus from Dynamic Effects), Moscow, Energiya, 1970.

[19] Besekerskii, V.A., Popov, E.P., Teoriya sistem avtomaticheskogo regulirovaniya (Theory of Automatic Control Systems), Moscow, Nauka, 1975.

[20] Izerman, R., Tsifrovye sistemy upravleniya (Digital Control Systems), Moscow, Mir, 1984.

[21] Kuo, B., Teoriya i proektirovanie tsifrovykh sistem upravleniya (Theory and Design of Digital Control Systems), Moscow, Mashinostroenie, 1986.

[22] Sadomtsev, Yu.V., Konstruirovanie sistem upravleniya s obratnoi svyaz'yu po kriteriyam tochnosti i grubosti (Designing Control Systems

with Digital Feedback Using the Accuracy Criterion), Saratov, Yuri Gagarin State Technical University, 2003.

[23] Fedorov, S.M., Litvinov, A.P., Avtomaticheskie sistemy s tsifrovymi upravlyayushchimi mashinami (Automatic Systems with Digital Control Machines), Moscow, Leningrad, Energiya, 1965.

[24] Aleksandrov, A.G., Metody postroeniya sistem avtomaticheskogo upravleniya, (Methods of Development of Automatic Control Systems), Moscow, Fizmatlit, 2008.

[25] Aleksandrov, A.G., Optimal'nye i adaptivnye sistemy (Optimal and Adaptive Systems), Moscow, Vysshaya shkola, 1989.

[26] Kuropatkin, P.V., Optimal'nye i adaptivnye sistemy (Optimal and Adaptive Systems), Moscow, Vysshaya shkola, 1980.

[27] Nebylov, A.V., Garantirovanie tochnosti upravleniya (Ensurance of Control Accuracy), Moscow, Fizmatlit, 1998.

Prediction of Experimental Data of Pendulum Compensation Accelerometer by Moving Average Method by Means of Neural Network

V.M. Nikiforov
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

D. M. Kalikhman
Branch of Academician Pilyugin Center – Production Association Korpus
Saratov, Russia

E. A. Deputatova
Branch of Academician Pilyugin Center – Production Association Korpus
Saratov, Russia

N. V. Tarakanov
Branch of Academician Pilyugin Center – Production Association Korpus
Saratov, Russia

A.D. Sheiko
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

K.A. Andreev
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

A.A. Gusev
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

***Abstract* — The paper considers the problem of predicting the experimental data of a pendulum compensation accelerometer using the moving average method by applying neural network algorithms as a duplication of the well-known method of simulating the natural storage of gyroscopic instruments and accelerometers, in particular, which is widely applied in the instrument-making industry.**

Keywords — pendulum compensation accelerometer, moving average method, neural network, simulation of natural storage.

I. Introduction

Nowadays, the application area of strapdown inertial navigation systems (SINSs) is expanding and the requirements to their accuracy parameters and, accordingly, to the command instrumentation included in SINSs are increasing [1, 2]. The basic parameters of accelerometers are determined during manufacture since accelerometers cannot be calibrated immediately before use in their main products [3–10]. Therefore, the task of confirming the stability of accelerometer parameters during the full service life (10 to 15 years) is relevant. In turn, estimation of the temporal stability of accelerometer parameters requires precision measurements over a prolonged period of time, and the subsequent mathematical processing is realized by algorithms that reduce the influence of measurement errors.

The objective of this study is to reduce the time of testing instrumentation and duplicate the method of simulating natural storage.

Three accelerometer transducers with a sampling time of 10 *sec* were considered as **the object of study**, and the sensing elements of accelerometers were simulated by means of precision electric current.

The subject of the study is linear neural network that allows predicting experimental results.

II. The Predicting Experimental Results

In modern instrument-making industry the method of natural storage is widely known and is widely used in practice in order to confirm the persistence of product parameters after 10 to 15 years of storage [4, 6]. For this purpose the product is stored in a thermal chamber at a temperature of +70°C for a long period of time. Considering that 11.3 days of storage is equivalent to 1 year of product operation, i.e. 113 to 170 days of storage in a thermal chamber are required to confirm 10 to 15 years of instrument performance stability. The method proposed in the paper is not a means of substitution, but is proposed as a means of duplicating and validating the results of the natural storage simulation method.

In this regard, a five-step moving average method implemented by means of a linear neural network is proposed.

Three accelerometer transducers with a sampling time of 10 *sec* were considered as the object of study, and the sensing elements of accelerometers were simulated by means of precision electric current. Five-hour measurements (~2000 samples) of the ~1 *g* parameter were carried out, and the results are presented in Fig. 1.

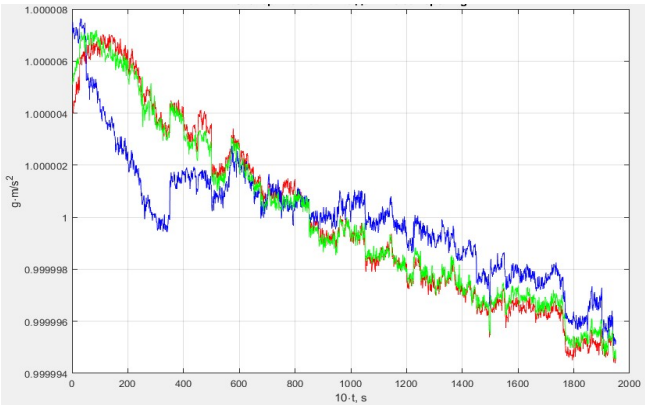


Fig. 1. Experimental values of the ~ 1 g parameter measurements from the TR1, TR2, and TR3 transducers

The structure of the linear neural network and the working fragment are presented in Fig. 2…6.

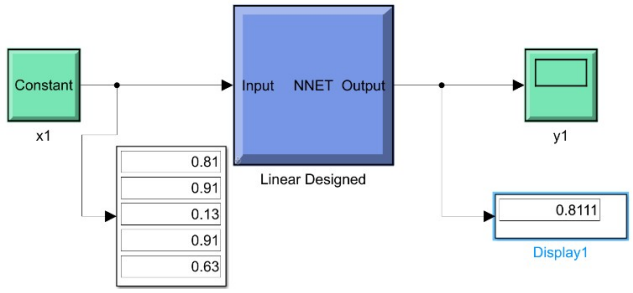


Fig. 2. General structure of the neural network

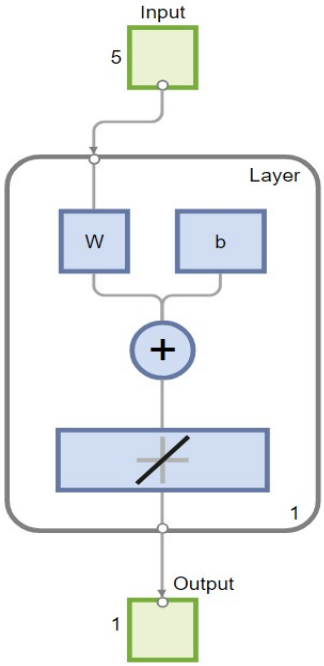


Fig. 3. Structural diagram of the linear neural network

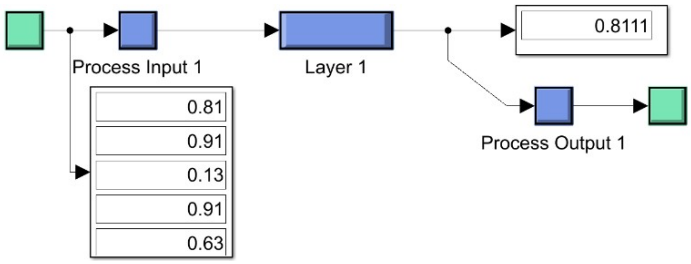


Fig. 4. General structural diagram of the linear layer

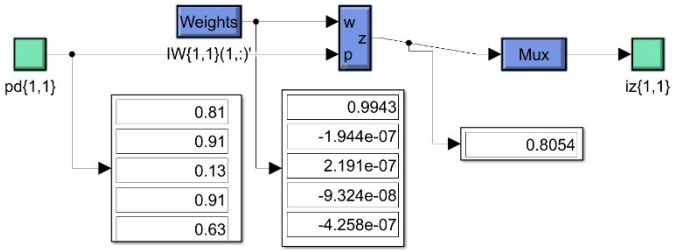


Fig. 5. Structural diagram of interaction of weight coefficients and input data

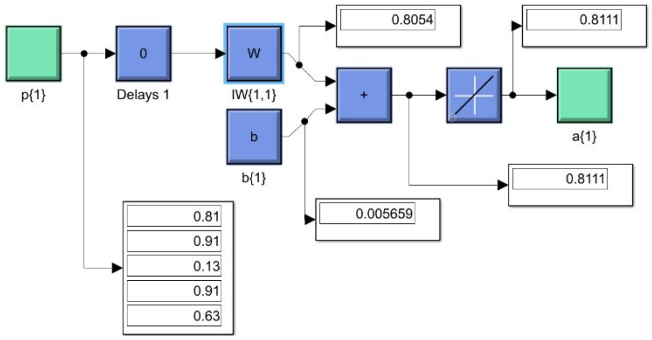


Fig. 6. Structual diagram of the linear layer

The weighting factors and the bias coefficient are respectively equal to:

$$W_{11} = 0.994341434384729,$$
$$W_{12} = -0.000000194449614,$$
$$W_{13} = 0.000000219102184,$$
$$W_{14} = -0.000000093238754,$$
$$W_{15} = -0.000000425764285,$$
$$b = 0.005659051043023.$$

The output signal is determined by the:

$$a\{1\}_{i+5} = p_i W_{11} + p_{i+1} W_{12} + p_{i+2} W_{13} + p_{i+3} W_{14} + p_{i+4} W_{15} + b_i,$$
$$i = 1 \ldots N,$$

where p_i - input data,

N – the number of experimental points.

Neural network (NN) training was performed on the first 1000 samples as shown in [11–17]. Conducted testing (green color) and validation (red color) showed good compatibility of experimental, testing, and validation data presented in Fig. 7.

When testing each new device, additional training of the neural network is required without changing the network structure.

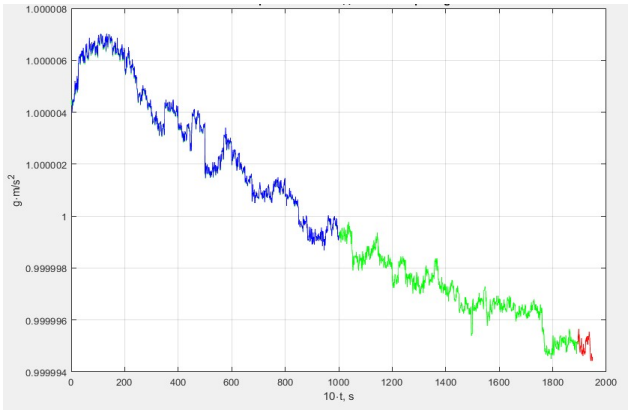


Fig. 7. Experimental, testing, and validation data of parameter measurements ~1 g

III. CONCLUSION

The proposed moving average method implemented by means of neural network can provide significant reduction of

the test times, ensure further prediction of experimental data and, ultimately, duplicate the method of natural storage widely used in the production of navigation instrumentation.

REFERENCES

[1] Izmailov, E.A., Modern Trends in the Development of Technologies of Inertial Sensing Elements and Aircraft Systems, Trudy FGUP NPTsAP, 2010, no. 1, pp. 27–35.

[2] Rivkin, B.S., Analiticheskii obzor sostoyaniya issledovanii i razrabotok v oblasti navigatsii za rubezhom (Analytical Review of the State of Research and Development in the Field of Navigation in Foreign Countries), St. Petersburg, CSRI Elektropribor, 2017–2023, iss. 1–8.

[3] Sinel'nikov, A.E., Nizkochastotnye lineinye akselerometry. Metody i sredstva proverki i graduirovki (Low Frequency Linear Accelerometers: Methods and Tools for Inspection and Calibration), Moscow, Izdatel'stvo standartov, 1979.

[4] Konovalov, S.F., Navigatsionnye akeselerometry. Uchebnoe posobie (Navigation Grade Accelerometers: Study Guide), Moscow, MGTU, 2023.

[5] Konovalov, S.F., Teoriya vibroustoichivosti akselerometrov (Theory of Vibration Resistance in Accelerometers), Moscow, Mashinostroenie, 1991.

[6] Nakhov, S.F., Kalikhman, D.M. et al., Opyt proektirovaniya i izgotovleniya blokov izmeritelei lineinogo uskoreniya na kvartsevykh mayatnikovykh akselerometrakh s analogovoi i tsifrovoi sistemami upravleniya (Experience in Design and Manufacture of Linear Acceleration Sensors Units on Quartz Pendulum Accelerometers with Analog and Digital Control Systems), Saratov, Yuri Gagarin State Technical University of Saratov, 2021.

[7] Kalikhman, D.M., Skorobogatov, V.V. et al., Quartz Pendulum Accelerometer with Digital Feedback as an Automatic Control System: Synthesis of Controllers and Software: Further Prospects, Navigatsiya i upravlenie letatel'nymi apparatami, 2019, no. 26, pp. 2–37.

[8] Grebennikov, V.I., Deputatova, E.A. et al., Pendulum Accelerometer with Digital Control and New Functional Capabilities, Izvestiya RAN. Teoriya i sistemy upravleniya, 2021, no. 2, pp. 73–95.

[9] Kalikhman, D.M., Samitov, R.M. et al., Flight Tests of the Six-Axis Unit of Specific Force Meters Based on Precision Quartz Pendulum Accelerometers with Digital Feedback as Part of the Soyuz MS-14 Spacecraft, 27th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2020, pp. 184–187.

[10] Gnusarev, D.S., Deputatova E.A., Kalikhman, D.M., Analysis of Noise Components in a Quartz Pendulum Accelerometer with Digital Feedback Amplifier, Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologii, mekhaniki, optiki, St. Petersburg, 2018, vol. 18, no. 6, pp. 1091 – 1098.

[11] Nikiforov, V.M., Zaitsev, S.A., Motion Control of Dynamic Systems Based on Neural Network Technologies, Trudy FGUP "NPTSAP". Sistemy i pribory upravleniya, 2011, no. 2.

[12] Nikiforov, V.M., Vashchenko, A.V., Grebenshchikov, D.V. et al., Application of Neural Network Technologies in Solving Problems of Terminal Control of Motion of Dynamic Systems, 8th International Science and Technology Conference on Gyrotechnology, Navigation, Motion Control and Aerospace Engineering, Kyiv, 2011.

[13] Nikiforov, V.M., Gladyshev, V.M., Malykhin, L.I., Chaikovskii, M.M., Terminal Control of Motion of Dynamic Systems Based on Neural Network Technologies, 7th International Aerospace Congress IAC12, Moscow, 2012.

[14] Ivashina, E.A., Korlyakova, M.O., Pilipenko, A.Yu., Neural Network Approach to the Formation of an Intelligent Video System of a Navigation Complex, 20th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2013.

[15] Nikiforov, V.M., Bakhmet'eva, O.V., Approximation of Scattering Fields of Synchronous Hysteresis Gyro Motor Using Neural Network, Trudy FGUP "NPTSAP". Sistemy i pribory upravleniya, 2015, no. 1.

[16] Nikiforov, V.M., Bespalova, N.M., Estimation of the Systematic Error of a Precision Angle Transducer Based on Neural Network Technologies, 3rd Russian Multiconference on Problems of Control, 26th Conference in Memory of Outstanding Gyroscopic Instruments Designer N.N. Ostryakov, 2010.

[17] Nikiforov, V.M., Bespalova, N.M., Dorokhin, Yu. P., Modeling of the Systematic Error of a Precision Angle Transducer and its Estimation Using Neural Networks, Trudy FGUP "NPTSAP". Sistemy i pribory upravleniya, 2010, no. 3.

Test Results of Compensation Pendulum Accelerometers

S. A. Volobuev
Concern CSRI Elektropribor, JSC
St. Petersburg, Russian Federation

E.V. Dranitsyna
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russian Federation
ORCID 0000-0003-2591-5816

Abstract – Concern CSRI Elektropribor has been developing pendulum-type accelerometers with torsion suspension since 2022. The design of such accelerometers allows avoiding expensive materials and complex technological operations during manufacturing, providing sufficient accuracy parameters for use in navigation grade systems. The report presents the design features of the developed accelerometers and results of dynamic tests of the manufactured samples, which showed high accuracy characteristics.

Keywords—sensors, compensation pendulum accelerometers, test results

I. INTRODUCTION

Concern CSRI Elektropribor has a vast experience in development, manufacture, and use of accelerometers in navigation systems as sensors of linear accelerations. The floating pendulum accelerometer on torsion suspension with damping fluid is used for precise systems to ensure zero buoyancy. The technological process of manufacturing these accelerometers contains a large number of labor-intensive operations. Since 2022, Concern CSRI Elektropribor has been developing a pendulum compensation accelerometer on torsion suspension (PCA) for use in navigation grade systems, the manufacture of which excludes the major part of labor-intensive operations and significantly reduces the price.

This paper describes the principle of operation and design features of the developed accelerometers and presents the results of dynamic tests of the manufactured samples, which showed high accuracy characteristics.

II. PCA DEVICE DESIGN

A. Structural features of the PCA

Structurally, the accelerometer consists of two parts: a sensitive element and a signal processing plate, mounted on a single base. The sensing element is covered by a screen. An optical pair (LED/photodiode) and a rod on the pendulum element are used as an angle sensor. A coil with a magneto-electric torque sensor winding suspended on a stretching string is used as the pendulum element. Fig. 1 shows a cross section of the 3D model of the PCA.

The sensing element is installed on strings 2 in the body 10 through the base plate 5. The sensing element is made in the form of a frame 6 with control windings and a rod 14. Strings are insulated from the body by ceramic pads 9 and serve as current conductors to the control windings. Strings are soldered to transition plates 1 also mounted on the body 10. The control windings together with permanent magnets 5 installed in the magnet cores 3 serve as a torque sensor compensating the acting acceleration. The four-section photodiode 18 and LED 12 act as an angle sensor. The pads 9 are pressed by screws 8 in the body 10. A thermistor 15 is mounted on the body 10 for temperature control. The case 16 is covered by a shield 20 and is attached to a base 17 with pins 16 through which power is supplied and output information is taken. The electronics plate is installed inside the base 17.

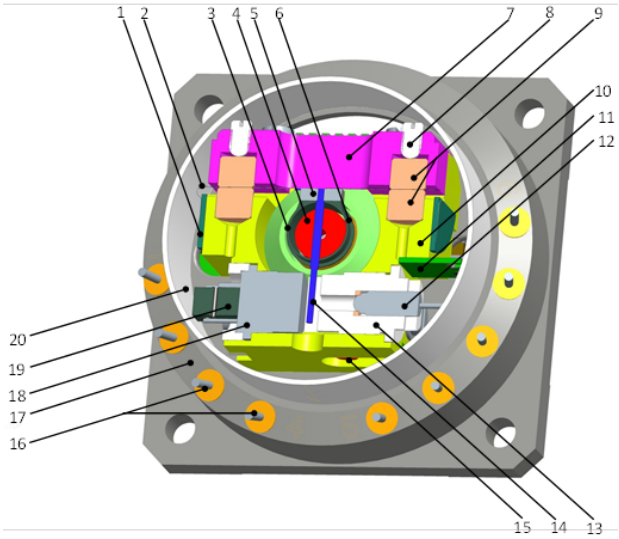


Fig. 1. 3D model of the PCA.

B. Principle of operation of the PCA

When the base is stationary (no accelerating motion), the rod of the movable part of the sensing element is between the photodiodes illuminated by the LED. As a result, the output currents of the photodiodes form equal voltages at the outputs of the amplifiers of the signal processing board. At the same time, there is a zero voltage at the outputs of the corrector and amplifier, creating a zero current in the torque sensor. When acceleration occurs, the rod moves, the currents of the photodiodes become unequal to each other and the current flows through the torque sensor, creating a magnetic field. It returns the rod to its original position between the photodiodes. In the signal processing plate at the output of the corrector due to the presence of an integrator voltage occurs, creating a current in the torque sensor, which is strictly proportional to the input acceleration. This current creates a voltage on the resistor, which is supplied through the repeater to the accelerometer output. At the adjustment stage, by selecting the elements in the signal processing plate, it is possible to achieve the required bandwidth, scale factors and bias of the accelerometer output signal [1]. Additionally, the PCA includes a temperature sensor consisting of a thermistor and an electronics board, which makes it possible to algorithmically take into account the dependency of the scale factor and bias of the accelerometer output signal on the temperature and its speed of change.

III. PCA TEST RESULTS

In order to assess the main accuracy characteristics, a series of tests were conducted on a pilot sample of PCA both separately and as part of the inertial measurement module in the operating range of linear accelerations and temperatures. Preliminary each sample was calibrated, which included determination of bias, scale factors and their dependencies on the temperature and its speed of change, measured by the built-in temperature sensor, as well as the angles of misalignment of the sensitivity axes. The calibration method is based on the method of comparison with the reference value, as a reference value the projections of the free fall acceleration on the device axes are taken, taking into account the transient accelerations caused by the rotation of the stand.

The main characteristics of the PCA obtained during the tests taking into account temperature compensation are summarized in Table I.

TABLE I. PCA CHARACTERISTICS

Parameter name, measurement units	**Value**
Measurement range, g	±5. no less
Bandwidth, Hz	220
The root of the noise power spectral density, $(m/s^2)/\sqrt{Hz}$	$3.5 \cdot 10^{-5}$
Measurement error at stable temperature[a], m/s^2	$5 \cdot 10^{-4}$
Scale factor, V/g	1.15±0.15
Scale factor instability in startup 24 hours[a], ppm	50, не более
Scale factor instability from startup to startup[a], ppm	±35, no less
Scale factor instability in the operating temperature range[b], ppm	±100, within
Bias instability in startup 24 hours (at 10 s averaging) [a], m/s^2	$\pm 2.5 \cdot 10^{-4}$, within
Bias instability from startup to startup[a], m/s^2	$\pm 5 \cdot 10^{-4}$, within
Bias instability at startup 24 hours in the operating temperature range (at 10 s averaging) [b], m/s^2	$\pm 7.5 \cdot 10^{-4}$, within
Instability of the measuring axis position relative to the mounting surfaces in the operating temperature range[b], "	±30

[a.] Temperature variation was maintained within ± 0,5°C;
[b.] Operating temperature range from 0 to 40°C

The bias and scale factor instabilities from startup to startup and in startup were determined at a stable temperature of 23°C on a three-axis motion simulator, the rotations of which allowed determining a set of bias and scale factors for each accelerometer installed in the inertial measurement module.

Figures 2-3 show the bias and scale factor estimates taking into account temperature compensation in the runs, accordingly. The plots show that the bias and scale factor instabilities from startup to startup do not exceed 0.0005 m/s^2 and 0.0035% (35ppm) and correspond to the instabilities in startup.

A distinct feature of the developed PCA devices is the presence of a stretching string and the possibility of its orientation in different positions relative to the direction of the gravity force vector during installation in the product or during operation. When changing the position of the string relative to the direction of the gravity force vector in the transformation model, the cross-coupling coefficients change significantly. Fig. 4.a shows the measurement errors of the PCA device of gravity acceleration at different positions of the string. In the graph, the blue color indicates the position when the string is placed across the gravity force vector, while the red color indicates the position along it.

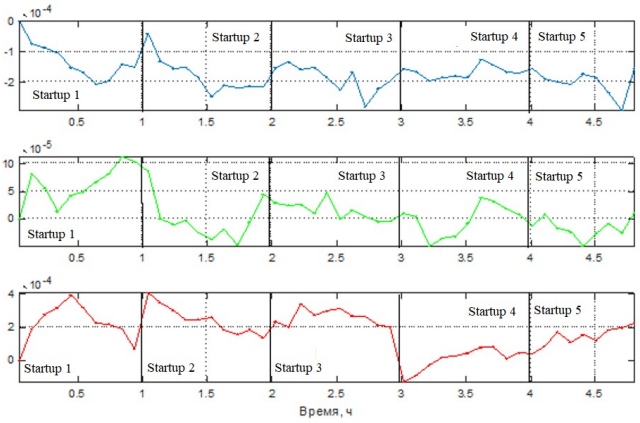


Fig. 2. PCA bias instability from startup to startup and in startup

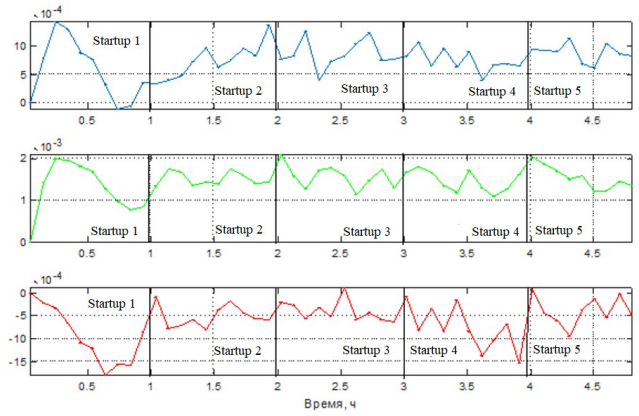


Fig. 3. PCA scale factor instability startup to startup and in startup.

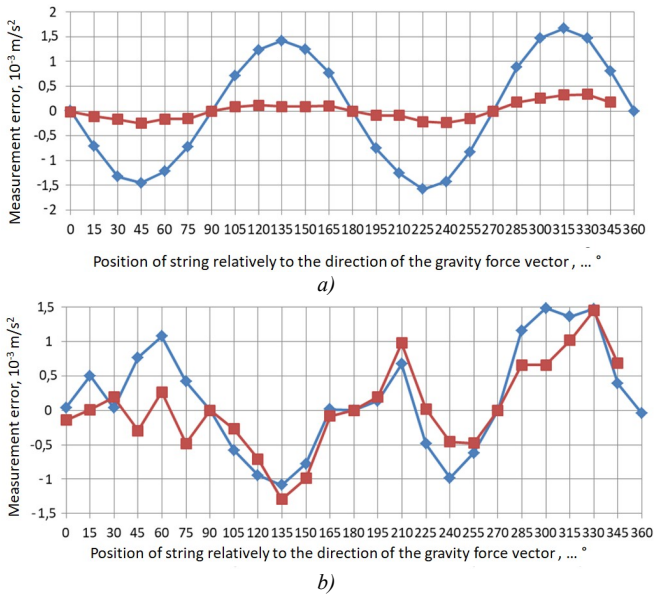


Fig. 4. Measurement error of the PCA device: a) without taking into account cross-correlations; b) taking into account cross-correlations.

However, the cross couplings are well compensated, Fig. 4.b shows the measurement errors after accounting for

them. The graphs show that the measurement error decreases by an order of magnitude and does not exceed $\pm 5 \cdot 10^{-4}$ m/s².

IV. CONCLUSION

The paper describes the design features of the newly developed accelerometers and test results of their prototypes. The tests have shown high accuracy of the developed accelerometers. Thus the bias instability from startup to startup and in startup at constant temperature is within $\pm 5 \bullet 10^{-4}$, m/s², and the scale factors are within 100 ppm. Tests of PCA devices as part of the inertial measurement module confirmed the possibility of building navigation grade systems based on them.

The obtained high accuracy of PCA devices made it possible to create a precise sensor of small inclination angles on its base with a significant reduction of the measurement range [2].

REFERENCES

[1] V.V. Sumarokov, E.A. Makhaev, E.A. Yakushina., Providing the Stability of the Pendulum Compensation Accelerometer and Development of its Adjustment Methodology // Proceedings of the 32 Saint Peterburg International Conference on Integrated Navigation Systems, St. Petersburg: Concern CSRI Elektropribor, JSC, 2025.

[2] E. A. Yakushina, S.G. Romanenko., The Influence of Design Features of the Auto-Compensation Device on the Errors of the Precise Small Tilt Angle Meter, Proceedings of the 32 Saint Peterburg International Conference on Integrated Navigation Systems, // Materials of XXXI Saint Peterburg international conference on integrated navigation systems St. Petersburg: Concern CSRI Elektropribor, JSC, 2025.

Providing the Stability of the Pendulum Compensation Accelerometer and Development of its Adjustment Methodology

V.V. Sumarokov
Concern CSRI Elektropribor, JSC
St. Petersburg, Russian Federation
biseg2@mail.ru

E.A. Makhaev
Concern CSRI Elektropribor, JSC
St. Petersburg, Russian Federation
machaev@yandex.ru

E.A. Yakushina
Concern CSRI Elektropribor, JSC
St. Petersburg, Russian Federation
yakushina.ev1997@yandex.ru

Abstract – The paper presents the features of the amplitude-frequency response of the compensatory-type pendulum accelerometer and the specific features of its tuning, considering the resonance frequencies occurring in the structure. In addition, the method of feedback tuning is presented, which minimizes the effect of technological variations of the assembly and to adjust the optimal amplitude-frequency response in terms of dynamic parameters.

Key words: accelerometer, amplitude-frequency response, dynamic stability.

I. Introduction

Providing stability of accelerometer performance is one of the most important tasks occurring at the stage of instrument tuning and regulation. There are several solution paths to achieve the required values [1,2]. The presented paper reflects two ways of solution working together: providing stability of the pendulum compensation accelerometer, at the stage of feedback circuit design and regulation of the parameters obtained after assembly of the sensing element.

II. Providing Accelerometer Stability by Feedback Correction

The compensating accelerometer of the pendulum type consists of an inertial mass m suspended on an elastic string (length L), an optical sensor of the inertial mass rotation angle, feedback and a torque sensor compensating the effect of the measured acceleration [3]. The amplitude-frequency response (AFR) of the sensitive element of such an accelerometer has two resonances: at frequencies f1 ≈ 8 Hz and f2 ≈ 600 Hz (the exact value is determined by the stretching tension and may differ by units of Hz). The stability of the accelerometer of such design is possible only at realization of the complex frequency correction in the feedback circuit and its further adjustment, allowing to compensate technological variations of parameters of the sensitive element.

The frequency response of the accelerometer and its components is shown in Figure 1. The AFR of the sensitive element (curve 1) in the operating frequency range of 20-200 Hz has a slope of 40 dB/dec. The feedback closure is simply closed looped with the closed-loop control gain, this results in excitation of the accelerometer. To achieve stability, a differentiator is introduced into the feedback (curve 2), the AFR of which creates an overall slope of 20 dB/dec in the frequency range of 20-200 Hz (curve 3). However, the use of the differentiator leads to an increase in the maximum of the AFR at the frequency f2 ≈ 600 Hz, and when the feedback is closed, the accelerometer is excited at the frequency f2.

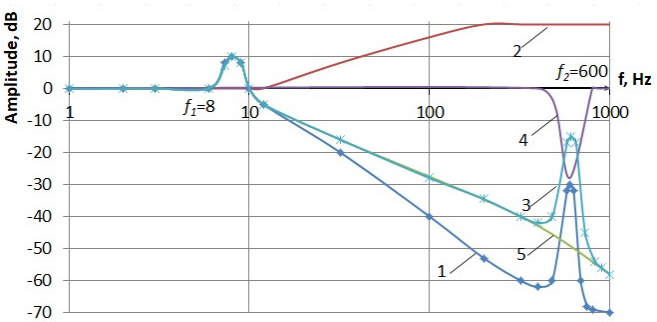


Fig. 1. Frequency response of the accelerometer: 1 - AFR of the sensing element, 2 - AFR of the differentiator, 3 - AFC of the corrected sensitive element, 4 - AFR of the barrier filter, 5 - AFR of the accelerometer with the barrier filter

To eliminate the influence of the arising resonance at frequency f_2 a barrier filter is introduced, the AFR of which (curve 4) removes the rise of the characteristic at frequency f_2 (curve 5) when it is accurately adjusted. In this case the accelerometer works with good quality at the bandwidth f_b=*100 Hz*. According to the results of accelerometer tests it is seen that the frequency of the maximum of the AFR f_2 changes significantly with heating, due to temperature variations, as well as in time. In this case, the frequency f_2 is equal to [3]:

$$f_2 = \frac{1}{2\pi}\sqrt{\frac{2F}{L \cdot m}}, \qquad (1)$$

where F is the tension force of the stretch

According to formula (1) the change of the frequency f_2 depends on the change of the tension force of the string, so the compensation of the maximum of the AFR of the sensing element at the frequency f_2 is not effective with the use of the barrier filter. The change of tension can occur when the ambient temperature changes - the frequency f_2 will change and will not be equal to the frequency of the barrier filter, as a result, excitation of the accelerometer can occur.

To ensure the stability of the accelerometer, instead of the barrier filter, a second-order low-pass filter with a cutoff frequency f_{sf}=300 Hz and an AFR decay of 40dB/dec, which has a simpler tuning, can be used. The AFR of such accelerometer and its components is shown in Figure 2. According to Figure 2, the use of a low-pass filter (curve 4) allows us to sharply reduce the maximum resonance at the frequency f_2 of the accelerometer's AFR (curve 1). In addition, the introduction of an integrator (curve 3) in the feedback circuit allows to reduce the error of constant acceleration

measurement and increase the measurement of vibration acceleration up to 3g.

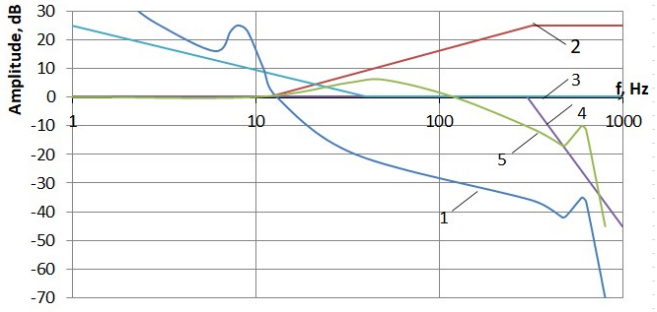


Fig. 2. Frequency response of the accelerometer: 1 - AFR of accelerometer with open loop feedback, 2 - AFR of the differentiator, 3 - AFR of the integrator, 4 - AFR of the low-pass filter, 5 - AFR of accelerometer with closed feedback

The implementation of the mentioned circuits with a gain of 30 dB and closing of the feedback makes it possible to obtain an accelerometer, the frequency response of which has a bandwidth of 200 Hz with a rise of 3 dB at a frequency of 60 Hz. The accelerometer is not tend to excite and is able to measure static acceleration up to 5g, and vibration acceleration up to 3g.

III. Ensuring the Stability of the Accelerometer by Aligning the Signals of the Angle Sensor

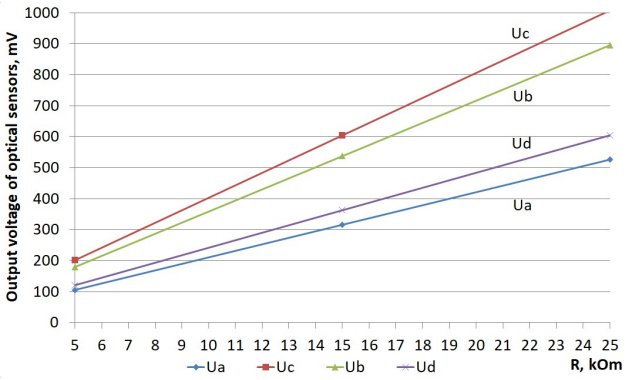
When manufacturing and assembling the sensing elements of a series of devices there is a variation of technological parameters, so in addition to ensuring the stability of the accelerometer must achieve a reduction in the variation of these parameters, which can also be realized by adjusting the feedback elements.

At the stage of development it was supposed to adjust the device by bringing its AFR to the required form by adjusting only the gain of the device. However, as test results of the first samples showed the parameters of the optical sensor, depending on the inequality of light fluxes, non-ideal geometry of gaps and pendulum, affect the static and dynamic characteristics of the device in a much greater degree than it was imagined at the modeling stage. Since there is a large variation of parameters of the input stage of the input electronics of the angle meter with an optical sensor, the adjustment technique became more complicated, time-consuming and labor-intensive, which made it absolutely unsuitable for a mass-produced device.

Therefore, it was necessary to adjust the input stage of the pendulum compensating accelerometer electronics, which provides equalization of the angle sensor signals. Application of this technique allows to adjust the accelerometer regardless of the variation of parameters of the optical sensor and allows the simplest way to bring the AFR of the devices to the desired form, thus adjusting the dynamic characteristics of the accelerometer in the entire range of operation.

Regulation is conducted on the device with open feedback on the signals of output voltages of two optical sensors in the extreme positions of the rod in the gap of the optical sensor (the rod is a component of the accelerometer pendulum overlapping the light streams). In the positions of measurement g and -g it is necessary to capture the dependency of the voltage change of the optical sensors on the resistance of the input stage amplifiers and we obtain four equations of the dependency of the output voltages on the corresponding resistances:

$$\begin{aligned} U_a &= U_{fd1} = f(R_{opt.s1}), \\ U_b &= U_{fd1} = f(R_{opt.s1}), \\ U_c &= U_{fd2} = f(R_{opt.s2}), \\ U_d &= U_{fd2} = f(R_{opt.s2}), \end{aligned} \qquad (2)$$



where U_{fd1}, U_{fd2} – voltages of the first and second optical sensors in the positions of acceleration measurement $-g$ or $+g$; $R_{opt.s1}$, $R_{opt.s2}$ – resistances of amplifiers of input stages of the first or second optical sensor (Figure 3). The presented dependencies have a linear form, so they can be represented by classical linear equations of the following type:

$$y = k \bullet x + b.$$

Then it is necessary to take the dependency of the working point - the point of dependencies intersection of the optical sensors output voltages on resistances:

$$e = U_{fd1} = U_{fd2} = f(R_{opt.s1}, R_{opt.s2}). \qquad (3)$$

The conditions for equality of output voltages of the input stage shoulders over the entire measurement range are the equation:

$$U_a - U_c = U_b - U_d. \qquad (4)$$

Such a ratio allows to provide a symmetrical response of the device under dynamic influences (Figure 4). The deviation of the inertial mass from the working point in any direction leads to the same signal increase at the output of the optical sensor.

Solution of Equations (3) and (4) gives an unambiguous dependence of the operating point on the resistances in the optical sensor arms. The optimal shape of the AFR curve is achieved by selecting the operating point and actually by selecting the gain, but at the same time with the regulation of the total gain there is also an alignment of the shoulders of the input stages of optical sensors which makes it possible to bring the AFR of the device to the form as in Figure 2 (curve 5) without complex regulation of the correction circuits.

Experimentally shown that the device obtains optimal characteristics in terms of overload capacity and stability in the whole frequency range when the amplitude of the signal in the closed system at frequency 210 Hz is 2.5±0.1 times higher than the amplitude at frequency 10 Hz. It is enough for the regulator to achieve the required shape of the AFR by controlling the amplitude at the frequency of 210 Hz, selecting with the help of resistors Ropt.d1 and Ropt.d2 the gain of the

input stage under the selected operating point (smoothly increasing the value of the operating point). The system of equations (3) and (4) allows us to determine such values of real elements on the board Ropt.d1 and Ropt.d2 and provide good dynamics due to the alignment of the optical sensor arms and at the same time to select the maximum gain at which there is no excitation of the control system (operating point).

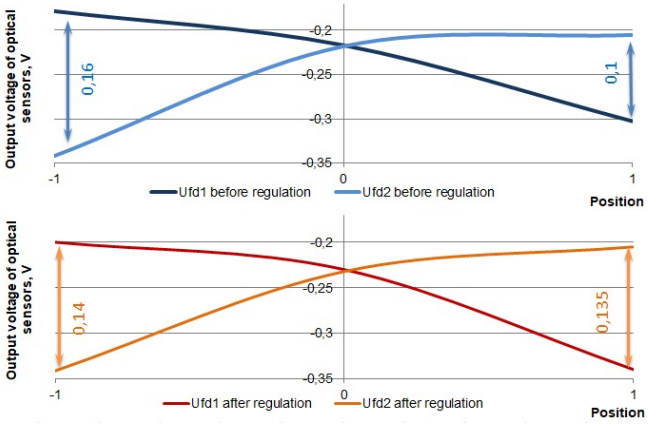


Fig. 4 Output signal of optical sensors at rod deflection before and after regulation

IV. Conclusion

The paper presents the features of the amplitude-frequency response of the compensatory-type pendulum accelerometer, its components, as well as the features of its tuning considering the resonance frequencies. In addition, the method of regulating the input feedback stage of the accelerometer is proposed. Its application allows to ensure the alignment of light fluxes of the angle sensor and to adjust the accelerometer regardless of the variation of the optical sensor parameters. In addition, regulation of the input stage of electronics, which converts the signal from the optical sensor, allows the simplest way to bring the AFR of the devices to the desired form, thereby adjusting the dynamic characteristics of the accelerometer in the entire frequency range.

References

[1] Konovalov S.F. Gyroscopic systems. Designing of gyroscopic systems. III. Accelerometers, angular velocity sensors, integrating gyroscopes and gyrointegrators: Textbook / E.A. Nikitin [et al.]. Edited by D.S. Pelpor. Moscow: Higher School, 1980. 128 c.

[2] Nakhov, S.F. Design and manufacturing experience of the linear acceleration meter units on the quartz pendulum accelerometers with the analog and digital control systems (in Russian) / D.M. Kalikhman [et al.]. Saratov: Izd-vo SSTU named after Gagarin Y.A., 2021. 240 c.

[3] Melnikov, V.E.; Melnikova, E.N.; Chernomorsky, A.I.; Goncharenko, G.G. Inertial information sensors. Textbook / Edited by A.I. Chernomorsky. Moscow: MAI PRINT Publishing House, 2011. 356 c.

The Influence of Design Features of the Auto-Compensation Device on the Errors of the Precise Small Tilt Angle Meter

E.A. Yakushina

Concern CSRI Elektropribor. JSC
St. Petersburg. Russian Federation
yakushina.ev1997@yandex.ru

S.G. Romanenko

Concern CSRI Elektropribor. JSC
St. Petersburg. Russian Federation
biseg2@mail.ru

Abstract – The paper deals with a small tilt angle meter (STAM) intended to be used in laboratory conditions to control the position of surfaces on which the calibration and testing of sensitive elements in navigation systems is carried out. The influence of the autocompensation device and its design on the accuracy of tilt angle determination is evaluated. The model of errors of the autocompensation device and the method of calibration of its coefficients are presented.

Key words: accelerometer, small inclination angle meter, rotary device, error model.

I. Introduction

The Small Tilt Angle Meter (STAM) is designed for use in laboratory conditions to control the position of base planes on which navigation systems and their sensitive elements are installed during testing and calibration. The use of STAM allows to separate the changes in the tilt of the base surface of the foundations during testing from the errors of the output signal of the investigated instrument during long-term tests. Nowadays, high accuracy requirements are imposed on modern instruments. On this basis, the accuracy of the STAM should be several tenths of arcsec, and the range of angle measurement should be (±600) arcsec. The paper presents an error model of a rotary device, a method for calibrating the coefficients of the obtained model and their stability.

II. Design of the STAM

The design of the small tilt angle meter is presented in Figure 1. In the considered design of the device the sensitive elements are two pendulum accelerometers of compensatory type (2), the sensitivity axes of which are mutually perpendicular, which allows to control tilt along two axes simultaneously.

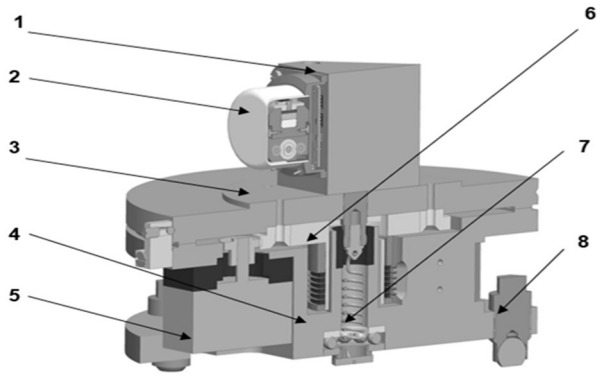


Figure 1. Design of the STAM
1 - base bracket; 2 - accelerometers; 3 - rotary platform; 4 - base; 5 - step motor; 6 - rotation gear; 7 - tensioning device; 8 - adjusting screws.

Measurement of tilt angles is provided by the auto-compensation device which includes a rotary platform (3) moving on the contact surface of the device base (4). Rotation of the rotary platform is carried out by a step motor (5). The use of the rotary platform allows simultaneous measurements in two axes for a long time and also compensates the accelerometer zero shift which allows increasing the accuracy of tilt angle measurement.

The value of the tilt angle of the controlled surface is detected in two opposite positions of the measurement axis (0–180°) and is determined as a half difference of the previously obtained values [1, 2]. The rotary device, as a separate complex design element, enable to compensate part of the errors of the sensing element, but at the same time introduces its own errors in the tilt angle measurement.

III. Model of Errors of the Autocompensation Device

A. Components of the error model

Let us represent the design of the rotary device as a combination of its components each of which contributes an additional tilt angle to the measured value. Therefore, it is necessary to take into account each of the components, their contribution to the tilt angle determination, to evaluate their stability and compensability during rotation. The arrangement of the planes of the construction elements described below is presented in Figure 2 on the example of one plane.

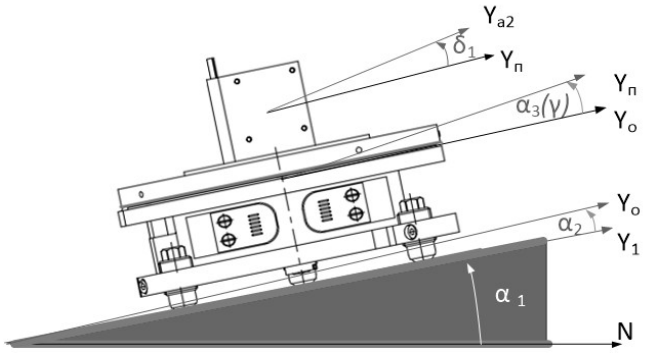


Figure 2. Planes of construction elements and angles between them.

ENh – horizon coordinate system with geographic orientation,

$X_1Y_1Z_1$ – coordinate system of the controlled surface,

α_1, β_1 – angles of controlled surface tilt (angles of transition from ENh coordinate system to $X_1Y_1Z_1$),

$X_oY_oZ_o$ – coordinate system associated with the surface of the STAM base,

α_2, β_2 – transition angles from the coordinate system $X_1Y_1Z_1$ to $X_oY_oZ_o$,

$X_пY_пZ_п$ – coordinate system associated with the sliding surface of the STAM rotary platform,

c(γ), *d(γ)* – angles of tilt of the rotary platform resulting from rotation by angle γ (angles of transition from coordinate system $X_oY_oZ_o$ to $X_пY_пZ_п$),

γ – the rotation angle of the rotary platform,

$X_{a1}Y_пZ_п$ – coordinate system associated with the sensitivity axis of the first accelerometer,

$X_пY_{a2}Z_п$ – coordinate system associated with the sensitivity axis of the second accelerometer,

δ_2, δ_2 – angles of non-orthogonality of the accelerometers sensitivity axes and measurement axes.

For the following analysis. we introduce the following assumptions and notations:

- $\alpha_1, \alpha_2, \beta_1, \beta_2 = const$,
- $\alpha = \alpha_1 + \alpha_2$, $\beta = \beta_1 + \beta_2$,
- $\beta < 1°$, $\alpha < 1°$, thus $\sin(\beta) \approx \beta, \cos(\beta) \approx 1$, $\sin(\alpha) \approx \alpha, \cos(\alpha) \approx 1$.

Taking into account the accelerometers error model [3, 4], the projections of free fall acceleration on the axes of accelerometers' sensitivity will have the following form:

for the first accelerometer:

$$\begin{aligned} a^x_{A1} &= a^0_X \cdot K_x + K_x \cdot g \cdot [\cos\gamma \cdot (\beta - \alpha \cdot \delta_1) + \\ &+ \cos\gamma \cdot (\beta_3(\gamma) - \alpha_3(\gamma) \cdot \delta_1) - \\ &- \sin\gamma \cdot (\alpha + \beta \cdot \delta_1) - \sin\gamma \cdot (\beta_3(\gamma) \cdot \delta_1 + \alpha_3(\gamma))] + \xi_{1c\text{п}}, \end{aligned} \quad (1)$$

for the second accelerometer:

$$\begin{aligned} a^y_{A2} &= a^0_Y \cdot K_y + K_y \cdot g \cdot [-\cos\gamma \cdot (\alpha + \beta \cdot \delta_2) - \\ &- \cos\gamma \cdot (\alpha_3(\gamma) + \beta_3(\gamma) \cdot \delta_2) + \\ &+ \sin\gamma \cdot (\alpha \cdot \delta_2 - \beta) + \sin\gamma \cdot (\alpha_3(\gamma) \cdot \delta_2 - \beta_3(\gamma))] + \xi_{2c\text{п}}. \end{aligned} \quad (2)$$

where K_X, K_Y – transformation coefficients, a^0_X, a^0_Y – zero shifts, $\xi_{1c\text{п}}$, $\xi_{2c\text{п}}$ – random component of the signal.

From expressions (1), (2) we can see that the required angles of tilt of the controlled surface are contained in the first and third summands in square brackets, the second and fourth summands describe the error of the rotary device.

When the platform rotates in the range of γ=0°...360°, as will be shown below, the STAM output signal has a complex periodic form containing n harmonic components:

$$\begin{aligned} a^x_{A1}(\gamma) &= \sum_{i=0}^{n} (A_{ci} \cos i\gamma + A_{si} \sin i\gamma), \\ a^y_{A2}(\gamma) &= \sum_{i=0}^{n} (B_{ci} \cos i\gamma + B_{si} \sin i\gamma), \end{aligned} \quad (3)$$

where $A_{ci}, A_{si}, B_{ci}, B_{si}$ – amplitudes of cosine and sine harmonic components of accelerometer output signals (taking into account conversion coefficients K_X, K_Y).

Taking into account expressions (1-3) the angles *α₃(γ), β₃(γ)* can be represented in the following form:

$$\begin{aligned} \alpha_3(\gamma) &= \sum_{i=0}^{n-1} (c_{ci} \cos i\gamma + c_{si} \sin i\gamma), \\ \beta_3(\gamma) &= \sum_{i=0}^{n-1} (d_{ci} \cos i\gamma + d_{si} \sin i\gamma), \end{aligned} \quad (4)$$

where $c_{ci}, c_{si}, d_{ci}, d_{si}$ – amplitudes of cosine and sine harmonic components of angles *α₃(γ), β₃(γ)*.

Expression (4) describes the interaction of the contacting surfaces and is caused by their design (technological allowance).

Since the tilt angle is estimated by measurements in two opposite positions of the measuring axis (for example, in the positions 0° and 180° of the measuring axis X), the half difference of these values enables us to determine the value of the measured tilt angle of the axis (taking into account expression 4):

$$\begin{aligned} \frac{a^x_{A1}(0°) - a^x_{A1}(180°)}{2K_x g} &= \beta - \alpha \cdot \delta_1 + \left(d_{c0} + d_{c2} + d_{c4} + ...\right) - \\ &- \delta_1 \cdot \left(c_{c0} + c_{c2} + c_{c4} + ...\right) + \xi_1. \end{aligned} \quad (5)$$

Expression (5) shows that the error in determining the angle *β* includes the amplitudes of the even harmonics of the angles *α₃(γ), β₃(γ).* A similar result applies to the angle α.

Considering expressions (1-4) the angles of tilt of the controlled surface will be represented as

$$\begin{aligned} \beta_{акс1} &= A_{c1} + \frac{A_{c3}}{2} + \frac{A_{c5}}{2}, \quad \beta_{акс2} = B_{s1} + \frac{B_{s3}}{2} + \frac{B_{s5}}{2}, \\ \alpha_{акс1} &= A_{s1} + \frac{A_{s3}}{2} + \frac{A_{s5}}{2}, \quad \alpha_{акс2} = B_{c1} + \frac{B_{c3}}{2} + \frac{B_{c5}}{2}. \end{aligned} \quad (6)$$

Thus, the amplitudes of harmonic components of the accelerometer output signal are the coefficients of the error model of the autocompensation node. Calibration of these coefficients before testing consists in determining the number of harmonics and their amplitudes for the purpose of their subsequent application in measurements. It is necessary to evaluate the stability of the obtained coefficients and the time interval for their redetermination.

B. Results of error model calibration

To evaluate the effect of the auto-compensation unit on the accuracy of tilt angles, the tests were carried out, which included measurements of tilt angles from the 360° platform rotation angle γ with a sampling rate of 1°. The measurement time in each position was 40 seconds. Harmonic components were extracted from the obtained experimental curves (Figure 3) which have a non-sinusoidal shape, using the least squares method. The obtained values of harmonic amplitudes (Table 1) show that the first six harmonic components are predominant in the accelerometers output signal. The first harmonic contains a useful signal (the angle of inclination of the controlled surface); the subsequent ones are related to the “fight” of the rotating platform (3).

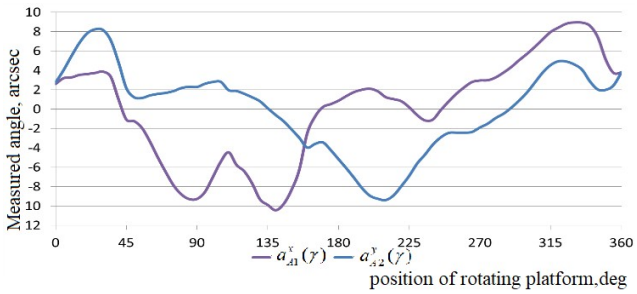


Figure 3. Dependency of angle change γ of the rotating platform tilt on the rotation angle γ

TABLE I. AMPLITUDES OF HARMONIC COMPONENTS OF ACCELEROMETER OUTPUT SIGNALS

A_{c1}	A_{s1}	A_{c2}	A_{s2}	A_{c3}	A_{s3}	A_{c4}	A_{s4}	A_{c5}	A_{s5}	A_{c6}	A_{s6}
2.8	-8.7	-2.5	0.1	-1.1	-0.5	0.2	0.9	-1.9	-0.5	-0.3	0.1
B_{c1}	B_{s1}	B_{c2}	B_{s2}	B_{c3}	B_{s3}	B_{c4}	B_{s4}	B_{c5}	B_{s5}	B_{c6}	B_{s6}
8.6	1.5	-0.6	-1.7	0.5	1	-0.3	0.3	-1.2	1.2	-0.6	0.7

According to expression (6) and the data in Table I. the tilts of the laboratory foundation were estimated by two accelerometers for each axis:

$$\beta_{acc1} = 2.3 \text{ arcsec}, \quad \beta_{acc2} = 2.6 \text{ arcsec},$$
$$\alpha_{acc1} = 8.7 \text{ arcsec}, \quad \alpha_{acc2} = 8.6 \text{ arcsec}.$$

The maximum error was 0.3 arcsec.

The evaluation of time stability of the obtained coefficients was carried out during three months of continuous operation. For the comparative evaluation, an electronic Zeromatic 2/2 gauge was installed on the same plate parallel to the STAM measurement axes. Figure 4 shows the results of the change in the tilt angle β along the X axis obtained from two accelerometers of the STAM and the electronic meter Zeromatic 2/2, as well as the graph of the change in the ambient temperature measured by the temperature sensor of the device Zeromatic 2/2.

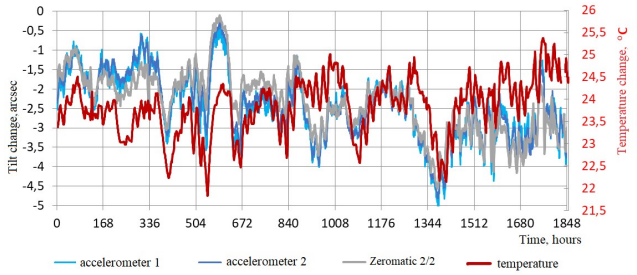


Figure 4. Changes in slope angle β and in ambient temperature

Figure 4 shows that the change in the tilt angle of the laboratory foundation over the three months varies in the range of 5 arcsec which is primarily due to the change of ambient temperature. The nature of the changes recorded by the three meters coincides. For the STAM accelerometers the difference in measurement along one axis does not exceed ±0.5 arcsec (Figure 5).

At the end of the long-term tests. the amplitudes of the harmonic components of the accelerometer output signals were re-evaluated (Table II).

TABLE II AMPLITUDES OF HARMONIC COMPONENTS OF ACCELEROMETER OUTPUT SIGNALS (RE-EVALUATION)

A_1	A_{s1}	A_{c2}	A_{s2}	A_{c3}	A_{s3}	A_{c4}	A_{s4}	A_{c5}	A_{s5}	A_{c6}	A_{s6}
2.8	-8.2	-3.0	-0.2	-0.9	-0.8	-0.2	0.9	-1.6	0	-0.44	0.59
B_{c1}	B_{s1}	B_{c2}	B_{s2}	B_{c3}	B_{s3}	B_{c4}	B_{s4}	B_{c5}	B_{s5}	B_{c6}	B_{s6}
8.2	1.5	-1.1	-1.4	0.2	1	-0.7	0.5	-1.7	1.2	-0.7	0.37

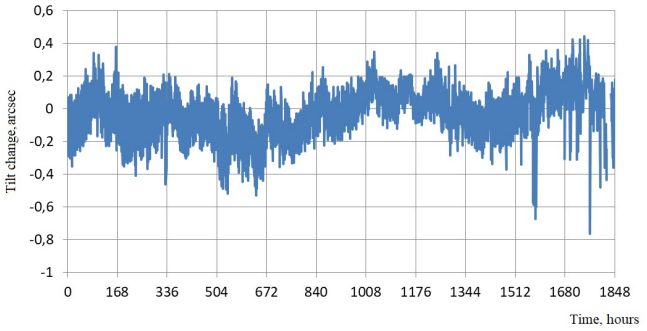


Figure 5. Difference in X-axis tilt angle measurement by STAM accelerometers readings

The comparison of Tables I and II shows that after three months, the values of the change in the coefficients of the error model of the autocompensation device does not exceed 0,5 arcsec. This confirms the stability of the coefficients and the possibility of practical application in order to reduce the error of slope angle estimation.

IV. CONCLUSION

The paper presents the results of the study of the design features of the rotary device of the small tilt angle meter. the application of which is relevant in the testing of navigation systems. The STAM makes it possible to measure the inclination angles of laboratory foundations with high accuracy and stability and evaluate their change during long-term tests. Taking into account the harmonic components of errors allows the accuracy of measurements to be significantly increased. The developed model and calibration technique can be used for further improvement of similar measurement systems.

REFERENCES

[1] Zeromatic dual-axis inclination sensor. [Электронный ресурс] URL: http://Wylerag.com (дата обращения 19.09.2022)

[2] Jun Luo. Zhiqian Wang. Chengwu Shen. et al.. Measurement science review. Volume 15. No. 5. 2015

[3] Konovalov. S.F. *Giroskopicheskie sistemy. Proektirovanie giroskopicheskikh sistem. Chast' III. Akselerometry. datchiki uglovoi skorosti. integriruyushchie giroskopy i girointegratory* (Gyroscopic systems. Designing of gyroscopic systems. Part. III. Accelerometers. angular velocity sensors. integrating gyroscopes. and gyrointegrators: Textbook. Edited by D.S. Pelpor. Moscow: Higher School. 1980.

[4] IEEE Std1293-2018- IEEE Standard Specification Format Guide and Test Procedure for Linear Single-Axis. Nongyroscopic Accelerometers

Determining the Characteristics of Motion Simulators Used in Navigation Equipment Tests

N. L. Yavorovskaya

JSC Concern CSRI Elektropribor,
St. Petersburg, Russia
metrolog@eprib.ru

T. N. Siraya

JSC Concern CSRI Elektropribor,
St. Petersburg, Russia
metrolog@eprib.ru

***Abstract* — The measurement problems of determining the dynamic characteristics of motion simulators are considered. An optical method for determining the dynamic characteristics of rotary devices was developed, and it was tested using a high-speed laser autocollimator. The paper also examines the prospects for the development of equipment and methods for determining dynamic characteristics using the optical methods, including the development of digital autocollimator for accurate angular measurements in dynamic mode.**

Keywords — motion simulator, rotating devices, rocking test bench, static characteristic, dynamic characteristics, measurement procedure, error, digital autocollimator

I. Introduction

Nowadays various motion simulators, which realize the motion with specified characteristics, are widely used in testing of navigation equipment. This group includes several kinds of devices, such as:

– rocking test benches,
– various rotating devices,
– centrifuges.

In order to ensure control of precise navigation equipment parameters it is necessary to carry out certification of motion simulators as test equipment [1, 2]. Thus, it is urgent to determine both static and dynamic characteristics of the simulators with high accuracy. According to [1, 2], the accuracy of the test bench should be only estimated by independent external devices that are certified as measuring instruments.

Moreover, the demands for the list and accuracy of simulators characteristics are steadily growing. Thus in practice still more refined and accurate characteristics are needed. As an example we can mention the control of the error of angular velocity set in the certification of rotary devices.

An approach to solve this problem involves the use of the precise optical instruments, which measure angle in dynamic mode. This method allows to estimate the errors and determine stability of the angular velocity within a period. Similar methods also allow to estimate the errors in setting angular amplitude of the rocking test benches.

The paper analyzes the methods and optical devices for determining the dynamic characteristics of motion simulators. The prospects for the development of similar methods are also considered.

II. Optical methods for determining the dynamic characteristics of motion simulators

The main types of the motion simulators, used for testing navigation devices, with main characteristics, are presented in Table 1.

Table 1. Characteristics of motion simulators

Parameter	Symbol	Error limits
1-3-Axis Rate Tables (positioning and constant speed rotation modes)		
Position accuracy	Δα	±2" (±5")
Rate accuracy over 360°	ΔΩ	±0,05 %
Rate stability over 360°	δΩ	±0,005 %
Rate accuracy over 10°	ΔΩ	±0,05 %
Rate stability over 10°	δΩ	±0,005 %
Rocking test benches (Positioning and harmonic oscillation modes)		
Position accuracy	Δα	±1' (±5')
Error of the period setting	ΔT	±10 %
Error of the amplitude setting	ΔA	±30'

The control of the simulators characteristics is often performed by means of indirect or joint measurement procedures of several parameters, for example, angles and frequency.

In order to use the procedures for testing equipment it is necessary to perform their metrological certification according to standard [3], first of all, to obtain a priori estimates of measurement errors.

Some specific features of the certification for this kind of procedures are discussed in [4, 5]. However, the scope of these methods is rather limited, as most of them provide just estimates of the average speed per revolution.

An approach to solve this problem involves the use of the precise optical instruments, such as autocollimators. Various types of autocollimators are generally used to control the positioning error of rotary devices. The most common models currently included in the Federal Information Fund of measuring instruments, and the main characteristics are listed in four columns of Table 2.

The Table 2 contains the main characteristics of the following models of autocollimators:

1 – Autocollimators АК-У (АК-01У, АК-05У, АК-1У) («Diagnostika» , Russia);
2 – Autocollimator АФ-0,05Ц («NPZ», Russia);
3 – Digital autocollimator TriAngle TA 500-57 («Trioptics GmbH», Germany);
4 – Laser autocollimators TriAngle TA HS 500-57 («Trioptics GmbH», Germany);

5 – Digital dynamic autocollimator (on development in JSC Concern CSRI Elektropribor, Russia).

TABLE 2. CHARACTERISTICS OF AUTOCOLLIMATORS

Parameter	Main characteristics of autocollimators				
	1	*2*	*3*	*4*	*5*
Angle range in horizontal plane	±720″	±20′	1240″	1240″	±20′
Angle range in vertical plane	±1020″	±20′	980″	980″	±20′
Error limits (accuracy)	±1,0″ in ±300″: ±0,1″ … ±0,5″	0,5″ (RMS)	±0,4″	±2,0″	±0,5″
Reading step	-	0,05″	0,02″	0,02″	0,02″
Frequency, Hz	-	-	160	10000	160
Frequency error	-	-	-	-	10^{-4}
Maximal distance for object, m	-	20	-	-	10

The developed method allows, using an autocollimator and a prism located in the center of rotation, to evaluate simultaneously both the static and dynamic characteristics of the rotary device.

Several series of experiments were carried out to verify the method. Most experiments have been carried out with photoelectric and laser autocollimators, TriAngle TA 500-57 and TriAngle TA HS 500-57, which provide information about the angular position of the article under test with a frequency from 1 to 10000 Hz.

The autocollimator was placed on a fixed base near the rotary table. A mirror prism, face nonflatness of which was no more than 0,0003 mm, was fixed in the center of the table. The autocollimator was placed so as to form an optical link between its axis and the face of the mirror prism. A scheme of measurements of the parameters mentioned above is shown in Fig. 1.

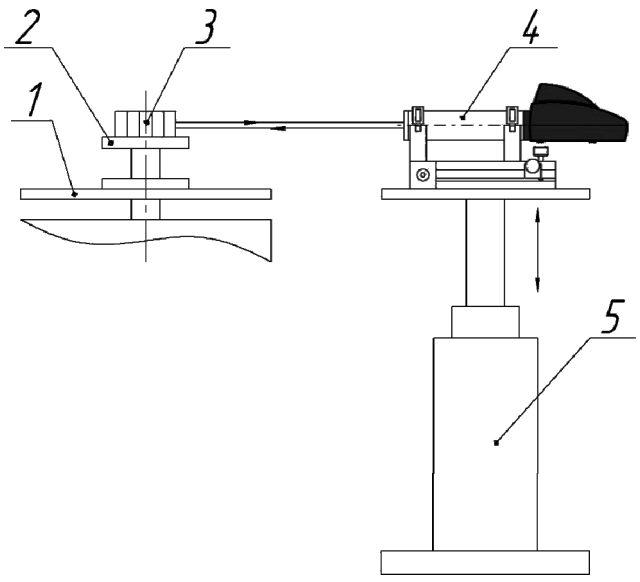


Fig. 1. Measuring the error in setting the period and rate: 1 – motion simulator; 2 – prism insertion device' 3 – mirror prism; 4 – autocollimator; 5 – fixed base.

While the mirror prism plane is in the autocollimator's field of view, varying readings on its horizontal scale are recorded.

Then the regions of implementation are identified in which the readings change linearly. These regions correspond to the passage of the mirror surface through the autocollimator's field of view.

Based on the processing of the linear regions of implementation, it is possible to determine the angular velocities over a period, and when using a prism, also the velocities within a revolution.

In addition, when using a multi-faceted prism, it is possible to evaluate the error in setting the pitching amplitude and other motion parameters during the certification of the motion simulator. A form of the signal from autocollimator during rocking test bench control is presented in figure 2.

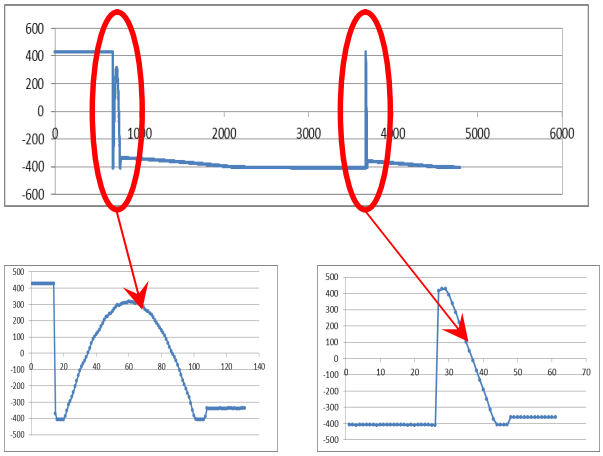


Fig. 2. Implementation of the signal during rocking test bench certification in dynamic mode

There are two steps in data processing. The first step is revealing the significant points, corresponding to successive regions. The second step includes the data processing of the sample obtained at the first step.

The processing of the measurement results showed that the relative error in setting the speed of the rotary device in the range from 30 to 720 °/s is within the limits ± 0,025 %.

It should also be noted that the main component of the method error is due to the sampling interval. Therefore, the laser autocollimator (based on the specified measurement range and the signal recording frequency) allows one to estimate the dynamic parameters in the speed range up to 720 °/s.

At the same time, the specified autocollimator is not certified for frequency and speed measurements. Therefore, it cannot be used for test bench certification.

Currently, JSC Concern Elektropribor is working on creating an optical measuring device - a digital dynamic autocollimator, which will allow for accurate angular measurements in dynamic mode. The characteristics of the autocollimator are given in the last column of Table 2. These are similar to the characteristics of the digital autocollimator TriAngle TA HS 500-57, but in addition the new autocollimator is standardized on frequency error.

This autocollimator is to be certified as a measuring instrument for angle, frequency and angular velocity measurements. When developing such a measuring instrument it is necessary to solve a number of problems, including:

– determine and standardize the static and dynamic characteristics of the measuring instrument (according [8]);

– develop the procedures for the static and dynamic characteristic control;

– conduct the tests to approve the type of measuring instrument and introduce it in the Federal Information Fund of Measuring Instruments.

Thus, these measuring instruments are relevant for certification of many motion simulators required for testing of precise navigation equipment.

Besides, such measuring instruments can be directly applied in tests of navigation equipment, for example, when estimating the stabilization systems parameters.

CONCLUSION

1 An analysis of measurement problems for determining dynamic characteristics of motion simulators reveals the necessity for development of the measurement instruments and procedures in dynamic mode.

2 An optical method for determining the dynamic characteristics of rotary devices is developed, and it is tested using a high-speed laser autocollimator. The method demonstrates high accuracy; but the measuring instruments should be certified for dynamic mode.

3 The development of digital dynamic autocollimator for accurate angular measurements and its certification in dynamic mode allows for solving the problems of determining the dynamic characteristics of motion simulators and navigation equipment.

REFERENCES

[1] GOST R 8.568-2017 State system for ensuring the uniformity of measurements. Verification of testing equipment. General provisions.

[2] GOST R 51672-2000 Metrological ensuring of product testing for the assurance of conformity. General principles.

[3] GOST R 8.563-2009 State system for ensuring the uniformity of measurements. Procedures of measurements.

[4] K. S. Gorohovsky, N. L. Yavorovskaya, O. A. Yakovenko. Testing of precise motion simulators – Sixth All-Russian conference «Measurements and testing in shipbuilding (Sudometrica 2016)», St. Petersburg, 2016, p. 125-127.

[5] V. V. Belyaeva, M. D. Kudryavtsev, T. N. Siraya, N. L. Yavorovskaya. Certification of measurement procedures in case of parametric errors. - Sixth All-Russian conference «Measurements and testing in shipbuilding (Sudometrica 2016)», St. Petersburg, 2016, p. 151-157.

[6] D. G. Gryazin, M. D. Kudryavtsev, N. L. Yavorovskaya, K. N. Usachev, F. Bellon, D. Rames, and J. Perdriat. Testing a C40-ST-09 Actidyn Systemes Centrifuge. - Gyroscopy and Navigation, 2013, Vol. 4, No. 4, pp. 233–237.

[7] N. L. Yavorovskaya, T. N. Siraya. Metrological assurance in development of navigation equipment. - 15 All-Russian conference «Metrological assurance in defence and security of Russian Federation», Povedniky, Moscow reg., 2024.

[8] ГОСТ 8.256-77 State system for ensuring the uniformity of measurements. Standartization and determination of dynamic characteristics of analogue measuring instruments. Basic principles.

Three-Axis Simulation and Calibration Test Bench with Inertial Sensing Elements in Three Rotational Axes

D. M. Kalikhman, E. A. Deputatova, N. V. Tarakanov
Yuri Gagarin State Technical University of Saratov,
Russia, Saratov;
Branch of Academician Pilyugin Center – Production Association Korpus,
lidkalihman@yandex.ru
deputatova@bk.ru
nikolaytarakanov87@gmail.com

Yu. V. Chebotarevskii
Yuri Gagarin State Technical University of Saratov,
Russia, Saratov

V. M. Nikiforov
Academician Pilyugin Scientific and Production Center of Automatics and Instrument-Making,
Russia, Moscow
v.m.nikiforov@gmail.com

Abstract — The paper considers the circuit solution for designing a small-size three-axis simulation and calibration test bench, which features a fiber optic gyroscope as an inertial sensing element along the vertical rotational axis, which performs controlled turns with a given angular rate setting, as well as two pairs of linear accelerometers included in the control loops of the other two axes, and a processor in the control loop of the test bench.

Key words — precision rotary simulation and calibration test bench, processor, fiber optic gyroscope, linear accelerometer, automatic control system.

I. INTRODUCTION

In modern navigation, the technologies of strapdown inertial navigation systems (SINSs) have become widespread [1, 2]. No less urgent is the development of test equipment for navigation instruments that are used as inertial sensing elements (ISEs) in SINS arrangements [3]. Angular rate sensors (ARSs) of various accuracy class and principle of operation, which are part of SINSs, have been consistently increasing their accuracy, as well as achieving reduced size and mass characteristics and lowered power draw [2]. In addition, in most cases digital control systems (DCSs) [4-8] have started to become applied in ISEs included in modern SINSs. In this connection, it is an urgent task to create not only digitally controlled single-axis calibration test equipment with ISEs for verification of ARS characteristics by testing the scale factor error, but also three-axis simulation and calibration test benches built on the same principle for testing both individual ARSs and SINSs based on them [3, 9–16].

A three-axis small-size test bench with a fiber optic gyroscope (FOG) taken as its base ISE along the vertical rotational axis (pitch axis) will make it possible to not only calibrate ARSs of any accuracy class and principle of operation, but also to investigate the ARS itself and the associated SINS under the action of programmed motions set along all three axes of the test bench [3, 15, 16]. This is due to the fact that modern FOGs have reached the navigation class of accuracy in terms of scale factor and zero signal errors, which allows one to select such instruments as a basis for building precision automatic control systems, including test benches with ISEs. The main disadvantage of a FOG taken as a SINS sensing element is its rather large dimensions determined by the optical fiber coil diameter, which drastically limits the possibilities for designing, for instance, a SINS with non-orthogonal orientation of the measuring axes. Another drawback is the FOG's dependence on thermal perturbations. However, these disadvantages are not of decisive importance when designing a test bench based on a FOG ISE in the pitch axis. Three-axis simulation and calibration test benches built according to the above principle were considered by the authors in [3, 15, 16] and were discussed at international conferences on integrated navigation systems [15].

The difference of the newly proposed solution for designing a simulation and calibration test bench from the previously developed ones is that a pair of linear accelerometers, included in the automatic control system (ACS) loop of the bench, is installed along the both yaw (Y) and rotation axes (R). In the previously considered solutions of test bench design, the oscillatory motions of the axes were controlled by the data from induction angle sensors. In the proposed ACS, both FOGs and any linear accelerometers of medium and navigation accuracy class, depending on the task to be solved, can be utilized in a three-axis simulation and calibration test bench.

II. ADVANTAGES OF THE PROPOSED TECHNICAL SOLUTION OVER THE PREVIOUSLY DEVELOPED ONES

The proposed technical solution has a number of advantages over previously developed ones.

1. In the previously proposed designs of three-axis simulation and calibration test benches, the yaw and rotation axes were controlled using data from induction angle sensors, and the feedback was built as a classical tracking system. This principle was typical for both analog and digital control loops. The introduction of accelerometers intro the control loop makes

it possible to significantly improve the robustness of the test bench ACS in the mentioned axes.

2. Accelerometers installed along the yaw and rotation axes can be used in the test bench control system as tilt meters, allowing for accurate specification of the angles of rotation of the Y and R axes relative to the zero position, thus expanding the functional capabilities of the test bench.

3. Accelerometers in the control loop of the yaw and rotation axes of the three-axis test bench, combined with angle sensors for the same axes, improve the accuracy characteristics of the bench when specifying oscillations for the Y and R axes.

4. Generally, not only linear but also angular accelerometers can be used in the proposed solution, albeit such a solution requires additional research.

III. Functional and Kinematic Diagram of the Three-Axis Simulation and Calibration Test Bench with ISEs and Digital Control

The functional and kinematic diagram of the proposed test bench is shown in Fig. 1.

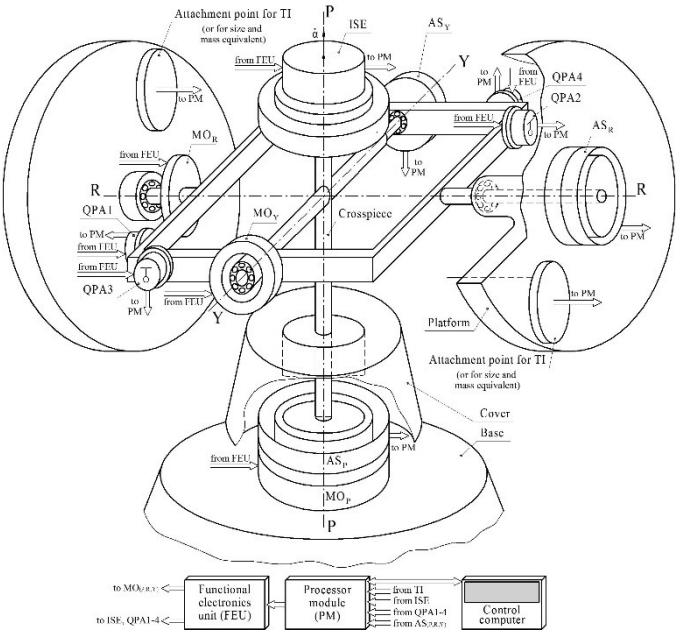


Fig. 1. Functional and kinematic diagram of the three-axis test bench with a digital control system.

In the diagram the following designations are introduced: TI – tested instrument; ISE – inertial sensing element; QPA – quartz pendulum accelerometer; PM – processor module; FEU – functional electronics unit; AS – angle sensor; MO – motor; P, R, Y – pitch, rotation and yaw axes, respectively; $\dot{\alpha}$- angular rate in the pitch axis.

Around the outer axis, the motion of the object along the pitch axis is simulated. The intermediate and inner axes of the gimbal are horizontal, and around these axes the oscillations of the object, which occur in the yaw and rotation axes, respectively, are simulated. Formation of laws of object oscillation, sinusoidal or random, for all three axes – pitch (P), yaw (Y) and rotation (R) is performed by the processor module. Angular oscillations excited in the P, Y and R axes are measured, respectively, by angle sensors (AS_P, AS_Y, AS_R), whose signals are transmitted both to the processor module and to the control computer. The inertial sensing element – FOG is included in the motor's digital control system in the P-axis. The Y and R axes are controlled by both pairs of quartz pendulum accelerometers (QPAs) 1-4, of which QPA 1 and QPA 2 are mounted on the Y-axis, and QPA 3 and QPA 4 – on the R-axis. To the control system, the signal is sent from both accelerometers and induction angle sensors AS_Y and AS_R, is combined and enters the automatic control system of the test bench through the processor module, which contains synthesized digital controllers; then through the functional electronics unit the signal enters the corresponding digital-to-analog converters, and from there it is sent to the drive motors of the corresponding gimbal axis. The adder implemented in the control processor for the P axis compares the codes proportional to the measured and driving voltages. The difference signal from the adder output is fed to the digital controller programmed in the processor, performing the P-axis control, and from there through the digital-to-analog converter goes to the drive motor of the P-axis. Therefore, in the proposed solution, the motor of the outer axis of the gimbal (the P-axis) receives separate control signals: the angular rate signal $F(\dot{\alpha})$, formed by the circuit, in which an ARS is taken as a sensing element, and the angle signal $F(\alpha)$, formed by the circuit, in which the angle sensor AS_P is taken as a sensing element. Analysis shows [15] that in such an arrangement the "slow" rotation around the P-axis can be specified more accurately.

The proposed small-size simulation and calibration test bench with a digital control system for the three gimbal axes can perform highly accurate rotation in the pitch axis (P), around which harmonic or random oscillations can be specified, as well as oscillations around the two axes Y and R, which can also be harmonic or random. The axes Y and R can rotate to precisely specified angles relative to the P-axis, which are formed by both angle sensors and accelerometers used in the initial rotation mode as tilt meters, and can oscillate according to a certain law in a given fixed angular position. Movements along the three axes are independent. The test bench allows one to calibrate angular rate sensors of different accuracy class and principle of operation and conduct dynamic testing of gyroscopic instruments and small-sized SINSs built on their basis. QPAs used in the proposed test bench design are taken as an example, they can feature analog or digital feedback, but in the case of accelerometers with digital feedback, the circuit solution of the bench control system will be different from the one proposed in this paper. The same applies if angular accelerometers are used.

IV. Structural Diagram of the Automatic Control System for the Yaw and Rotation Axes

The structural diagram of the Y-axis ACS has the form shown in Fig. 2.

Fig. 2. shows the following transfer functions:

- of the moving part of the test bench having a degree of freedom in relation to the Y-axis:

$$W_{\alpha Y}(s) = \frac{1}{I_{\alpha Y} \cdot s^2 + n_{\alpha Y} \cdot s}, \quad (1)$$

where $I_{\alpha Y}$ is the moment of inertia, $I_{\alpha Y} = 245$ *gcmc*^{2}; $n_{\alpha Y}$ is the damping factor, $n_{\alpha Y} = 40$ *gcmc*;

- of the correction circuit in the Y-axis:

$$W_{ccY}(s) = \frac{k_{ccY} \cdot s}{(T_{ccY} \cdot s + 1)^2}, \quad (2)$$

where k_{ccY} is the transfer coefficient, $k_{ccY} = 50$; T_{ccY} is the time constant, $T_{ccY} = 10^{-5}$ *s*;

- of the motor in the stabilization channel in the Y-axis:

$$W_{stY}(s) = \frac{k_{stY} \cdot (T_{st1Y} \cdot s + 1)}{s \cdot (T_{st2Y} \cdot s + 1) \cdot (T_{st3Y} \cdot s + 1)}, \quad (3)$$

where k_{stY} is the transfer coefficient, $k_{stY} = 1000$ *gcm*/(*V*·*s*); T_{st1Y}, T_{st2Y}, T_{st3Y} are the time constants, $T_{st1Y} = 0{,}5$ *s*, $T_{st2Y} = 5 \cdot 10^{-3}$ *s*, $T_{st3Y} = 5 \cdot 10^{-4}$ *s*.

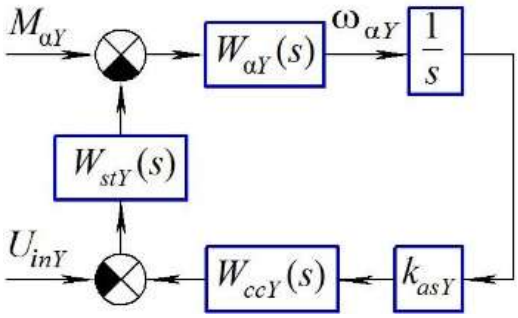


Fig. 2. . Structural diagram of the mathematical model of the Y-axis with an angle sensor and analog control system.

Fig. 2 has the following designations: U_{inY} – input voltage, *V*; $M_{\alpha Y}$ – sum of disturbing moments along the Y-axis, *gcm*; $\omega_{\alpha Y}$ – angular rate, *rad*/*s*; k_{asY} – angle sensor conversion ratio, $k_{asY} = 2{,}1$ *V*/*rad*.

Let us consider the digital stabilization circuit of the test bench for the Y-axis with an angle sensor and accelerometers as ISEs, the structural diagram of which is shown in Fig. 3.

Fig. 3, for simplicity, does not show external disturbing influences, except for the input U_{inY}, and does not take into account the dynamics of the motor and the accelerometer circuit. This simplification is quite justified [7, 9–11] since the time constants of these functional elements are several orders of magnitude less than the time constant of the moving part of the test bench – the Y-axis frame with elements mounted on it.

In addition to those introduced earlier, Fig. 3 has the following designations: k_{moY} – transfer coefficient of the motor in the stabilization circuit; $W_{conY}(z)$ – discrete transfer function of the controller; $k_{\omega Y}$, $k_{\varepsilon Y}$ – transfer coefficients of the control channels relying on the signals of *n* accelerometers (i.e. *n*=*2* for the studied case) and the angle sensor; k_{accY} – accelerometer transfer coefficient.

The method of calculation of the digital controller for the stabilization circuit is similar to that given in [10]. Fig. 4 shows the results of mathematical simulation of the bench operation, namely, the transient plots of the Y-axis angular rate in the three-axis test bench with analog and digital control systems and with control channels based on signals of an angle sensor and a pair of accelerometers.

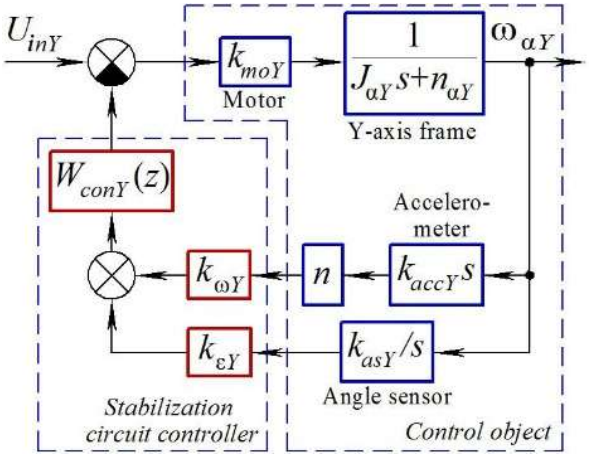


Fig. 3. Structural diagram of the mathematical model of the Y-axis with an angle sensor, accelerometers, and a digital control system.

Analysis of the plots shows that both automatic control systems of the test bench are stable, only the value of overshoot changes qualitatively (25% for the analog control system and 3% for the digital one), which does not deteriorate the quality of control.

Calculation of the control loop for the R-axis is not given here, since it is similar to the Y-axis ACS.

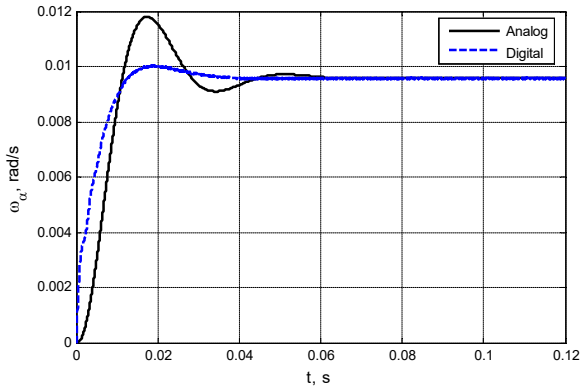


Fig. 4. Transient plots of the Y-axis angular rate of the three-axis test bench with analog and digital control systems.

V. CONCLUSIONS

Synthesis of digital controllers allows one to introduce flexibility into the structure of the bench itself, providing fast adaptation to any type of inertial sensing element, makes it possible to specify not only constant and harmonically changing input influences, but also influences that change according to

any law formed by the control computer. An inertial sensing element can be swapped without the need of changing the overall bench design, and the regulator of the motor control system is programmed for each inertial sensing element and the selection is carried out algorithmically. The use of accelerometers for the Y and R axes improves the accuracy characteristics of the ACS for these axes, improves the quality of control and, accordingly, its robustness. Application of such a test bench significantly reduces the cost of test equipment. The proposed test bench can be manufactured at enterprises specializing in the manufacture of inertial sensing elements.

Test bench characteristics: mass – not more than 50 kg; instability of the specified angular rate in the P-axis – not more than 10^{-4} °/*s*; amplitude of the specified harmonic or random oscillations on each of the three axes – not more than 10° with accuracy of positioning relative to the angular position of several angular minutes.

If precision angular rate sensors are used as inertial sensing elements for the P-axis control system, the instability of the specified angular rate can be improved up to 10^{-6} °/*s*. The accuracy characteristics of the bench equipment developed according to the proposed direction are improved in parallel with the betterment of the accuracy characteristics of SINSs and their ISEs, while the advancement of computer processing technologies ensures improvement of digital controllers, possibility of application of various algorithms of filtering and processing of signals supplied to the bench ACS from the ISEs, thus improving the ACS robust properties.

REFERENCES

[1] Izmailov, E.A., Modern Trends in the Development of Technologies of Inertial Sensing Elements and Aircraft Systems, Trudy FGUP NPTsAP, 2010, no. 1, pp. 27–35.

[2] Rivkin, B.S., Analiticheskii obzor sostoyaniya issledovanii i razrabotok v oblasti navigatsii za rubezhom (Analytical Review of the State of Research and Development in the Field of Navigation in Foreign Countries), St. Petersburg, CSRI Elektropribor, 2017–2023, Iss. 1–8.

[3] Kalikhman, D.M., Peshekhonov, V.G., Ed., Pretsizionnye upravlyaemye stendy dlya dinamicheskikh ispytanii giroskopicheskikh priborov (Precsision Controlled Test Benches for Dynamic Testing of Gyroscopic Instruments), CSRI Elektropribor, 2008.

[4] Besekerskii, V.A., Popov, E.P., Teoriya sistem avtomaticheskogo regulirovaniya (Theory of Automatic Control Systems), Moscow, Nauka, 1975.

[5] Izerman, R., Tsifrovye sistemy upravleniya (Digital Control Systems), Moscow, Mir, 1984.

[6] Kuo, B., Teoriya i proektirovanie tsifrovykh sistem upravleniya (Theory and Design of Digital Control Systems), Moscow, Mashinostroenie, 1986.

[7] Sadomtsev, Yu.V., Konstruirovanie sistem upravleniya s obratnoi svyaz'yu po kriteriyam tochnosti i grubosti (Designing Control Systems with Digital Feedback Using the Accuracy Criterion), Saratov, Yuri Gagarin State Technical University, 2003.

[8] Fedorov, S.M., Litvinov, A.P., Avtomaticheskie sistemy s tsifrovymi upravlyayushchimi mashinami (Automatic Systems with Digital Control Machines), Moscow, Leningrad, Energiya, 1965.

[9] Deputatova, E.A., Kalikhman, D.M. et al., Digital Motion Stabilization of Precision Test Benches with Inertial Sensing Elements. Part 1. Application of a Float-Type Angular Rate Meter, Izvestiya RAN. Teoriya i sistemy upravleniya, 2011, no. 1, pp. 120–132.

[10] Deputatova, E.A., Kalikhman, D.M. et al., Digital Motion Stabilization of Precision Test Benches with Inertial Sensing Elements. Part 2. Application of a Float-Type Angular Rate Meter and Pendulum Accelerometers, Izvestiya RAN. Teoriya i sistemy upravleniya, 2011, no. 1, pp. 131–146.

[11] Deputatova, E.A., Kalikhman, D.M. et al., New Generation Precision Rotary Test Benches with Inertial Sensing Elements and Digital Control, Izvestiya RAN. Teoriya i sistemy upravleniya, 2014, no. 2, pp. 130-146.

[12] Kalikhman, D.M., Deputatova, E.A. et al., Development of a Design Concept for a Class of Precision Mechatronic Test Benches with Inertial Sensing Elements Combined with Precision Angle Sensors, 29th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2022, pp. 212–217.

[13] Kalikhman, D.M., Deputatova, E.A. et al., Design Concept for a Class of Precision Test Benches with Inertial Sensing Elements in the Feedback Loop, Giroskopiya i navigatsiya, 2022, vol. 30, no. 3, pp. 41–64.

[14] Kalikhman, D.M., Deputatova, E.A. et al., Development of Control Algorithms for a Precision Test Bench with Inertial Sensing Elements and Digital Control System and Analysis of Its Errors, Izvestiya TulGU. Tekhnicheskie nauki, 2023. no. 9., pp. 66–80.

[15] Kalikhman, D.M., Kalikhman, L. Ya. et al., Three-Axis Simulation and Calibration Multipurpose Test Bench with Digital Control System for Gyroscopic Instruments, 19th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2012, pp. 91–94.

[16] Kalikhman, D.M., Kalikhman, L. Ya. et al., On the Possibility of Creating a Three-Axis Test Bench with Fiber Optic Gyroscope as an Inertial Sensing Element in the Vertical Rotational Axis, Aviatsionnaya promyshennost', 2008, no. 1., pp. 11–17.

QD-Type Sun Sensor Calibration Methods for Small Spacecraft Orientation

N. A. Kiz

Geoscan LTD,
Peter the Great St. Petersburg Polytechnic University
St. Petersburg, Russia
ORCID 0009-0000-4705-2703

Abstract – This work examines research and development of calibration algorithms for the solar sensor, used in orientation systems for small spacecraft. An alternative measurement model is introduced. A sensor field-of-view determination method has been implemented. A new algorithm has been synthesized based on the proposed measurement model and the assumption of linearity in calibration data. A comparative analysis of the used algorithms has been carried out.

Keywords – calibration, small spacecraft, Sun sensor, orientation system, field-of-view, measurement model, Sun direction vector, orientation vector

I. INTRODUCTION

The quadrant-type Sun sensor (QD, Quadrant Detector) is one of the primary components in the sensing part of orientation and stabilization systems for small spacecraft (CubeSat) that are used for a variety of scientific, research and educational purposes. For stable and efficient operation of this system, it is necessary to develop a calibration algorithm for the sensor, enabling the satellite orientation module to determine the Sun direction vector (the orientation vector) parameters.

This vector is defined in the coordinate system of the solar panel, onto which the sensor is mounted. The geometric scheme of this sensor is shown in Figure 1.

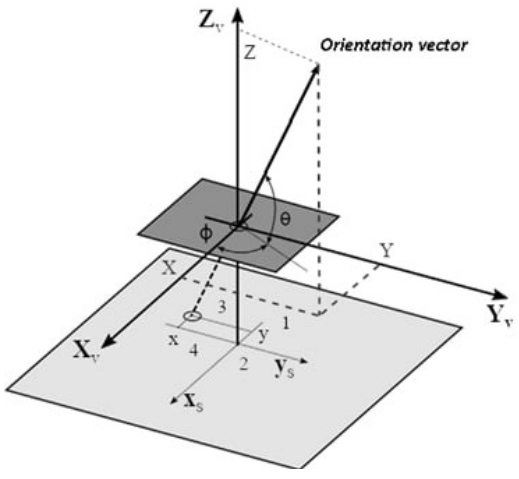


Fig. 1. Geometric diagram of the sensor

Two coordinate systems are considered: X_v, Y_v, Z_v – axes of the solar panel coordinate system, X, Y, Z – orientation vector coordinates in that system; x_s, y_s – axes of the sensor coordinate system and x, y – the light spot center coordinates in that coordinate system. Each of the sensor's photodiodes actually represents a quadrant inside of its coordinate system and is denoted 1-4 in Figure 1.

In the small spacecraft industry, the calibration problem so far has been studied for other types of sensors essentially [1,2]. There also researches related to QD-sensors [3]. A disadvantage of existing solutions is the measurement model, in which the parameters of the orientation vector are represented as two spherical angles. In this work, an alternative measurement model is proposed, which enabled the development of a new calibration algorithm. Previously known methods have been adapted to the new measurement model.

II. PROBLEM STATEMENT

The input data are the digital values of the light intensity falling on each of the sensor's photodiodes. Based on these measurements, it is required to determine the Sun direction vector with an average accuracy of at least 1° within sensor's field of view. The calibration process consists of finding a transformation that maps the coordinates of the light spot center of the sensor into the coordinates of the orientation vector, based on the dataset, collected during laboratory sessions using two-axis rotary platform. The coordinates of the light spot center in Figure 1 (x, y) are determined using the following formulas (1):

$$x = \frac{(I_2 + I_4) - (I_1 + I_3)}{\sum_{i=1}^{4} I_i}$$
$$y = \frac{(I_1 + I_2) - (I_3 + I_4)}{\sum_{i=1}^{4} I_i}, \quad (1)$$

where I_i are digital values of the sunlight intensity on the corresponding photodiodes.

In the selected measurement model, the Sun direction vector $\begin{bmatrix} X & Y & Z \end{bmatrix}^T$ is assumed to be a unit vector, so the problem is reduced to determining two of it's coordinates, and can be expressed as (2):

$$\begin{bmatrix} x \\ y \end{bmatrix} \xrightarrow{A} \begin{bmatrix} X \\ Y \end{bmatrix}, \quad (2)$$

where A is some mapping, the form of which depends on the chosen calibration method.

III. SENSOR FIELD-OF-VIEW DETERMINATION

The field-of-view of the sensor is defined as the maximum zenith angle deviation from the perpendicular to the sensor surface at which the error between the real

orientation vector and the vector obtained by calibration algorithm does not exceed the required threshold. Experiments have shown that the sensor shows better measurement accuracy at smaller zenith angles. The error increases as the zenith angles grow. The error level – an absolute angular value – is determined by the following formula (3):

$$error = \cos^{-1}(\mathbf{V}_{real} \cdot \mathbf{V}_{pred}), \quad (3)$$

where $\mathbf{V}_{real}$ is real orientation vector and $\mathbf{V}_{pred}$ is the predicted vector, obtained by the calibration algorithm.

The sensor's field-of-view is the region where the transformation A remains unambiguous. It is limited by the zenith angle of the sensor relative to the light source. The rotary platform used in the experiments enables capturing the dependence of the light spot center coordinates on the orientation vector coordinates, meaning that the sough mapping A is the inverse of the transformation that can be formed based on the collected data. For preliminary analysis of the field of view, an auxiliary function (4) is introduced:

$$f_{FOV} = r(R), \quad (4)$$

where r is the magnitude of the radius vector, formed from the coordinates of the light spot center (x, y) and R – the magnitude of the vector, formed from the components of the orientation vector (X, Y).

This function simplifies field-of-view estimation by reducing the problem to a study of a single variable function. It can be assumed, that to define the field-of-view boundary, it is sufficient to find the point of the local extremum of function (4). Due to the sensor's anisotropic properties, a series of measurements at various azimuth angles is required. The goal is to find the minimum value of R, at which one of the functions (4) reaches it's maximum value. The corresponding zenith angle can then be calculated.

Preliminary analysis results are shown in figure 2.

In order to define sensor's field-of-view more rigorously, it is necessary to analyze the gradients of the functions that relate each coordinate of the sensor's light spot center (x, y) to the orientation vector coordinates (X, Y). Figures 3 and 4 partially show the results of such analysis for two different sensors: one without any defects, and one that got a defect during solar panel manufacturing process.

Two functions are considered separately: $x = f_1(X, Y)$ and $y = f_2(X, Y)$. For each laboratory measurement, the gradient of these functions at the point is determined. Next, the average gradient direction is found (represented by the red vector on the figures, scaled for better visibility).

Then, the angle between the gradient at each point and the average gradient vector is computed. An increase in this angle indicates proximity to a local extremum of the function, which can be interpreted as the field-of-view boundary in this context.

Concentric circles on the angle distribution plots represent zenith angle levels with 10° increments. These levels define actual field of view of the sensor. For the defective sensor, a field of view is narrowed to a cone with an angle of less than 50°. Meanwhile, the field-of-view of the defect-free sensor is significantly larger. Data from defective sensors was not used in this research

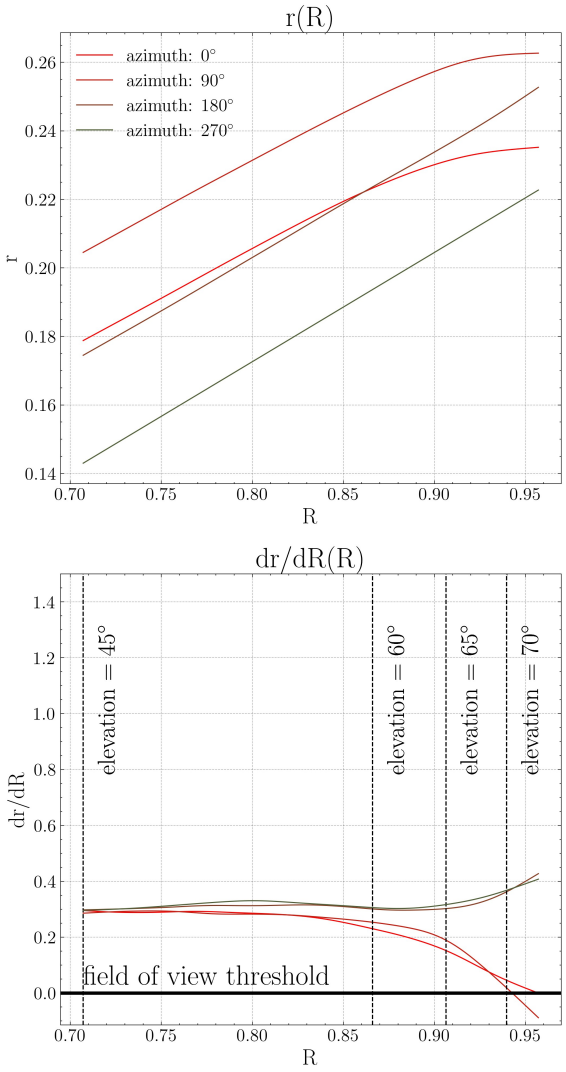


Fig. 2. Preliminary sensor field-of-view analysis

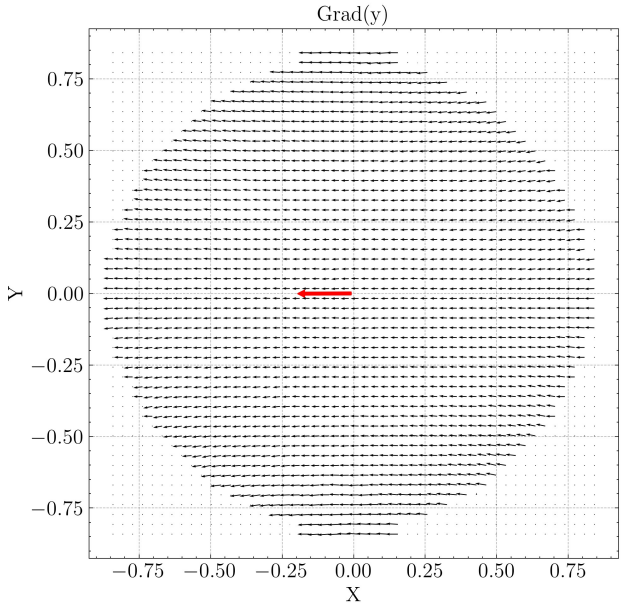


Fig. 3a. Gradients of function f_2 for the defect-free sensor

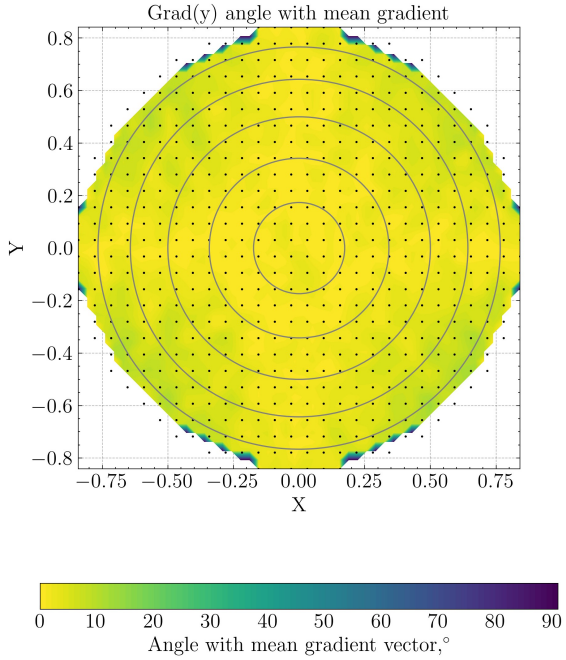


Fig. 3b. Gradient angles heatmap of function f_2 for the defect-free sensor

IV. Research on Calibration Algorithms

A. Linear transformation algorithm

Due to the sufficient linearity of the calibration data, it is assumed that the transformation A can be fairly accurately approximated in the following form (5):

$$\begin{bmatrix} X \\ Y \end{bmatrix} = A \times \begin{bmatrix} x + x_c \\ y + y_c \end{bmatrix}, A \in \mathbb{R}^{2\times 2}, \quad (5)$$

where x_c, y_c are the linear shifts of the sensor coordinate system relative to the solar panel's coordinate system, which arise due to mounting inaccuracies.

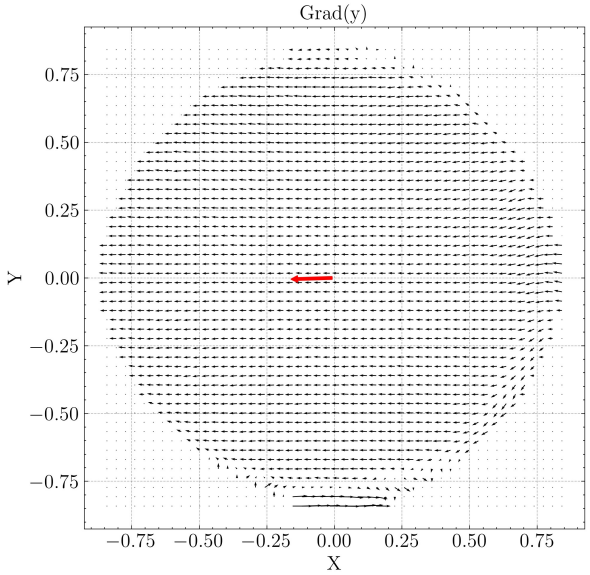


Fig. 4a. Gradients of function f_2 for defective sensor

To determine the coefficients of the matrix A, a system of equations (6) can be constructed based on the dataset obtained from the rotary platform:

$$\begin{cases} X_1 = a_{11}\cdot x_1 + a_{12}\cdot y_1 + a_{21}\cdot 0 + a_{22}\cdot 0 + x_c\cdot 1 + y_c\cdot 0 \\ Y_1 = a_{11}\cdot 0 + a_{12}\cdot 0 + a_{21}\cdot x_1 + a_{22}\cdot y_1 + x_c\cdot 0 + y_c\cdot 1 \\ \vdots \\ X_n = a_{11}\cdot x_n + a_{12}\cdot y_n + a_{21}\cdot 0 + a_{22}\cdot 0 + x_c\cdot 1 + y_c\cdot 0 \\ Y_n = a_{11}\cdot 0 + a_{12}\cdot 0 + a_{21}\cdot x_n + a_{22}\cdot y_n + x_c\cdot 0 + y_c\cdot 1 \end{cases}, \quad (6)$$

where X_i, Y_i are the known orientation vector components, x_i, y_i are the light spot center coordinates of the sensor and a_{jk} are the coefficients of the matrix A.

This system is overdetermined and is solved numerically using the least squares method. The result of calibration for each sensor is a vector of six values:

$$\left[a_{11}, a_{12}, a_{21}, a_{22}, x_c, y_c\right]^T$$

The calibration results are validated using a test dataset that was not involved in the calibration process. This algorithm provides an accuracy within the range of 1.2–2°. An example of the error distribution within sensor field-of-view is shown in figure 5

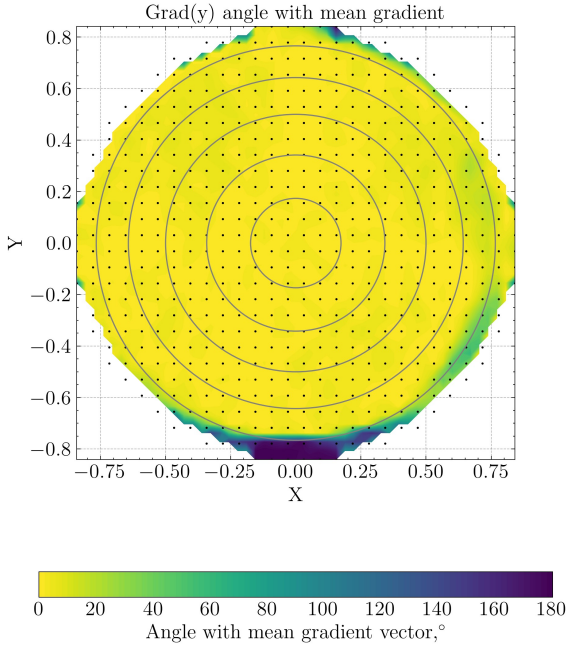


Fig. 4b. Gradient angles heatmap of function f_2 for defective sensor

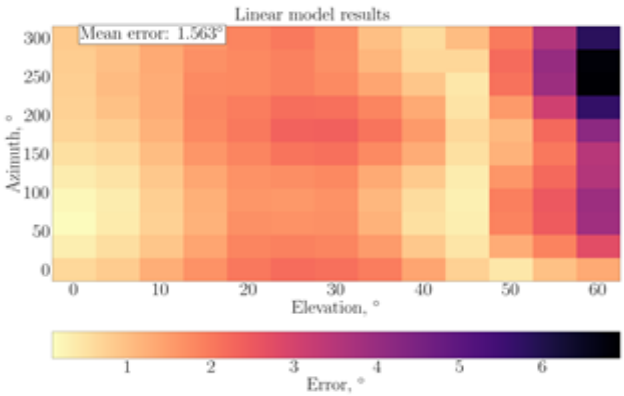


Fig. 5. Error distribution within sensor field-of-view for linear model

B. *Algorithm based in geometric transformations*

A calibration method based on geometric transformations of light spot center coordinates is widely used to solve this kind of problem [1, 2]. This transformation links two coordinate systems – the sensor's system and the spacecraft solar panel system. By adapting this method to the selected measurement model, the transformation can be represented as (7):

$$X = \sin^{-1}\left(\tan^{-1}\left(\left(\frac{x_c^* - x}{d_0}\right)\cdot\cos\beta_0\right) + \alpha_0\right)$$

$$Y = \sin^{-1}\left(\tan^{-1}\left(\left(\frac{y_c^* - y}{d_0}\right)\cdot\cos\alpha_0\right) + \beta_0\right)$$

$$\begin{bmatrix} x_c^* \\ y_c^* \end{bmatrix} = \begin{bmatrix} \cos\varphi_0 & -\sin\varphi_0 \\ \sin\varphi_0 & \cos\varphi_0 \end{bmatrix} \times \begin{bmatrix} x_c \\ y_c \end{bmatrix}, \quad (7)$$

where d_0 is the distance between the mask of the sensor mounted in the solar panel and the photodiode elements and $\alpha_0, \beta_0, \varphi_0$ are the Euler angles describing the angular shifts between the coordinate systems.

The angular displacements between these coordinate systems are negligibly small, which is confirmed by the minimal difference in model accuracy for the model that considers these angles, and the model where these angles were set to zero:

- With $\alpha_0, \beta_0, \varphi_0$: error range 0.9–1.6°
- Without $\alpha_0, \beta_0, \varphi_0$: error range 0.9–1.8°

An example of error distribution within sensor field-of-view based on testing dataset is shown in figure 6.

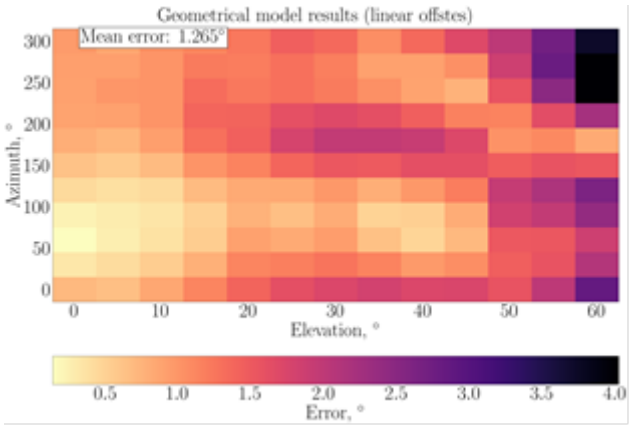


Fig. 6. Error distribution within sensor field-of-view for geometrical model

C. *Polynomial regression algorithm*

One of the most effective methods for solving the calibration problem – widely used in scientific literature [1–4] – is an algorithm based on multivariable polynomial regression with a MSE (Mean Squared Error) loss function. In this case, the transformation is represented as (8):

$$X = \sum_{i=0}^{n}\sum_{j=0}^{n-i} c_{ij}^X x^i y^j \quad (8)$$

$$Y = \sum_{i=0}^{n}\sum_{j=0}^{n-i} c_{ij}^Y x^i y^j,$$

where c_{ij}^X, c_{ij}^Y are the polynomial coefficients and n is the degree of the polynomial.

When using a third-degree polynomial, the results satisfy the calibration accuracy requirements. Increasing the polynomial degree beyond $n = 3$ poorly improves model accuracy:

- For n = 3, the error is in the range 0.21 – 0.59°
- For n = 4, the error is 0.18 – 0.43°

However, increasing the polynomial degree increases the total number of coefficients required for Sun direction vector estimation exponentially. This affects the memory efficiency of the algorithm since it demands storing more coefficients on the satellite's compute module memory. The validation results for this model on one of the sensors are shown in figure 7.

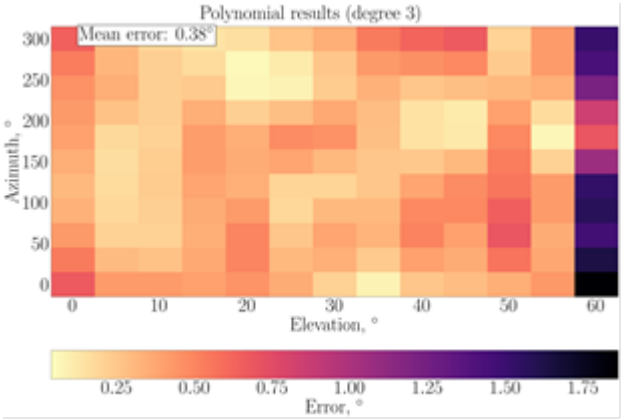


Fig. 7. Error distribution within sensor field-of-view for polynomial regression algorithm

V. Conclusion

In this paper, a method of determining and analyzing the field-of-view of a QD-type solar sensor has been implemented. The algorithms for its calibration have been examined and tested. The new algorithm, based on the proposed measurement model and calibration data linearity was synthesized. The best calibration accuracy was achieved using the algorithm based on multivariable polynomial regression. Its accuracy on the calibration dataset was:

- 0.18 – 0.43° for a fourth-degree polynomial
- 0.21 – 0.59° for a cubic polynomial

All research was conducted using data from 30 sensors that were calibrated during the development of the orientation system for spacecraft at Geoscan LTD. The calibration results have been integrated into the satellite orientation system and are currently being used on real satellites that were launched into orbit on November 5, 2024.

References

[1] Degtyarev A., Grassi M., Perrotta A., Ovchinnikov M. Testing and calibration of a miniature sun sensor prototype developed using APS technology. Preprints of Keldysh Institute of Applied Mathematics, 2005. 30 pages.

[2] Rufino G., Perrotta A., Grassi M. Laboratory test of an APS-based sensor prototype. Proceedings of the 5th International Conference on Space Optics (ICSO). 2004. Pages 551-559.

[3] Grigorov P. Laboratory research of the characteristics of a sun sensor based on semiconductor crystal for microsatellite orientation determination systems. Moscow Institute of Physics and Technology (MIPT), 2014. 28 Pages.

[4] Sun Q., Gutiérrez J. L. R., Yu X. Deep Neural Network-Based 4-Quadrant Analog Sun Sensor Calibration. Space: Science & Technology. 2023. 13 pages.

Development of Methodology and Study of Operational and Functional Characteristics of Microelectromechanical Sensors

Leonid Chelyshev

Institute of Nano- and Microsystems Engineering (NMST)
National Research University of Electronic Technology (MIET)
Moscow, Russia
leonid.chelyshev@live.ru

Sergei Timoshenkov

Institute of Nano- and Microsystems Engineering (NMST)
National Research University of Electronic Technology (MIET)
Moscow, Russia
spt111@mail.ru

***Abstract*— The report describes a method for analyzing the characteristics of MEMS accelerometers, including zero-shift data, scale factor, and temperature dependencies.**

Keywords*— *MEMS, measurement, accelerometers, acceleration output characteristic, accuracy

I. Introduction

A critical factor affecting the operational accuracy of MEMS is the influence of temperature and mechanical stresses. Temperature variations can induce significant changes in the physico-mechanical properties of MEMS materials and the performance of embedded electronics, leading to parameter deviations and reduced operational reliability. Consequently, investigating thermodynamic processes in MEMS is essential for the correct interpretation of their functional parameters.

In the Russian Federation, research on MEMS device development is actively conducted; however, progress is constrained by several technological limitations. Notably, the lack of advanced methodologies for monitoring thermal parameters in MEMS sensors complicates rapid and accurate quality assessment of manufactured devices. This may result in undetected errors during their operational use.

II. Development of an Automated Checking System

To study the functional and operational characteristics of MEMS accelerometers, it is necessary to create a system that meets several parameters:

1. The system must include the test article that complies with modern trends and possesses the corresponding characteristics.
2. The test article must be subjected to specific influences to alter its output signal. It should be possible to determine at any stage of the experiment which influence is being applied to the sensor. For convenience, a standard testing methodology should be developed, along with the capability to conduct experiments in a "manual" mode where the operator selects the influences applied to the sensor at any given moment.
3. The acquired data must be collected by the system without loss. Given the variety of data formats transmitted by modern sensors, the system should be capable of receiving any or most of the commonly used formats.
4. The operator or developer should receive the data on widely used equipment (PC or any other computer) via a compatible interface. Additionally, to enhance usability during the experiment, real-time monitoring should be available.
5. The data received by the computer must undergo processing and be prepared for further analysis in a format convenient for the operator (developer).

A CNC two-axis rotary platform is required to detect the acceleration due to gravity (g) and, after analysis, obtain the output signal characteristics of the accelerometer or gyroscope. The dimensions and specifications of the rotary platform allow it to be used inside a thermal chamber, enabling automated data acquisition under varying temperature conditions. A CNC tilt platform can apply influences exceeding ±g and track changes in the scale factor. A control module is necessary to record the parameters of the tested accelerometers and implement the test algorithm using data interpreted by the rotary and tilt platforms. Additionally, it must be possible to save and transfer data to a computer for further analysis of the tested sensors' behavior. Figure 1 shows the test workstation setup.

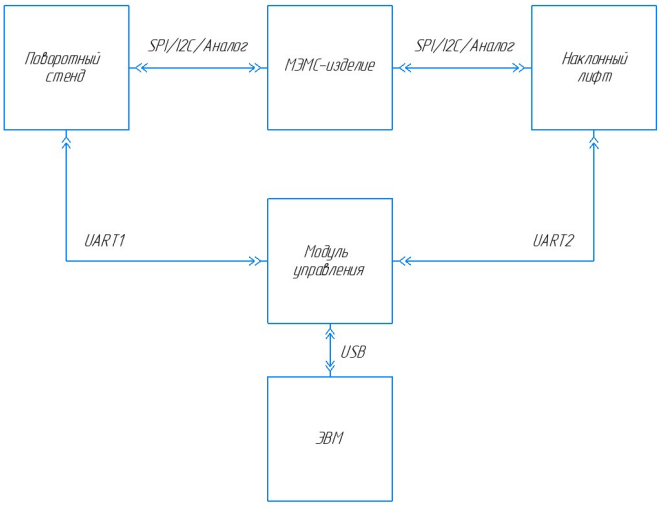


Fig. 1. Structural diagram of the developed solution.

III. Conducting Experiments

The following accelerometer was selected for the study: the analog ADXL335 model, for which a testing algorithm was developed and characteristics were calculated.

First, it is necessary to convert the voltage signal from the sensor into acceleration values. For this conversion, the following formula is used:

$$A = \frac{A_v - V_{0g}}{S_{en}} * 1000 \qquad (1)$$

where A is acceleration in mg, Av is acceleration in volts, Vog is zero-g voltage, and Sen is axis sensitivity. According to ADXL335 documentation, Vog = 1.5V and Sen = 300 mV/g.

The mean zero offset for the accelerometers was:

AoffX = 188.8475, AoffX = 217.9334722, AoffX = 213.3153333 for Sample 1,

AoffX = 263.9681667, AoffX = 210.9816389, AoffX = 298.083037 for Sample 2.

The mean maximum values were:

AmaxX = 1081.964167, AmaxX = 1102.893861, AmaxX = 1135.834 for Sample 1,

AmaxX = 1072.196, AmaxX = 1131.927694, AmaxX = 1135.908296 for Sample 2.

To evaluate sensor performance under vibration effects, we calculated the standard deviation from the mean in the data sample. The RMS values for X, Y, Z axes of ADXL335 (two samples at normal temperature +20°C in the thermal chamber) were:

s1x= 23.524 mg, s1y= 13.6151 mg, s1z= 13.2915 mg for Sample 1,

s2x= 22.2049 mg, s1y= 14.6568 mg, s2z= 11.5348 mg for Sample 2.

The obtained data indicate relatively low noise levels characteristic of stationary state, stable output signal with minimal variations. The Y-axis shows higher RMS compared to other axes, which is explained by measurements taken during acceleration peaks (13,400-15,400 samples out of 24,000) rather than stationary state. For other axes, larger ranges were selected: 12,000 samples for X-axis and 9,000 for Z-axis.

We calculated RMS for the difference between actual and simulated data for A and Aresult, obtaining:

SA1X = 169.1904233 mg, SA1Y = 249.4242765 mg, SA1Z = 280.2979866 mg.

SA2X = 251.8298834 mg, SA2Y = 254.901225 mg, SA2Z = 361.6492017 mg.

SAresult1X = 26.30828309 mg, SAresult1Y = 12.42640397 mg, SAresult1Z = 14.82111543 mg.

SAresult2X = 34.04352478 mg, SAresult2Y = 14.03348902 mg, SAresult2Z = 15.85671757 mg.

This allows calculating the difference between processed and unprocessed data relative to expected results:

Sample 1, X-axis: 6.43

Sample 1, Y-axis: 20.07

Sample 1, Z-axis: 18.91

Sample 2, X-axis: 7.4

Sample 2, Y-axis: 18.16

Sample 2, Z-axis: 22.81.

Thus, we conclude that the implemented procedures improve final result accuracy severalfold, making accelerometer output data reproducible with error margins approaching normal sensor noise levels (15-30 mg).

One frequently overlooked parameter during calibration is output signal variation under different thermal conditions. Therefore, we conducted MEMS accelerometer studies considering various temperature effects, using -40...+60°C range as optimal for typical accelerometer operation.

Temperature variation tests were performed using Espec MC-811P thermal chamber (shown in Figures 14), with 10°C increments. After reaching target temperature, samples were thermally stabilized either until temperature equilibrium or for 30 minutes. Resulting temperature dependence graphs are presented in Figures 2-4, with polynomial approximations developed for intermediate value determination.

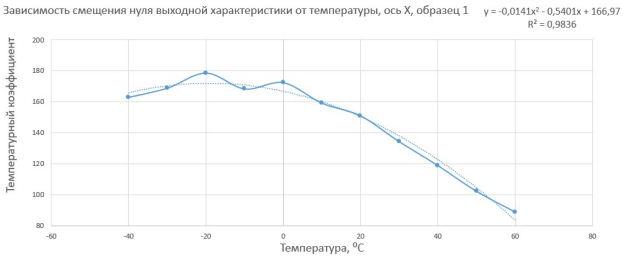


Fig. 2. Output characteristic zero offset versus temperature, X-axis, sample 1

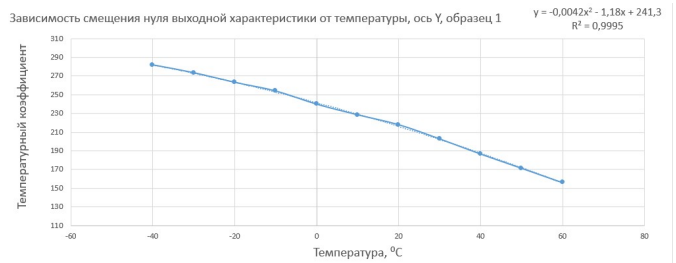


Fig. 3. Output characteristic zero offset versus temperature, Y axis, sample 1

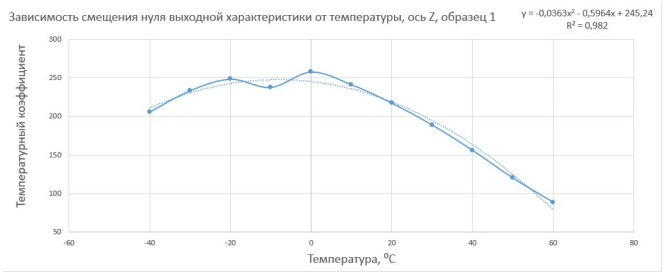


Fig. 4. Output characteristic zero offset versus temperature, Z axis, sample 1

IV. Conclusions

A module was developed to implement the proposed algorithm and interface with the final products, enabling comprehensive performance studies that could significantly improve the quality and reliability of these devices. Experimental testing was conducted on the developed automated system for analyzing primary error sources in MEMS devices.

The obtained data proves sufficient for multiple-fold accuracy improvement in analog accelerometers, including the studied ADXL335 models. The research revealed consistent trends of sensitivity reduction with temperature variations. While demonstrating high inter-axis correlation, each axis exhibited unique behavioral characteristics. The automated testing method yielded excellent results both in terms of testing speed (the described 24,000 measurements were completed in approximately 1 minute) and achieved multiple-fold accuracy enhancement of the tested accelerometers.

The developed system enables high-speed automated acquisition of key parameters in MEMS accelerometers. Compared to similar solutions, the system features relatively low cost while providing cost-effective calibration with high accuracy-to-cost ratio. It can serve as a training tool for microelectronics developers and as an essential instrument for identifying substandard products.

REFERENCES

[1] Kochurina E.S., Anchurin S.A., Timoshenkov A.S., Boev L.R., Putilov Ya.I., Nalsky A.A., Galkin A.A., Lebedev A.A. Avtomatizatsiya ispytaniy mikromekhanicheskikh datchikov [Automation of Testing for Micromechanical Sensors]. Intellektual'nye sistemy i mikrosistemnaya tekhnika. Nauchno-prakticheskaya konferentsiya: sb. trudov [Intelligent Systems and Microsystem Technology. Scientific-Practical Conference: Proceedings]. Moscow: MIET, 2024, pp. 5-13. (in Russian)

[2] Filatov Yu.V., Boronakhin A.M., Dao V.B., Le V.Ch., Podgornaya L.N. Issledovanie statisticheskikh pogreshnostey triady mikromekhanicheskikh akselerometrov v rezhime kvazigarmonicheskikh kolebaniy [Study of Statistical Errors in Triads of Micromechanical Accelerometers in Quasi-Harmonic Oscillation Mode]. Giroskopiya i navigatsiya. Nauchno-tekhnicheskiy zhurnal [Gyroscopy and Navigation. Scientific-Technical Journal]. St. Petersburg: CNII "Electropribor", 2016, pp. 35-46. (in Russian)

[3] Nemec D., Janota A., Hruboš M., Šimák V. Intelligent Real-Time MEMS Sensor Fusion and Calibration. IEEE Sensors Journal, 2016, vol. 16, no. 19, pp. 7150-7160.

[4] Li Y., Georgy J., Niu X., Li Q., El-Sheimy N. Autonomous Calibration of MEMS Gyros in Consumer Portable Devices. IEEE Sensors Journal, 2015, vol. 15, no. 7, pp. 4062-4072.

[5] Shen S.-Ch., Chen Ch.-J., Huang H.-J. A New Calibration Method for Low Cost MEMS Inertial Sensor Module. Journal of Marine Science and Technology, 2010, vol. 18, no. 6, Article 5. DOI: 10.51400/2709-6998.1939

[6] Frosio I., Pedersini F., Borghese A. Autocalibration of MEMS Accelerometers. IEEE Transactions on Instrumentation and Measurement, 2009, vol. 58, no. 6, pp. [pages missing].

[7] Fontanella R., Accardo D., Lo Moriello R.S., Angrisani L., De Simone D. MEMS Gyros Temperature Calibration through Artificial Neural Networks. Sensors and Actuators A: Physical, 2018. https://doi.org/10.1016/j.sna.2018.04.008

Studying the Modulation Factor Stability in the Magnetometric System of the Quantum Rotation Sensor

Viktor Chalkov, Alexader Shevchenko
Concern CSRI Elektropribor, JSC,
St. Petersburg, Russia
rztbro@gmail.com

Abstract— The operation of a magnetometric system of a quantum rotation sensor is studied. The influence of the resonance quality factor fluctuation in the cell on the optimal modulation factor of the magnetometric system under different operating modes is analyzed. Various possible methods for stabilizing the modulation factor are considered.

Keywords— quantum rotation sensor

I. INTRODUCTION

Quantum rotation sensors (QRS) based on the effect of nuclear magnetic resonance in noble gases and paramagnetic resonance in alkali metals show high accuracy, have no moving parts and are resistant to external effects [1–4]. One of the key components of such systems is a magnetometric system intended to register the Larmor precession of atoms of the working substance, sensitive to external magnetic fields and the angular rate of rotation of the base.

The performance of a magnetometric system largely depends on its initialization mode, including the value of the modulation factor η which determines the ratio between the amplitude of the alternating and magnetostatic magnetic field along the sensitivity axis. In practice, the choice of η depends on the resonance quality factor ξ, which can vary from run to run.

This work is aimed to study the stability of the magnetometric system output signals in three start-up modes, different in the approaches to the choice of modulation parameters.

II. BASIC PRINCIPLES OF OPERATION OF A MAGNETOMETRIC SYSTEM

The QRS magnetometric system is an optically pumped atomic magnetometer placed in the same cell as the noble gas isotopes (e.g. 131Xe). The system is designed to register the magnetic moment precession of an alkali metal (usually cesium or rubidium) in a resulting magnetic field that includes a constant component B_0 and an alternating modulation field $B_{rf}(t)$, as well as a field caused by the precession of the macroscopic magnetic moment of an ensemble of noble gas atoms. This precession is registered by the change in polarization of the detecting light caused by the Faraday effect.

A. Influence on the magnetometric system sensitivity

The sensitivity of the magnetometric system is significantly affected by the modulation factor

$$\eta=B_0/B_{rf}, \qquad (1)$$

At the same time, characteristics of the signal at the system output depend not only on η, but also on the quality factor of the resonance in the cell:

$$\xi=\omega_0\cdot\tau. \qquad (2)$$

where ω_0 is the Larmor frequency of the spins of alkali metal atoms; τ is the coherence time of the atomic state.

B. Output signal of the magnetometric system

As already shown, the output signal of the magnetometric system depends on η, the resonance quality factor ξ in the cell, and the phase shift between the signal components along the X- and Y-axes, perpendicular to the sensor sensitivity axis. Taking into account the phase difference, the amplitudes A_x and A_y at the outputs of the magnetometric system are defined by the following formulas:

$$A_x(\eta) \approx |\hat{A}_x^1(\eta)\cdot\cos(\theta_x^1(\eta)-\theta_\gamma^1(\eta))|\,|, \qquad (3)$$

$$A_\gamma(\eta) \approx |\hat{A}_\gamma^1(\eta)\cdot\cos(\theta_x^1(\eta)-\theta_\gamma^1(\eta))|\,. \qquad (4)$$

The sought expressions for the theoretical first-harmonic amplitudes $\hat{A}_x^1$ and $\hat{A}_\gamma^1$, representing the intrinsic harmonic response of the magnetometric system, can be given in the form [5]:

$$\hat{A}_x^1 = \sqrt{[(W_0+W_2)^2+(Q_2-Q_0)^2]}\,, \qquad (5)$$

$$\hat{A}_\gamma^1 = \sqrt{[(W_0-W_2)^2+(Q_2+Q_0)^2]}. \qquad (6)$$

The phase tangents θ_x^1, θ_γ^1 of these signals are described by the following dependencies [6]

$$\tan(\theta_x^1) = (W_2+W_0)/(Q_2-Q_0), \qquad (7)$$

$$\tan(\theta_\gamma^1) = (W_2-W_0)/(Q_2+Q_0). \qquad (8)$$

W_0, W_2, Q_0, and Q_2 in (5–8) represent infinite sums the values of which are determined by the following expressions [6, 7]:

$$W_2(\eta) = \Sigma_n[J_n(\eta)\cdot J_{n+1}(\eta)\cdot(n+1)\cdot\xi]/[1+(n+1)^2\cdot\xi^2],$$

$$Q_2(\eta) = \Sigma_n\,[J_n(\eta)\cdot J_{n+1}(\eta)]/[1+(n+1)^2\cdot\xi^2],$$

$$W_0(\eta) = \Sigma_n\,[J_n(\eta)\cdot J_{n-1}(\eta)\cdot(n+1)\cdot\xi]/[1+(n+1)^2\cdot\xi^2],$$

$$Q_0(\eta) = \Sigma_n\,[J_n(\eta)\cdot J_{n-1}(\eta)]/[1+(n+1)^2\cdot\xi^2],$$

where n is from $-\infty$ to ∞.

III. SETTING-UP OF THE MAGNETOMETRIC SYSTEM

Since ξ can change from run to run, the choice of modulation parameters must take into account the current state of the system. In this paper, we consider three modes of the magnetometric system start-up which are different in the approaches to the choice of η and resistance to external effects.

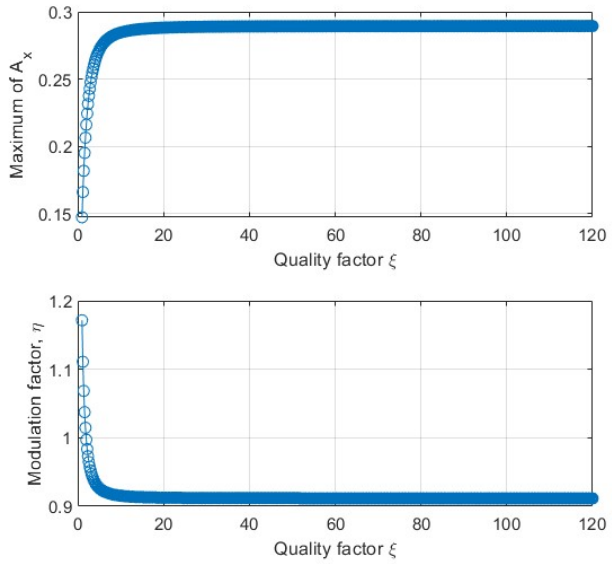


Fig. 1. A graph (top) of the dependence of the maximum value of the amplitude of signals Ax from the magnetometric system on the resonance quality factor in the cell and a graph (bottom) of the dependence of the value of the optimal tuned modulation factor on the resonance quality factor in the cell

A. *Mode 1: initialization at the signal amplitude maximum*

The most universal method of initializing a magnetometric system is to start it when the value of the modulation factor η corresponding to the maximum signal amplitude in one of the magnetometer's outputs, usually the X output (3), is reached, which ensures an optimal signal-to-noise ratio under the condition of stationary white noise. According to numerical calculations and experimental data [8], the optimal value of the modulation factor in this mode is $\eta \approx 0.91$.

B. *Mode 2: phase-insensitive stabilization (at $\xi \gg 1$)*

At a high value of the resonance quality factor ($\xi \gg 1$), it is possible to use the phase-insensitive stabilization mode [7]. This configuration minimizes the influence of crosstalk and phase drifts. In this mode, the modulation factor η is selected so that the output amplitude $A_x(\eta)$ (3) becomes zero. The optimal value of η for this case is approximately 1.82 (see Fig. 2).

C. *Mode 3: symmetrical start (at $A_x = A_y$)*

Mode 3 is based on the selection of the modulation factor η at which the output amplitudes $A_x(\eta)$ (3) and $A_y(\eta)$ (4) become equal. This mode is used in systems where monitoring is performed simultaneously on both outputs. According to the calculations, the value of the modulation factor at which the condition of $A_x = A_y$ is fulfilled in the system under consideration is η=2.4.

IV. Analysis of the Modulation Factor Stability

Estimation of the optimal modulation factor stability is an important step in choosing the method for starting the magnetometric system. External conditions (e.g. temperature, magnetic field gradients or variations of current in the coils) can significantly affect the cell parameters, which causes a change in the resonance quality factor; for example, an increase in the magnetic field strength in the QRS sensitive element by 1 μT may lead to a deviation of $\Delta\xi \approx 1.7493$, which can significantly affect the operation of the sensor. This requires estimation of the sensitivity of the selected start-up mode to such deviations.

The figures show the dependences of the signal amplitudes A_x and A_y on the resonance quality factor ξ, as well as the corresponding changes in the optimal value of the modulation factor η for each of the three modes. Based on the simulation, we have made the following observations.

A. *Analysis of the modulation factor stability in mode 1*

Significant variations of both the signal amplitude and the optimal value of η are observed in the range of $\xi = 1–50$. Only at $\xi > 50$, the parameters stabilize and become almost invariable. Thus, when using mode 1 with sensors at $\xi < 50$, it is necessary to carry out an additional check of the current value of ξ and corresponding correction of η at each start-up.

B. *Analysis of the modulation factor stability in mode 2*

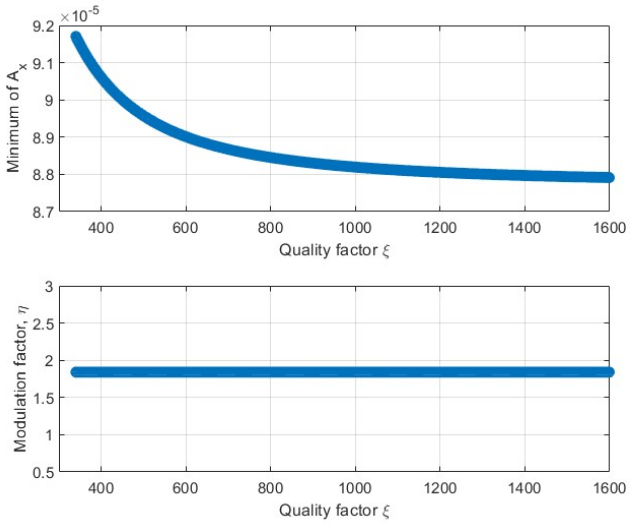


Fig. 2. A graph (top) of the dependence of the minimum value of the amplitude of signals Ax from the magnetometric system on the resonance quality factor in the cell and a graph (bottom) of the dependence of the value of the optimal tuned modulation factor on the resonance quality factor in the cell

Changes in the quality factor ξ do not practically affect the accuracy and stability of the magnetometric system in this mode. The optimal value of η remains practically invariable, and the signal amplitude remains within the permissible deviations. However, the realization of this mode requires a high resonance quality factor in the cell: stable operation is achieved at $\xi \gtrsim 100$, and its most complete implementation, at $\xi \sim 500$.

It should be noted that for the quality factor values less than ~500, the signal amplitude A_y tends to zero, as does A_x, which makes the phase-insensitive stabilization start-up potentially dangerous: the A_y can disappear completely, especially in the presence of noise, phase instability, or large external influences.

C. *Analysis of the modulation factor stability in mode 3*

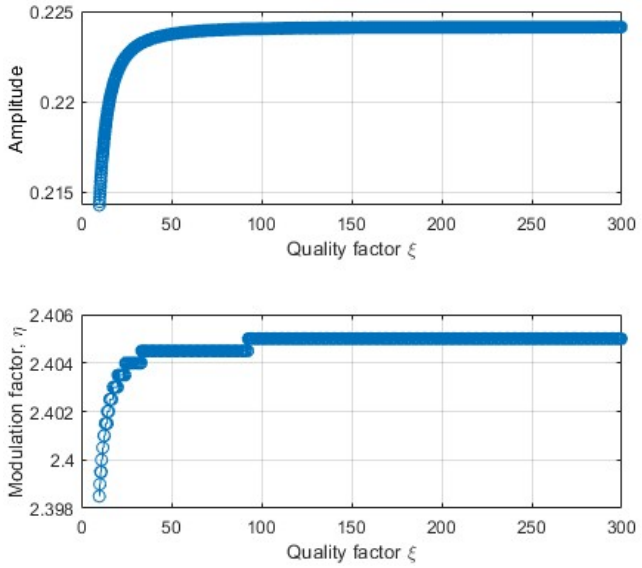


Fig. 3. A graph (top) of the signal amplitude values of Ax and Ay when they are equal, depending on the resonance quality factor in the cell and a graph (bottom) of the dependence of the optimal modulation factor (for mode 3) on the resonance quality factor

Mode 3 shows high stability over the entire scale of ξ variation. The shift of the optimal value of η does not exceed 0.05, and the signal amplitude remains stable. This allows a fixed value of η obtained during calibration to be used without having to redetermine it at each start-up. Thus, mode 3 is a reliable and practically drift-insensitive mode to start the magnetometric system.

V. SUMMING UP THE SILULATION RESULTS

Summing up the estimation of the optimal modulation factor η stability, it is necessary to take into account the mode, or more precisely, ξ, which was measured during the sensor assembly. For the sensor operation in modes 2 and 3, the optimal η can be taken as the certified value so that there is no need to redetermine it at start-up, limiting ourselves to only a rough estimation in the case of a possible impact on the QRS elements. For mode 1, this is also true, but only under the condition that ξ exceeds 50; however, in this case, the check of ξ at startup should be carried out with higher accuracy than that for modes 2 and 3.

In any other case, it is necessary to stabilize η [9, 10] in order to maintain the signal-to-noise ratio at the outputs of the QRS magnetometric system.

CONCLUSIONS

The operation of a magnetometric system of a quantum rotation sensor has been studied. Stability of three modes of the system startup to changes in the resonance quality factor has been analyzed, including the estimation of the stability of the optimal modulation factor value. It is shown that the choice of the mode should take into account the current value of the resonance quality factor registered during the assembly of the sensor or measured before its startup. The obtained results can be used to select a start-up strategy and design algorithms for automatic tuning of the magnetometric system parameters.

REFERENCES

[1] Peshekhonov, V.G., The Outlook for gyroscopy. *Gyroscopy Navig.* 2020, vol. 11, no. 3 pp. 193–197. https://doi.org/10.1134/S2075108720030062

[2] Vershovskii, A.K., Litmanovich, Y.A., Pazgalev, A.S., and Peshekhonov, V.G., Nuclear magnetic resonance gyro: Ultimate parameters. *Gyroscopy Navig.,* 2018, vol. 9, no. 3, pp. 162–176. https://doi.org/10.1134/S2075108718030100

[3] Bloch, F. Nuclear Induction. *Physical Review*, 1946, vol. 70, pp. 460–473. https://doi.org/10.1103/PhysRev.70.460

[4] Popov, E.N., Barantsev, K.A., Litvinov, A.N. Theoretical simulation of a signal for a scheme of an atomic spin gyroscope with optical detection. *Kvantovaya Elektronika*, 2019, vol. 49, no. 2, pp. 169–177. https://doi.org/10.1070/QEL16720

[5] Eklund, E. Microgyroscope based on spin-polarized nuclei. Ph.D. Thesis. 2008.

[6] Zhang, K., Luo, Z., Tang, F., Zhao, N., Wang, Y. Experimental optimization of atomic magnetometer in nuclear magnetic resonance gyroscope. *Japanese Journal of Applied Physics*, 2020. https://doi.org/10.1088/1361-6455/ab2a99;

[7] Luo, W., Zhang, H., Liu, Y., Wei, X., Zhang, X., Wang, Y., Yang, R. A dual-axis high-order harmonic and single-axis phase-insensitive demodulation atomic magnetometer for in situ NMR detection of Xe. *Journal of Applied Physics*, 2022, vol. 132, no. 14, article 144401. https://doi.org/10.1063/5.0096583

[8] Shevchenko, A.N., Chalkov, V.V. Method of adjusting the parametric modulation of the magnetometric system of a quantum rotation sensor. In: *XXIII* konferentsiya molodykh uchenykh "Navigatsiya i upravlenie dvizheniem" (*Proc. XXIII Conference of Young Scientists "Navigation and Motion Control"*), St. Petersburg: CSRI Elektropribor, 2024

[9] Li, S., Ma, D., Wang, K., Gao, Y., Xing, B., Fang, X., Han, B., Quan, W. High sensitivity closed-loop Rb optically pumped magnetometer for measuring nuclear magnetization. *Optics Express*, 2022, vol. 30, pp. 43925–43937. https://doi.org/10.1364/OE.472405

[10] Jian, Q., Li, J., Wang, Z., Zhang, Y., Luo, H. The characteristic analysis of the built-in vector atomic magnetometer in a nuclear magnetic resonance oscillator. *AIP Advances*, 2019. https://doi.org/10.1063/1.5082152

Calibrating Accelerometers During Operation of a Control System for Space Launch Vehicles

Ya. Savel'eva
JSC Academician N.A. Semikhatov Scientific and Production Association of Automatics
Yekaterinburg, Russia
avt@npoa.ru

L. Vodicheva
JSC Academician N.A. Semikhatov Scientific and Production Association of Automatics
Yekaterinburg, Russia

Yu. Parysheva
JSC Academician N.A. Semikhatov Scientific and Production Association of Automatics
Yekaterinburg, Russia

Abstract—Arrangement and technique for calibrating accelerometers during operation of a control system for space launch vehicles are considered. Vector and scalar calibration methods are studied. The results of comparative error analysis of the methods are presented.

Key words—space launch vehicle, inertial measurement unit, calibration, error analysis.

I. INTRODUCTION

Determination of initial orientation of a sensor frame of an inertial navigation system (INS) relative to a reference navigation frame (initial alignment) and estimation of accuracy parameters of an inertial measurement unit (calibration) during operation of a control system can be considered as a single task of determination of INS initial conditions [1]. According to the terminology used at Scientific and Production Association of Automatics (SPAA), this task is implemented by a special functional subsystem of a control system (CS), which is defined as a system for determination of initial conditions (SDIC). SDIC operates during the prelaunch preparation and operational testing.

A technique for precision initial self-alignment was considered by the authors in the paper presented at the 26th St Petersburg Conference of Young Scientists [2]. In this paper a technique for calibration of INS accelerometers developed at SPAA, peculiarities of its implementation for control systems of space launch vehicles (SLV), and the results of error analysis are discussed.

II. STATEMENT OF THE PROBLEM

SLV navigation is based on the inertial frame mechanization. A reference navigation inertial frame is a coordinate system Earth-fixed at the launching moment. In gimballed INS used in CS for operational and prospective SLV, the navigation frame is realised physically with the help of three-axis gyroscopic stabilizer (TGS) with three pendulum accelerometers (PA) mounted on its platform. A platform is stabilized relative to the inertial space, which does not require a precise control path of its angular position.

The mathematical model of TGS errors is developed by the manufacturer, which is Scientific and Production Association of Electromechanics (Miass). Accelerometer errors to be estimated include biases, scale factor error, and misalignments (orthogonality errors) between their sensitive axes (SA).

Misalignments between accelerometer SA $O\bar{X}$, $O\bar{Y}$, $O\bar{Z}$, which are nominally mutually orthogonal and coincide with the axes OX, OY, OZ of the orthogonal sensor frame, are defined by three small angles β_{zx}, β_{xy}, β_{yz} as it shown in Fig. 1.

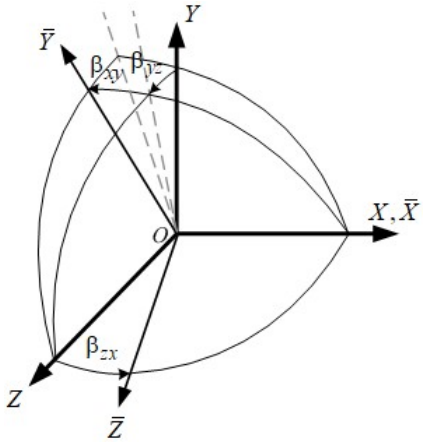


Fig. 1. Orientation of PA sensitive axes

According to Fig. 1, orthogonality errors are parametrised by the matrix of direction cosines of PA axes in the SF, which takes the form:

$$M = \begin{bmatrix} 1 & 0 & 0 \\ -\sin\beta_{xy} & \cos\beta_{yz}\cdot\cos\beta_{xy} & \cos\beta_{xy}\cdot\sin\beta_{yz} \\ \sin\beta_{zx} & 0 & \cos\beta_{zx} \end{bmatrix} \quad (1)$$

Each error is a sum of systematic and random components. Systematic error is, in turn, a sum of a value determined by a TGS manufacturer and correcting component, which is estimated during previous calibrations of TGS mounted in a control system. Systematic errors are used to correct accelerometer readings. Random errors are induced by run-to-run stability and have to be estimated during a current calibration.

III. CALIBRATION ARRANGEMENT

SDIC operates during the prelaunch preparation and various modes of control system testing. In Fig. 2, the SDIC timing

diagram, which is unified for all modes, is shown. A gyroscopic stabilized platform (GSP) is set in various programme positions with the help of servomotors and gyro torquers. After setting in a programme position, GSP is stabilised relative to the inertial space. Over stabilisation intervals, fine determination of parameters is carried out. Programme positions and times of intervals are determined so as to provide observability of estimated parameters and required calibration accuracy.

For the given error model, nine orientations are needed to calibrate accelerometers. Positions of accelerometer axes in various programme orientations are shown in Table I. In the first three programme orientations, SA of each pair of accelerometers are disposed in a vertical plane symmetrically relative to the direction of the gravity, which results in the minimum estimation error of SA misalignments. In the following six orientations, the sensitive axis of each accelerometer is aligned vertically upwards and downwards, which minimises estimation errors of biases and scale factors [3].

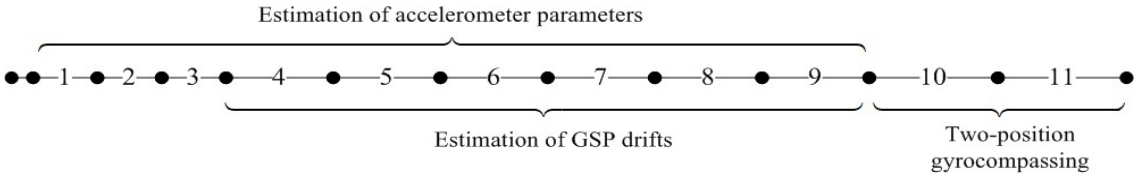


Fig. 2. SDIC timing diagram

TABLE I

Orientation of accelerometer axes in various programme positions								
Estimation of SA misalignments			***Estimation of biases and scale factors***					
f=1	f=2	f=3	f=4	f=5	f=6	f=7	f=8	f=9
$\bar{x}$ $\bar{y}$	$\bar{x}$ $\bar{z}$	$\bar{y}$ $\bar{z}$	↑ $\bar{z}$	↑ $\bar{x}$	↓ $\bar{y}$	↓ $\bar{z}$	↓ $\bar{x}$	↑ $\bar{y}$

When a platform is stabilised in the inertial space, data processing is carried out in two stages. At the first stage, which is the same for initial alignment and calibration, sensor sampled measurements are processed and generic parameters are estimated. These parameters are functions of measurement errors and misalignments between actual platform axes and their preset programme orientation. At the second stage at the end of all orientations, TGS errors are algebraically determined from generic parameters.

Calibration is implemented in the same run as a flight; results obtained are used in initial alignment and are translated to the flight mode.

IV. CALCULATION OF GENERIC PARAMETERS

Specific force projections are measured by accelerometers mounted on a platform which does not rotate in the inertial space. The vector of specific force due to the gravity is rotating with respect to accelerometer axes, which is taken into account in the algorithms.

Measurement data processing results in obtaining residuals and generic parameters, which allows linearizing measurement equations and reducing a large amount of measurement data to a few parameters, which completely describe estimated errors.

Calculation of generic parameters consists of four stages. At the first stage, systematic errors of accelerometer biases, scale factors and GSP drift rates are compensated in the specific force projections measured by accelerometers.

At the second stage, residual measurements as differences between specific force measured by accelerometers and its predicted values are calculated.

Residual measurements are calculated in the inertial frame whose axes coincide with the accelerometer axes at the beginning of a data processing interval. This allows separating data processing for each axis and isolating errors physically equivalent to the linear acceleration from errors equivalent to the GSP drift rates.

Predicted values of accelerometer measurements are calculated on the base of the reference specific force due to the gravity, its rotation relative to the inertial space and actual orientation of accelerometer axes, which is determined using gimbal angles measured by angular sensors. Residual measurements obtained under the assumption of zero alignment errors are as follows

$$\begin{bmatrix} \Delta\dot{w}_{\bar{x}}(t) \\ \Delta\dot{w}_{\bar{y}}(t) \\ \Delta\dot{w}_{\bar{z}}(t) \end{bmatrix} = \begin{bmatrix} \hat{\dot{w}}_{\bar{x}}(t) \\ \hat{\dot{w}}_{\bar{y}}(t) \\ \hat{\dot{w}}_{\bar{z}}(t) \end{bmatrix} - \begin{bmatrix} g_{\bar{x}}(t) \\ g_{\bar{y}}(t) \\ g_{\bar{z}}(t) \end{bmatrix} =$$

$$= \begin{bmatrix} -\rho_{\bar{x}} & 0 & 0 \\ -\beta_{xy} & -\rho_{\bar{y}} & \beta_{yz} \\ \beta_{zx} & 0 & -\rho_{\bar{z}} \end{bmatrix} \cdot R_f \cdot W(t) \cdot \begin{bmatrix} 0 \\ g \\ 0 \end{bmatrix} + \begin{bmatrix} \tau_{\bar{x}} \\ \tau_{\bar{y}} \\ \tau_{\bar{z}} \end{bmatrix}$$

$$W(t) = \begin{bmatrix} \cos^2\varphi_c + \sin^2\varphi_c \cos\Omega t & \sin\varphi_c \cos\varphi_c (1-\cos\Omega t) & \sin\varphi_c \sin\Omega t \\ \sin\varphi_c \cos\varphi_c (1-\cos\Omega t) & \cos^2\varphi_c \cos\Omega t + \sin^2\varphi_c & -\cos\varphi_c \sin\Omega t \\ -\sin\varphi_c \sin\Omega t & \cos\varphi_c \sin\Omega t & \cos\Omega t \end{bmatrix} \quad (2)$$

where $\rho_{æ}$, $\tau_{æ}$, $æ = \bar{x}, \bar{y}, \bar{z}$ are estimated random components of accelerometer scale factor errors and biases; R_f is a matrix

calculated using angular sensor outputs at the beginning of orientation; φ_c is a launch site latitude; Ω is the value of Earth angular rate; f is a number of orientations.

At the third stage, residual acceleration is twice integrated to smooth effectively measurement noise and input disturbances. To provide the required level of methodical errors, it is sufficient to assume that the desired acceleration is a linear function of time. In this case, the double integral of it is equivalent to linear displacement and, for each axis, can be represented by a third order polynomial:

$$s_{æ}(t) = s_{æ,0} + v_{æ,0} \cdot t + a_{æ,0} \cdot \frac{t^2}{2} + b_{æ,0} \cdot \frac{t^3}{6}, \; æ = \overline{x}, \overline{y}, \overline{z}. \quad (3)$$

At the fourth stage, generic parameters, which are the coefficients of the polynomial (3), are estimated with the help of Kalman filter.

Generic parameters are subdivided into two groups. They are $a_{æ,0}$ equivalent to the linear acceleration, which are functions of levelling errors and accelerometer errors, and $b_{æ,0}$ equivalent to the GSP drift rates, which are functions of azimuth alignment error and GSP drift rate components.

Using simple algebraic expressions, accuracy parameters of accelerometers are estimated using generic parameters $a_{æ,0}$ obtained in nine orientations.

V. Methods for Calibration of Accelerometers

In operating CS, it is the vector method that is used for the calibration of accelerometers. In this case, the calibration standards are projections of the gravity vector onto the accelerometer sensitive axes. This requires a precise fine levelling, which is provided by gyro torquers and estimation of SLV tilts prior to calibration. Errors of alignment of accelerometer axes relative to a programme orientation do not exceed some arc min.

The vector method allows for decomposition of measurement equations and expressions for error estimation take a simple form. Scale factor errors and biases of each accelerometer are determined from two orientations when their axes are aligned vertically upwards and downwards. The expressions for estimation of scale factor errors and biases are as follows

$$\rho^A_{æ} = \frac{a^{\downarrow}_{æ,0} - a^{\uparrow}_{æ,0}}{2 \cdot g}, \; \tau^A_{æ} = \frac{a^{\downarrow}_{æ,0} + a^{\uparrow}_{æ,0}}{2}, \; æ = \overline{x}, \overline{y}, \overline{z}. \quad (4)$$

Relative estimation errors are of the second order of smallness, which allows determining these parameters with 3-4 true significant decimal digits.

Orthogonality errors are evaluated in two stages in the vector calibration method. Here are the expressions for estimating orthogonality error between the accelerometer axes *X* and *Z*, which is observable in the second orientation.

At the first step, the intermediate parameter, which is the sum of generic parameters of the axes *X* and *Z*:

$$\phi_{xz} = \frac{a^{(2)}_{\overline{x},0} + a^{(2)}_{\overline{z},0}}{g} = \frac{1}{g} \cdot \left(\cancel{\frac{\sqrt{2}}{2} \cdot g \cdot \Delta\psi} - \frac{\sqrt{2}}{2} \cdot g \cdot \rho^A_{\overline{x}} + \tau^A_{\overline{x}} \right) + \\ + \frac{1}{g} \cdot \left(\cancel{-\frac{\sqrt{2}}{2} \cdot g \cdot \Delta\psi} + \frac{\sqrt{2}}{2} \cdot g \cdot \beta_{zx} - \frac{\sqrt{2}}{2} \cdot g \cdot \rho^A_{\overline{z}} + \tau^A_{\overline{z}} \right) \quad (5)$$

It can be seen from Expression (5) that these parameters are functions of the levelling error $\Delta\psi$ about axis orthogonal to the plane *OXZ* and errors of accelerometers. Since the levelling error is included in the generic parameters with different signs under chosen orientations, it is eliminated from their sum.

At the second step, scale factor errors and biases calculated before are taken into account:

$$\Delta\beta_{zx} = -\frac{\sqrt{2}}{2} \cdot \left(\rho^A_{\overline{x}} + \rho^A_{\overline{z}} \right) + \frac{1}{g} \left(\tau^A_{\overline{x}} + \tau^A_{\overline{z}} \right)$$
$$\beta_{zx} = \sqrt{2} \cdot \left(\phi_{xz} - \Delta\beta_{zx} \right) \quad (6)$$

The scalar method, which is widely used for calibration of strapdown inertial measurement units, was studied for calibration of accelerometers for SLV control systems under development as well. In this case, the calibration standard is the module of a measured vector, which does not require SA orientations to be known.

The measurement equations for the used version of scalar calibration are similar to those described in the paper of the authors at this conference [5] and take the form of the system of nine linearized equations with nine unknowns. In these equations, averaged accelerometer outputs are used. In the SDIC algorithms under consideration, averaging is implemented on the level of residuals. The entire measurements, which are required for the scalar method, are restored with the help of the generic parameters $a_{æ,0}$, $æ = \overline{x}, \overline{y}, \overline{z}$, equivalent to the linear acceleration and predicted values of measured specific force, which has been subtracted from accelerometer outputs.

To choose a calibration method, their comparative analysis was implemented.

VI. Results of Error Analysis

Calibration errors were estimated with the help of mathematical simulation using the C++ programme. In the calculations, TGS instrument errors and SLV body oscillations due to wind at a launch complex, which results in distortion of outputs of accelerometers, were simulated. Calibration errors were calculated separately for each error source. Total variance was calculated as a sum of variances. Results of estimation of calibration errors are shown in Table II.

TABLE II

Calibration errors			
Parameter	***Biases, m/s²***	***Scale factor errors, %***	***Misalignments, "***
Run-to-run stability	$6 \cdot 10^{-3}$	6·10–2	45
Errors of vector calibration	$(1,2\text{-}1,5) \cdot 10^{-5}$	$(1,5\text{-}1,9) \cdot 10^{-4}$	0,5-0,7
Errors of scalar calibration	$(1,2\text{-}1,5) \cdot 10^{-5}$	$(1,6\text{-}1,8) \cdot 10^{-4}$	0,5-0,7
Requirements of calibration accuracy	$2 \cdot 10^{-4}$	$2 \cdot 10^{-3}$	5

It can be seen from Table II that errors of the methods are close to each other and are an order less than accuracy requirements. Random components of biases and scale factor errors specified for simulation coincide with the estimation results with accuracy of three significant decimal digits, misalignments with accuracy of two significant decimal digits.

As a result of the comparative analysis, it was decided to include both vector and scalar calibration methods in SDIC of control systems for SLV being developed, which improves SDIC reliability.

VII. Conclusion

The technique including the vector and scalar methods for calibration of accelerometers developed at SPAA allows estimating their errors with accuracy from two to three true significant decimal errors during the operation of the SLV control system.

Comparative error analysis of the vector and scalar calibration methods has shown that errors of both methods are almost the same. Accuracy of the vector method is provided by rather precise alignment of accelerometer axes to the required orientation and minimization of calibration errors due to alignment errors by choice of orientations and algorithms.

To improve the reliability of SLV control systems, it is worth using both methods.

Литература

[1] Belsky, L. and Vodicheva, L., Rapid High Precision Initial Alignment and Calibration of Inertial Navigation System of a Flying Vehicle on a Moving Base, Giroskopiya I Navigatsiya, 2001, no. 4, pp. 3–18 (in Russian).

[2] Vodicheva, L. and Savel'eva, Ya., Precision Initial Self-alignment of a Gimballed Inertial Navigation System, 26th St. Petersburg Conference of Young Scientists, St. Petersburg, CSRI Elektropribor, 2024, pp. 239-241 (in Russian).

[3] Vodicheva, L. and Parysheva, Yu., Estimating the Accuracy Parameters of Sensors in a Strapdown Inertial Measurement Unit with the Use of a Relatively Coarse Turntable, Gyroscopy and Navigation, 2019, Vol. 10, No. 4, pp. 303-312.

[4] Belsky, L., Vodicheva, L., and Parysheva, Yu., Estimation of Generic Parameters in a Technique for Initial Alignment and Calibration of INS for Space Launch Vehicles, 28th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2021. https://ieeexplore.ieee.org/document/9470830

[5] Vodicheva, L., Parysheva, Yu., Savel'eva, Ya., Kabanova, D., and Shmakov, N., Calibration of a Strapdown Inertial Measurement Unit with Two Angular Rate Sensors, 32nd St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2025.

Research on the Technology for Improving Vibration Performance of High Precision Fiber Optic Gyroscope

Pei Zhang
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China
1184526321@qq.com

Yichen Wu
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Shitao Ji
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Zerun Zang
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Bo Xie
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Wei Hong
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Bin Ren
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Yihua Wang
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Honggang Chen
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, China

Abstract*—The fiber optic gyroscope (FOG) will produce nonreciprocal phase error under vibration environment, which will directly affect the accuracy of FOG, and is the main factor restricting the further engineering of high precision FOG. Firstly, the mathematical model of vibration error of the FOG is studied in this paper, and the main cause of the vibration error of FOG is identified as the change of fiber length and fiber refractive index caused by the deformation of fiber coil under vibration environment. To solve this problem, the structure of high precision FOG is analyzed , and the main factor affecting the deformation of fiber coil is determined to be the fiber coil module. The fiber coil module is mainly composed of skeleton, cover plate, fiber coil, Y-waveguide, coupler and so on. The skeleton provides the installation benchmark and support for the fiber coil, and completes the overall encapsulation of the fiber coil with the cover plate. In the vibration environment, the stiffness of the fiber coil module directly affects the deformation of the fiber coil, which in turn affects the vibration performance of the FOG. Nextly, the finite element simulation model for random vibration of fiber coil module is established. According to the structure form and composition of optical fiber ring module, the finite element model of optical fiber ring module assembly is established based on ANSYS software. On this basis, the mode analysis of the fiber coil module is carried out first, and then the random vibration simulation analysis is carried out. The vibration performance of the FOG can be quantitatively analyzed by extracting the deformation of fiber coil under random vibration condition. Then, the influence factors such as the fixed connection method between the skeleton and the cover plate, the structure form of the fiber coil module, the distribution and number of installation holes in the skeleton, and the thickness of the skeleton support wall are analyzed, and the structure optimization scheme of the fiber coil module is determined. On this basis, the structural optimization design of the fiber coil module is carried out, and the first order natural frequency of the optimized fiber coil module is increased from 2327Hz to 5353Hz, and the deformation of fiber coil in vertical direction decreased from 1.53×10^{-4}mm to 4.94×10^{-5}mm under 6.06g random vibration environment. Finally, based on the above simulation results, an optimized physical prototype of the fiber coil module is built and assembled into the FOG for random vibration test verification. The test results show that the vibration performance of the FOG is improved from 0.2183°/h to 0.0262°/h, which is of great significance for further realizing the engineering of high precision FOG.

Keywords—fiber optic gyroscope, vibration performance, fiber coil module, finite element simulation, optimal design

I. INTRODUCTION

The FOG has been widely used in fields of missile, ship, satellite and other carriers for its advantages such as all solid state, high dynamic, low cost and long lifetime. However, as the core component of fiber optic gyroscope, the fiber coil is very sensitive to vibration. The FOG will produce nonreciprocal phase error under vibration environment, which will directly affect the accuracy of FOG, and is the main factor restricting the further engineering of high precision FOG[1-4]. At present, the main methods to suppress the vibration error of FOG are to change the fiber coil wrapping process, design shock absorbers, optimize software algorithms and the material of the gyroscope body. Y.J. Wu effectively reduced the impact of mechanical vibration on the performance of FOG through the fourth-order symmetric winding method[5]. In order to reduce the influence of random vibration on the accuracy of FOG during rocket launch, L.H. Xu designed a three-axis equal-stiffness rubber shock absorber to install the FOG bracket in the space of eight points, which effectively suppressed the magnitude of random vibration[6]. R. Song proposed a

modeling method for FOG vibration error based on improved local mean decomposition, which effectively suppresses the noise and drift in the output of FOG under vibration environment[7]. J.X. Chai proposed a high-precision modeling method for the random vibration error of FOG, which combines the empirical mode decomposition method, time series modeling method and Kalman filtering algorithm, and realizes the high-precision fitting of the vibration error of fiber optic gyro, which can be applied to the high-precision compensation of the random vibration error of fiber optic gyro in engineering environment[8]. Other studies have shown that replacing aluminum alloy with beryllium aluminum materials with higher specific stiffness can effectively improve the overall stiffness of FOG[9-10].

Firstly, the mathematical model of vibration error of the FOG is studied in this paper, and the main cause of the vibration error of FOG is identified as the change of fiber length and fiber refractive index caused by the deformation of fiber coil under vibration environment. Then the structure of high precision FOG is analyzed , and the main factor affecting the deformation of fiber coil is determined to be the fiber coil module. Nextly, the finite element simulation model for random vibration of fiber coil module is established, which can quantitatively calculate the deformation of fiber coil under random vibration condition. On this basis, the influence factors such as the fixed connection method between the skeleton and the cover plate, the structure form of the fiber coil module, the distribution and number of installation holes in the skeleton, and the thickness of the skeleton support wall are analyzed, and the structure optimization scheme of the fiber coil module is determined. Finally, based on the above simulation results, an optimized physical prototype of the fiber coil module is built and assembled into the FOG for random vibration test verification. The test results show that the vibration performance of the FOG is improved from 0.2183°/h to 0.0262°/h, which effectively improve the vibration performance of high precision FOG.

II. THEORY

A. Fiber Coil Module

The fiber coil module is the core component of FOG, which forms a closed-loop feedback fiber optic gyroscope with digital demodulation circuits, light sources, etc., to measure the external angular velocity. The function of the fiber coil module is to introduce the optical signal of the light source into the FOG for Sagnac effect, and then send the interference signal to the photodetector for photoelectric conversion. The fiber coil module is mainly composed of fiber coil, Y-waveguide, coupler, skeleton, and cover plate. Its realization principle is based on the Sagnac effect. It uses the interference fringes formed by the optical path to sense the angular velocity input signal, resulting in a phase change of light intensity, so as to measure the external angular velocity. The main functions of the fiber coil module are as follows:

1) Provide interfaces with light source and photodetector.

2) Realize the polarization and beam splitting of the output light of the light source.

3) Complete the modulation of the positive and negative beams in the fiber coil.

4) Provide optical transmission path to detect weak angular velocity inputs.

The schematic diagram of the fiber coil module is shown in Fig. 1.

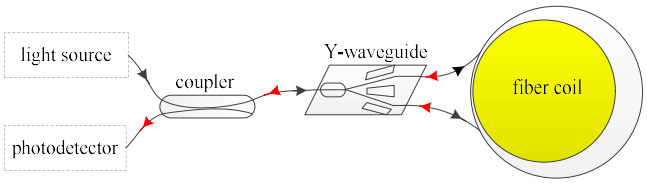


Fig. 1. The schematic diagram of fiber coil module

The fiber coil is fixed in the annular groove of the skeleton, and the cover plate and skeleton are installed together to complete the encapsulation of the fiber coil. The Y-waveguide and coupler are installed on the skeleton. The structure of the fiber coil module is shown in Fig. 2.

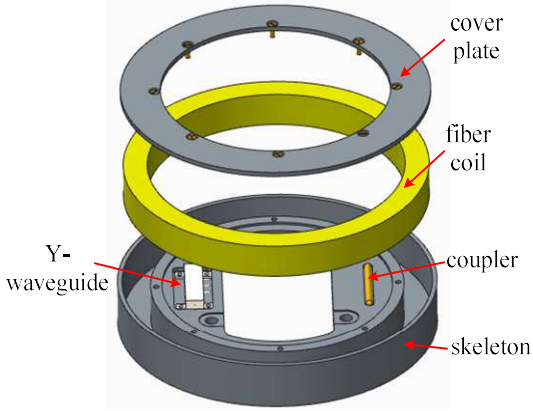


Fig. 2. Structure of the fiber coil module

B. Theoretical Analysis of Vibration Error of FOG

The vibration sensitivity of FOG is caused by the change of fiber length and fiber refractive index caused by the deformation of fiber coil under vibration environment. When the deformation caused by vibration is asymmetrically distributed along the coil, it is equivalent to a phase modulation, resulting in nonreciprocal phase error, which can be expressed as equation 1:

$$\begin{aligned}\Omega_e(t)=&\frac{n^2}{DL}\int_0^L\frac{\partial\varepsilon_\varphi(s,t)}{\partial t}(L-2s)ds\\&+\frac{nK_r}{DL}\int_0^L\frac{\partial\varepsilon_r(s,t)}{\partial t}(L-2s)ds\\&+\frac{nK_\varphi}{DL}\int_0^L\frac{\partial\varepsilon_\varphi(s,t)}{\partial t}(L-2s)ds\\&+\frac{nK_s}{DL}\int_0^L\frac{\partial\varepsilon_s(s,t)}{\partial t}(L-2s)ds\end{aligned}\quad\square\square\square\quad(1)$$

where

$K_r=-\frac{n^3}{2}(p_{12}|\cos\theta|+p_{11}|\sin\theta|)$,

$K_\varphi=-\frac{n^3}{2}p_{12}$,

$K_s=-\frac{n^3}{2}(p_{12}|\sin\theta|+p_{11}|\cos\theta|)$,

θ twisting angle of fiber,

n refractive index of fiber under no stress,

ε_r strain coefficient in the radial direction of fiber coil in cylindrical coordinate system,

ε_φ strain coefficient in the angular direction of fiber coil in cylindrical coordinate system,

ε_s strain coefficient in the axial direction of fiber coil in cylindrical coordinate system,

s variable of fiber length,

t time.

The first term of equation (1) represents the error caused by changes in fiber length, while the last three terms represent the error caused by the elastic optical effect.

It can be seen from equation (1) that the nonreciprocal phase error caused by vibration is related to the deformation of the fiber coil. To reduce the error of FOG, the deformation of the fiber coil in the vibration environment should be minimized. The deformation of fiber coil mainly depends on the overall stiffness of fiber coil module, so it is necessary to optimize the structure of fiber coil module.

III. Establish the Random Vibration Simulation Platform of Fiber Coil Module

A. Physical Parameter Analysis of Fiber coil Module

The fiber coil used for high precision FOG is a skeleton free fiber coil. The fiber is wound layer by layer on the special coil winding tool according to the multi-stage symmetric winding method by brushing winding adhesive. After the winding is completed, the winding adhesive is cured, and then the fiber coil is removed from the winding fixture to form a skeleton free fiber coil. The fiber is composed of cladding and coating layer. The rectangular cross section of the fiber coil is shown in Fig. 3.

The fiber coil used in high precision FOG is skeleton free fiber coil. It's a hollow cylindrical structure with rectangular cross section is shown in Fig. 3.

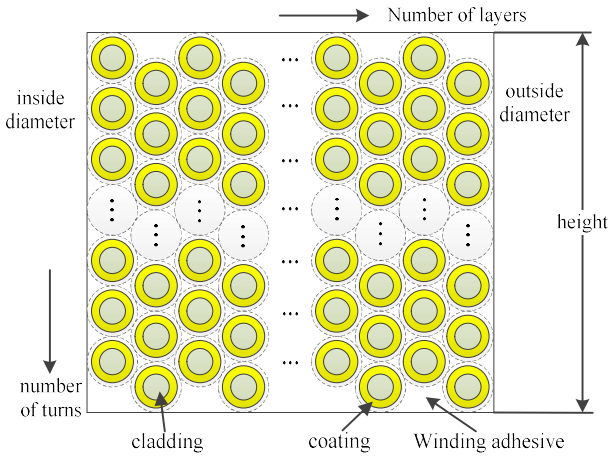


Fig. 3. The rectangular cross section of fiber coil

This scheme uses 80/135 specification optical fibers for the fiber coil, which is wound with 16-pole symmetrical method. The number of layers is 80, the number of turns is 109, and the inner diameter of the fiber ring is 113mm. In order to facilitate calculation, the fiber coil is regarded as a whole when the random vibration simulation analysis of the fiber coil module is carried out. The equivalent physical parameters of the fiber coil are calculated by the volume percentage of each component. The skeleton and cover plate materials are made of iron nickel alloy, and the simulated physical parameters are shown in Table I.

TABLE I. Physical Parameters of Fiber Coil Module

Physical Parameters	V(%)	Density (g·cm-3)	Elastic modulus (GPa)	Poisson's ratio
Cladding	30.7	2.65	70	0.168
Coating	56.8	1.2	2	0.38
winding adhesive	12.5	1	1	0.45
Fiber coil	-	1.62	22.75	0.32
skeleton	-	8.6	206	0.3
cover plate	-	8.6	206	0.3

B. Random Vibration Simulation Analysis of Fiber Coil Module

The random vibration simulation platform of fiber coil module is built with ANSYS software. First, perform modal analysis on the fiber coil module, set the physical parameters corresponding to the fiber coil, skeleton, and cover plate and fix the mounting holes of the skeleton. On this basis, the simulation analysis of 6.06g random vibration in the vertical direction of the fiber coil module is carried out, and the power spectrum of random vibration is shown in Fig. 4.

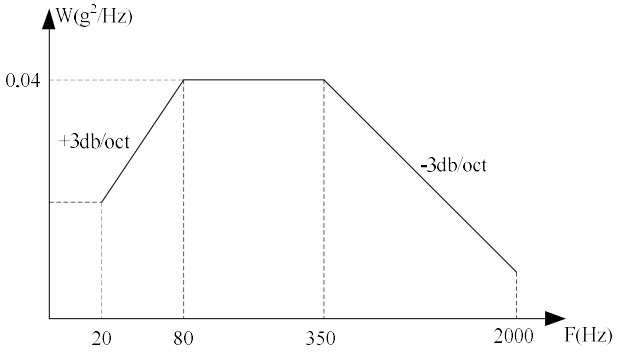
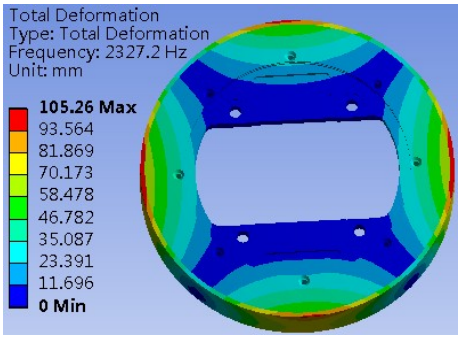
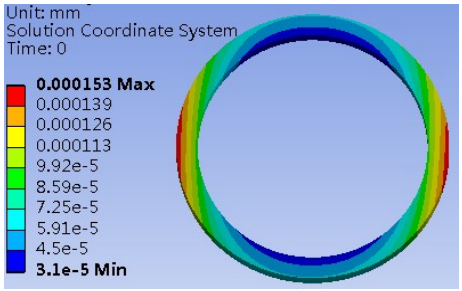


Fig. 4. The power spectrum of random vibration

Based on the built random vibration simulation platform, the fiber coil module is simulated and analyzed. The first order natural frequency of the fiber coil module is 2327Hz, and the maximum vertical deformation of the fiber coil is 1.53×10^{-4}mm. The simulation results are shown in Fig. 5.



(a) First order natural frequency and mode shape



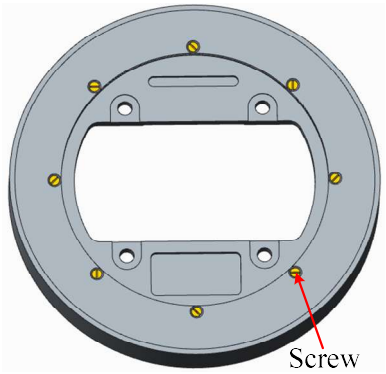
(b) Fiber coil deformation in vertical direction

Fig. 5. Simulation results of random vibration

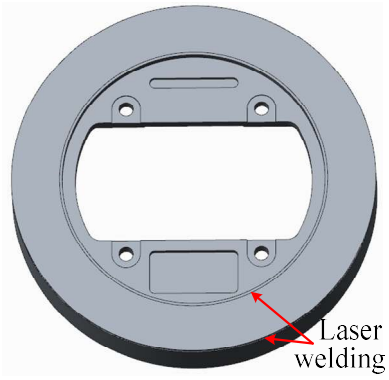
IV. Optimization Design of Fiber Coil Module Structure

A. Optimization Design of Connection Mode

Optimize the structural connection between the skeleton and cover plate by changing the screw fixation connection to laser welding of the inner and outer rings, as shown in Fig. 6.



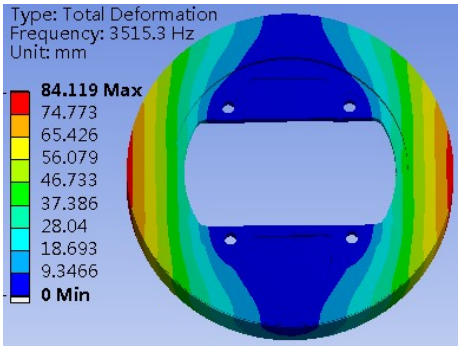
(a) Screw fixation connection



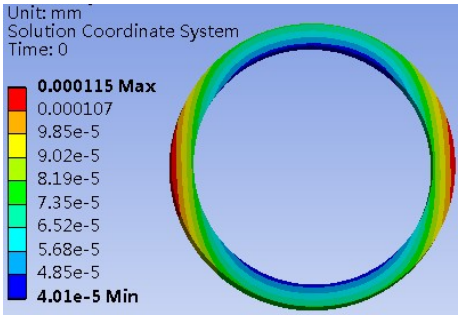
(b) Laser welding

Fig. 6. Optimization design of connection mode

By fixing the skeleton and cover plate through laser welding, the first order natural frequency of the fiber coil module increased from 2327Hz to 3515Hz, and the maximum vertical deformation of the fiber coil decreased from 1.53×10^{-4}mm to 1.15×10^{-4}mm. The simulation results are shown in Fig. 7.



(a) First order natural frequency and mode shape



(b) Fiber coil deformation in vertical direction

Fig. 7. Simulation results of laser welding

B. Optimization Design of Structural Form

In order to further meet the lightweight requirements, the structure of the fiber coil module is optimized and the transition plate is designed based on the laser welding connection between the skeleton and the cover plate. The Y-waveguide and coupler are installed on the transition plate. The transition plate is made of aluminum alloy, and the transition plate is fixed to the skeleton through screws, as shown in Fig. 8.

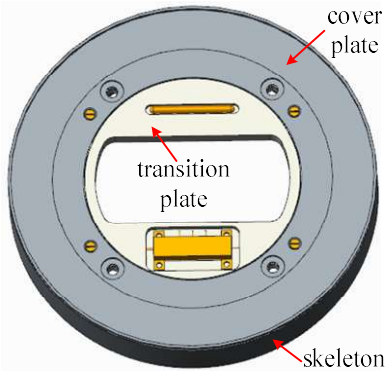
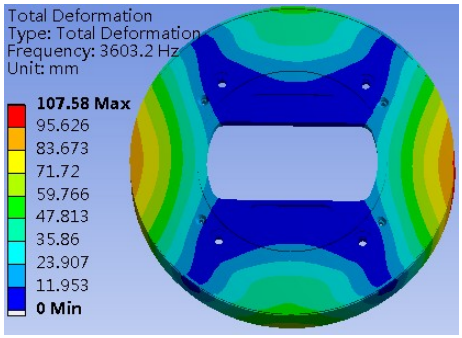
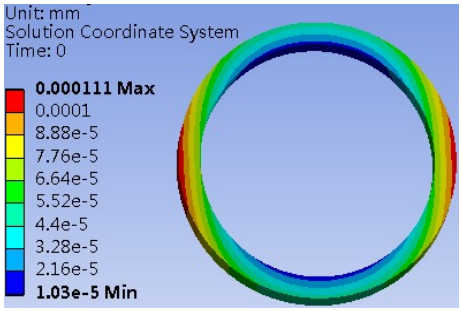


Fig. 8. Optimization design of structural form

By designing the transition plate to optimize the structural form of the fiber coil module, the first order natural frequency of the fiber coil module increased from 3515Hz to 3603Hz, and the maximum vertical deformation of the fiber coil decreased from 1.15×10^{-4}mm to 1.11×10^{-4}mm. The simulation results are shown in Fig. 9.



(a) First order natural frequency and mode shape

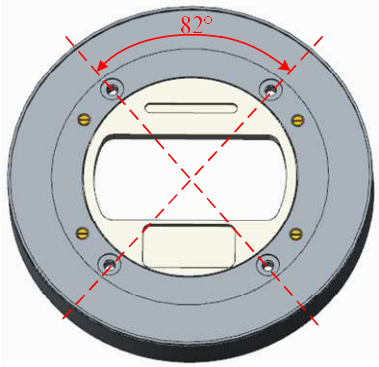


(b) Fiber coil deformation in vertical direction

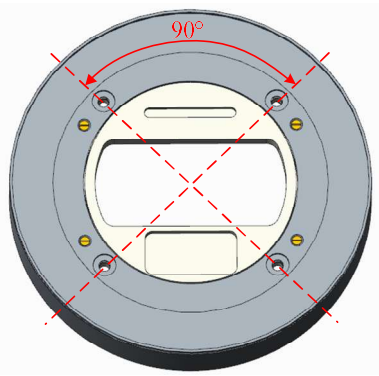
Fig. 9. Simulation results of optimized structural form

C. *Optimization Design of Skeleton Installation Holes Distribution*

On the basis of optimizing the structural form, the distribution angle of the installation holes on the skeleton is changed from 82 ° to 90 °, that is, the four installation holes are evenly distributed around the circumference, as shown in Fig. 10.



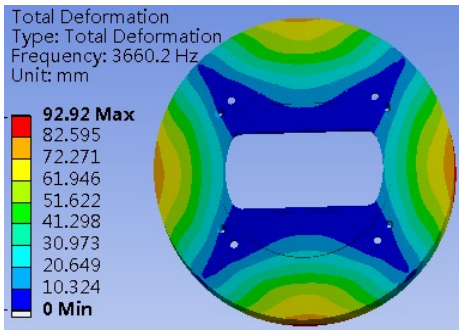
(a) Before optimization



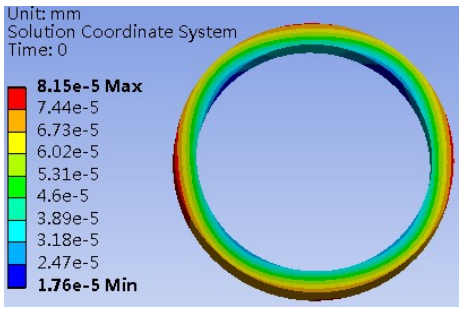
(b) After optimization

Fig. 10. Optimization design of skeleton installation holes distribution

By designing the four installation holes of the skeleton to be evenly distributed along the circumference, the first order natural frequency of the fiber coil module is increased from 3603Hz to 3660Hz, and the maximum vertical deformation of fiber coil is reduced from 1.11×10^{-4}mm to 8.15×10^{-5}mm. The symmetrical distribution of the mounting holes can effectively reduce the deformation of the fiber coil in vibration environments. The simulation results are shown in Fig. 11.



(a) First order natural frequency and mode shape



(b) Fiber coil deformation in vertical direction

Fig. 11. Simulation results of 4 installation holes evenly distributed

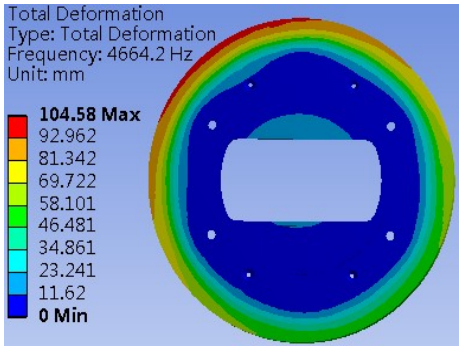
D. *Optimization Design of the Number of Skeleton Installation Holes*

The scheme of different number of skeleton installation holes is designed, and then random vibration simulation analysis of fiber coil module is carried out. The design scheme and analysis results are shown in Table II.

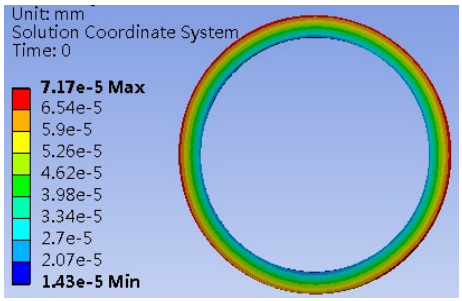
TABLE II. DESIGN SCHEME AND SIMULATION RESULTS

Number of skeleton installation holes	Angle between installation holes	First order natural frequency (Hz)	Maximum vertical deformation of fiber coil (mm)
4	90°	3660	8.15×10^{-5}
6	60°	4664	7.17×10^{-5}
8	45°	4876	6.31×10^{-5}
12	30°	4979	5.78×10^{-5}

The random vibration simulation results of fiber coil module with 6 skeleton installation holes are shown in Fig. 12.



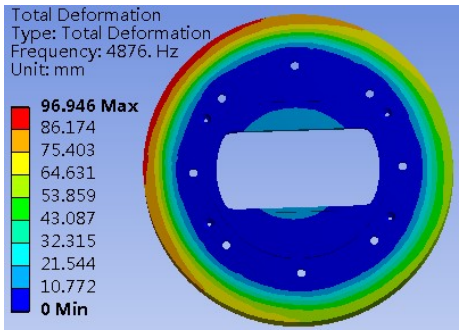
(a) First order natural frequency and mode shape



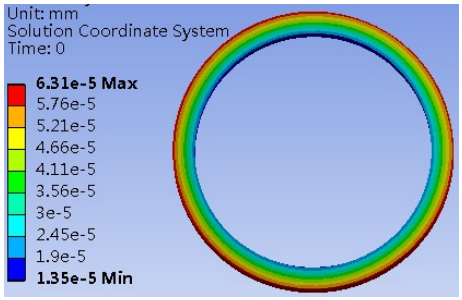
(b) Fiber coil deformation in vertical direction

Fig. 12. Simulation results of 6 installation holes evenly distributed

The random vibration simulation results of fiber coil module with 8 skeleton installation holes are shown in Fig. 13.



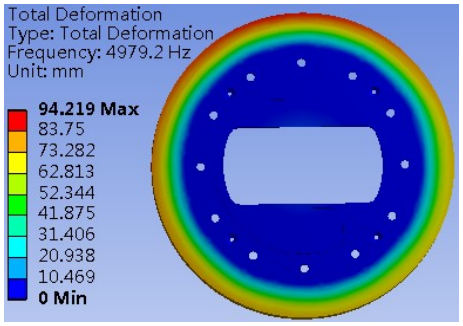
(a) First order natural frequency and mode shape



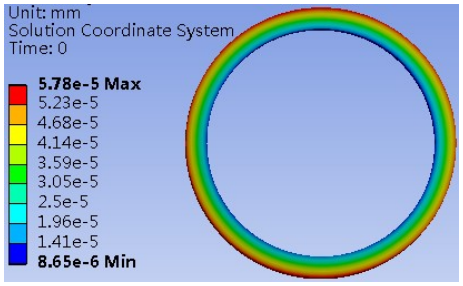
(b) Fiber coil deformation in vertical direction

Fig. 13. Simulation results of 8 installation holes evenly distributed

The random vibration simulation results of fiber coil module with 12 skeleton installation holes are shown in Fig. 14.



(a) First order natural frequency and mode shape



(b) Fiber coil deformation in vertical direction

Fig. 14. Simulation results of 12 installation holes evenly distributed

The simulation results show that increasing the number of skeleton installation holes can effectively improve the overall stiffness of fiber coil module and reduce the deformation of fiber coil.

E. Optimization Design of Skeleton Annular Groove Wall Thickness

The annular groove structure for installing fiber coil on the skeleton is shown in Fig. 15.

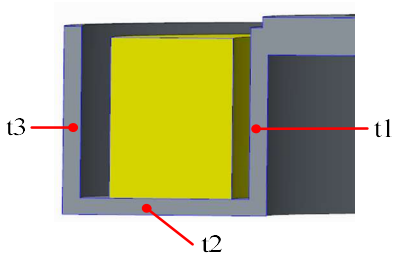


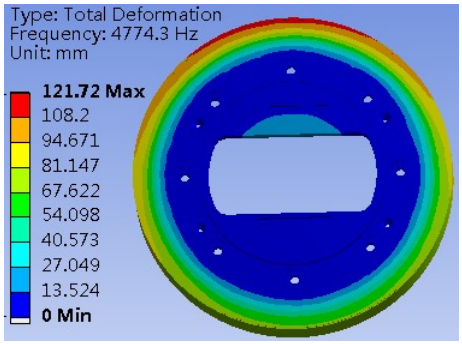
Fig. 15. Skeleton annular groove structure

The wall thickness of the skeleton annular groove for the 8 skeleton installation holes scheme is optimized, and the design scheme and random vibration results of the fiber coil module are shown in Table III.

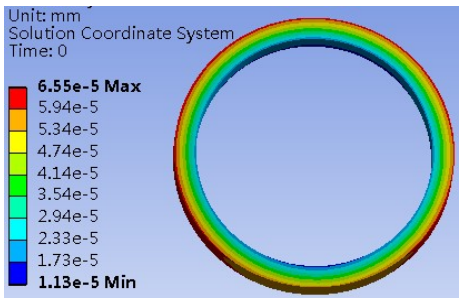
TABLE III. DESIGN SCHEME AND SIMULATION RESULTS

Scheme	Scheme I	Scheme II	Scheme III	Scheme IV
t1 (mm)	1	2	2	2
t2 (mm)	1	1	2	2
t3 (mm)	1	1	1	2
First order natural frequency (Hz)	4774	5085	5268	4869
Maximum vertical deformation of fiber coil (mm)	6.55×10^{-5}	5.46×10^{-5}	5.22×10^{-5}	6.36×10^{-5}

The simulation results of random vibration of fiber coil module when t1=1, t2=1 and t3=1 are shown in Fig. 16.



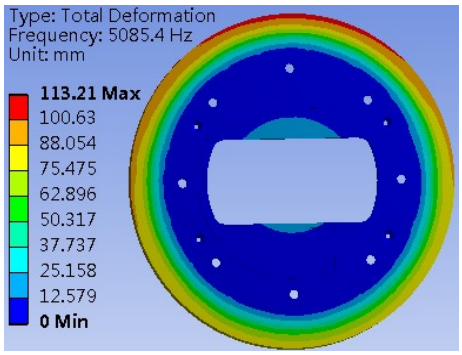
(a) First order natural frequency and mode shape



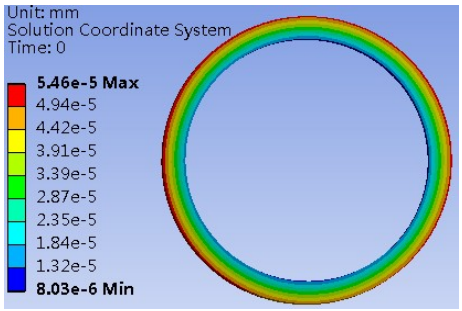
(b) Fiber coil deformation in vertical direction

Fig. 16. Simulation results (t1=1, t2=1, t3=1)

The simulation results of random vibration of fiber coil module when t1=2, t2=1 and t3=1 are shown in Fig. 17.



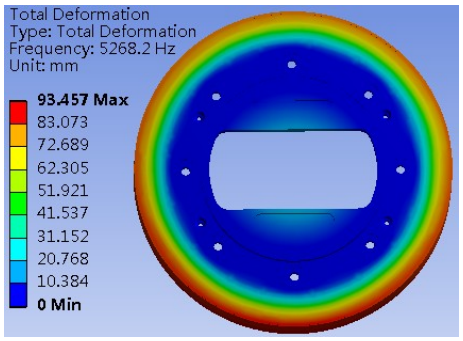
(a) First order natural frequency and mode shape



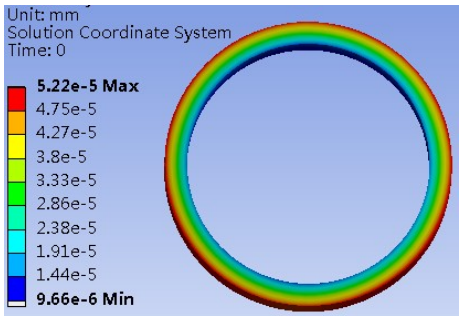
(b) Fiber coil deformation in vertical direction

Fig. 17. Simulation results (t1=2, t2=1, t3=1)

The simulation results of random vibration of fiber coil module when t1=2, t2=2 and t3=1 are shown in Fig. 18.



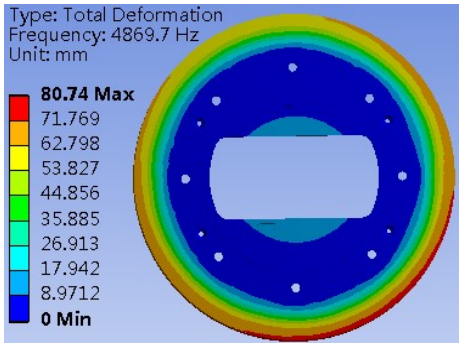
(a) First order natural frequency and mode shape



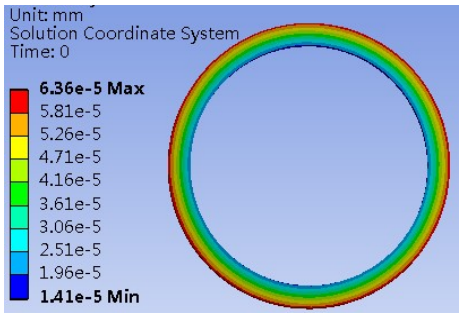
(b) Fiber coil deformation in vertical direction

Fig. 18. Simulation results (t1=2, t2=2, t3=1)

The simulation results of random vibration of fiber coil module when t1=2, t2=2 and t3=2 are shown in Fig. 19.



(a) First order natural frequency and mode shape



(b) Fiber coil deformation in vertical direction

Fig. 19. Simulation results (t1=2, t2=2, t3=2)

The simulation results show that increasing the thickness of t1 and t2 and decreasing the thickness of t3 can effectively improve the overall stiffness of the fiber coil module and reduce the deformation of fiber coil.

F. Optimization Design of the Fiber Coil Module Structure

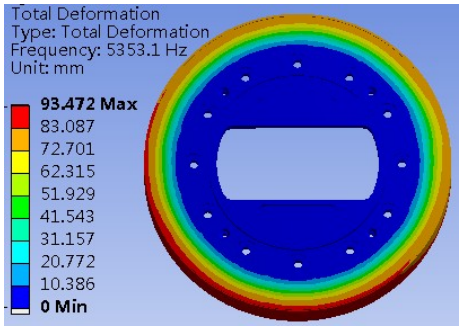
Based on the above analysis results and considering the requirements for fiber coil module volume and weight, the

optimized structure of the fiber coil module is determined as shown in Table IV.

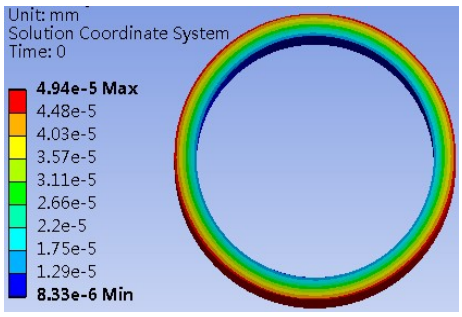
TABLE IV. OPTIMIZATION SCHEME FOR FIBER COIL MODULE STRUCTURE

Scheme	Before optimization	After optimization
Connection between skeleton and cover plate	Screw	Laser welding
Structural form	Skeleton, Cover plate	Skeleton, Cover plate and Transition plate
Number of skeleton installation holes	4	12
Angle between installation holes	82°	30°
The circular groove wall thickness of skeleton (mm)	t1=t2=t3=1.5	t1=2, t2=2, t3=1
First order natural frequency (Hz)	2327	5353
Maximum vertical deformation of fiber coil (mm)	1.53×10^{-4}	4.94×10^{-5}

The first order natural frequency of the fiber coil module is increased from 2327Hz to 5353Hz, and the maximum vertical deformation of the fiber coil is reduced from 1.53×10^{-4}mm to 4.94×10^{-5}mm after optimization. The simulation results are shown in Fig. 20.



(a) First order natural frequency and mode shape



(b) Fiber coil deformation in vertical direction

Fig. 20. The optimized simulation results

V. RANDOM VIBRATION TEST

The 6.06g random vibration test is conducted on the optimized fiber coil module after building FOG, as shown in Fig. 21.

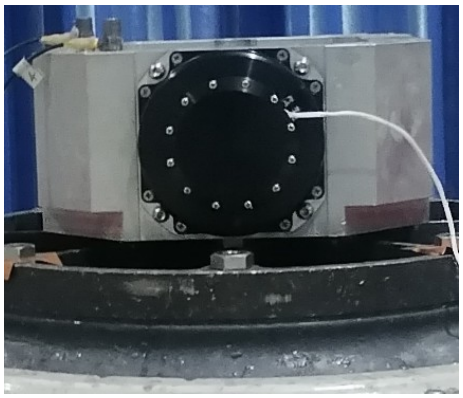


Fig. 21. Random vibration test

The results of the random vibration test show that the vibration performance of FOG has increased from 0.2183 °/h to 0.0262 °/h. The test results are shown in Table V.

TABLE V. RANDOM VIBRATION TEST RESULTS OF FOG

Random vibration	Before optimization	After optimization
Zero bias variation in X direction（°/h）	0.2183	0.0262
Zero bias variation in Y direction（°/h）	0.1655	0.0197
Zero bias variation in Z direction（°/h）	0.1871	0.0213

VI. CONCLUSION

The mathematical model of vibration error of the FOG is studied in this paper, and the main cause of the vibration error of FOG is identified as the deformation of fiber coil under vibration environment. By analyzing the structure and composition of the fiber coil module, the finite element simulation analysis model of the fiber coil module random vibration is established, which can quantitatively analyze the overall stiffness of fiber coil module and the deformation of fiber coil under random vibration condition. On this basis, the influence factors such as the fixed connection mode between the skeleton and the cover plate, the structure form of the fiber coil module, the distribution and number of installation holes in the skeleton, and the thickness of the skeleton support wall are analyzed, and the structure optimization scheme of the fiber coil module is determined. Finally, the optimized fiber coil module is assembled into FOG and verified through random vibration test. The test results show that the vibration performance of the FOG is improved from 0.2183°/h to 0.0262°/h, which effectively improve the vibration performance of high precision FOG.

REFERENCES

[1] G.C. Zhang, The Principles and Technologies of Fiber-Optic Gyroscope, Beijing, National Defense Industry Press, 2008.

[2] Y.J. Wu, The current situation and counter measures of high precision fiber optic gyro, Navigation Positioning &Timing, 2012, vol. 2, no. 4, pp. 53-57.

[3] L.L. Xu, Y.P. Shen, L.J Song, et al, Measurement error analysis of FOG's scale factor, Development and Review of Foreign Navigation Technology in 2019, Navigation and Control, 2020, vol. 19, no. 2, pp. 1-9.

[4] LEFEVRE H C, The fiber- optic gyroscope, London, UK: Artech House, 2014.

[5] Y.J. Wu, X.L. Huang, Influence of mechanical vibration on FOG and reducing measure, Infrared and Laser Engineering, 2008, vol. 37, no. 2, pp. 363-365.

[6] L.H. Xu, Q. Li, M.Q. Sun, et al, Analysis of Shock Absorber Characteristics of Fiber Optic Gyroscope Support for Rocket, Journal of Projectiles, Rockets, Missiles and Guidance, 2019, vol. 39, no. 6, pp. 134-137.

[7] R. Song, X.Y. Chen, Vibration noise processing method for digital closed-loop fiber optic gyroscope, Journal of Southeast University(Natural Science Edition), 2016, vol. 46, no. 3, pp. 489-493.

[8] J.X. Chai, X.L. Wang, D. Wang, et al, A High Precision Modeling Method for Random Vibration Error of Fiber Optic Gyroscope, Aero Weaponry, 2017, vol. 4, pp. 49-54.

[9] H.X. Liu, Y.F. Jiang, N.F. Song, et al, Effect of structural material upon the dynamic performance of FOG, Journal of Beijing University of Aeronautics and Astronautics, 2014, vol. 40, no. 1, pp. 39-43.

[10] B. Zhang, N.F. Song, X. Pan, Light Weight Design Based on Modal Analysis for Triaxial Integrated FOG, Journal of Beijing University of Technology, 2014, vol. 40, no. 11, pp. 1603-1608.

Method for FOG Dead Band Detection on a Stationary Base

N.V. Moiseev
Concern CSRI Elektropribor, JSC
St. Petersburg, Russia
moiseevnv@yandex.ru

R.G. Lyukshonkov
Concern CSRI Elektropribor, JSC
St. Petersburg, Russia
lukroma@yandex.ru

A.A. Untilov
Concern CSRI Elektropribor, JSC
St. Petersburg, Russia
untillov_aa@mail.ru

S.V. Pavlova
Concern CSRI Elektropribor, JSC
St. Petersburg, Russia
bagaevasv@yandex.ru

***Abstract* — A method for determining the dead band of a fiber-optic gyroscope (FOG) on a stationary base using a test input equivalent to an external physical angular rate is considered. The proposed method eliminates the need for a high-precision turntable and significantly reduces the testing time.**

Key words—fiber optic gyroscope, error, dead band, dead zone

I. INTRODUCTION

One of the disadvantages of compensation fiber-optic gyroscopes is the dead zone, which is the range of angular rate values in which the FOG output signal is zero. This complicates the use of FOG-based inertial measurement units in tasks that require high sensitivity to small angular rates. Examples of such applications include gyrotheodolites and navigation modules of spacecraft.

The causes of the appearance of a dead band include electrical interference on the photodetector from the modulating feedback voltage [1,13], nonlinearity of the digital-to-analog converter (DAC) and electro-optical modulator [2,3], the influence of a parasitic interferometer created by back reflections and back scattering [4], as well as polarization errors in the optical circuit.

There are various methods of dead band suppression. They can be conventionally classified into technological and algorithmic ones. Technological methods consist in reducing parasitic coupling between the photodetector and the control signal generation circuit by means of galvanic isolation and shielding, in diminishing the influence of the parasitic interferometer by reducing back reflections, and in reducing the coherence length of the light source. Algorithmic methods consist in stabilizing the parameters of the control system elements by using auxiliary phase modulation and additional feedback [5,12], in introducing pseudo-random noise (dithering) into the control loop [6], and in introducing a tone signal into the control loop that creates an equivalent angular rate acting on the FOG [7]. The methods considered do not always allow one to completely get rid of the dead band, so it is necessary to control the quality of the dead band suppression.

A method for determining the dead band is known, which uses the Earth angular rate [8]. The FOG under study is mounted on a turntable so that its sensitivity axis is in the horizontal plane and perpendicular to the Earth's rotation axis. In this position, the FOG is not affected by the Earth angular rate. Next, using the turntable, the position of the sensitivity axis in the horizontal plane is slowly changed within the required range relative to the zero position. The table rotation angle is converted into the projection of the Earth's rotation velocity onto the FOG sensitivity axis. The dependence of the FOG output signal on the angular rate acting on the FOG is plotted, and the dead band value is visually determined. The smaller the dead band to be measured, the lower the rotation speed of the turntable should be. The disadvantages of this method include the need to use a high-precision turntable; the method is not applicable on board a vehicle. For the FOG triad of the inertial measurement unit, the dead band is determined sequentially for each FOG, which increases the testing time threefold. A three-axis turntable is also required or the equipment for mounting the sensitivity axis of each triad FOG in a horizontal position on a single-axis turntable.

II. METHOD FOR DETERMINING THE DEAD BAND ON A STATIONARY BASE

The method is based on introducing a test signal into the control loop, creating an equivalent physical angular rate of rotation. Test signals are widely used in self-diagnostics (Built-in Self-Test) of sensitive elements of inertial sensors, which determines the operability of the micromechanical structure [9,11]. The articles [10,14,15] describe the use of test signals to determine the dynamic characteristics of the FOG, adjust the control loop regulator, and determine the quality of transient processes.

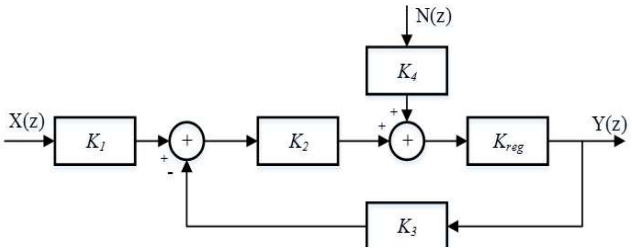


Fig. 1. Simplified structural diagram of the compensation FOG

The linearized structure of the FOG control system with test signal N is shown in Figure 1.

The transfer functions of the control system for physical $G_X(Z)$ and test $G_N(z)$ signals are identical under the condition $K_4=K_1K_2$.

$$G_X(z) = \frac{K_1 K_2 K_{reg}}{1 + K_2 K_{reg} K_3} \qquad (1)$$

$$G_N(z) = \frac{K_4 K_{reg}}{1 + K_2 K_{reg} K_3} \qquad (2)$$

Coefficient K_1 is determined by the parameters of the coil: the length of the fiber L and the diameter D, K_2 – the transfer factors of the interferometer, photosensor, transimpedance amplifier and ADC, K_{reg} – PID controller coefficients, K_3 – DAC and MIOC (multifunctional integrated optical circuit) coefficients.

The method for determining the dead band consists of two stages. At the first stage, the coefficient K_4 is determined. For this purpose, when K_4=1, the test signal N changes in a given range and the value Y of the FOG output signal is recorded. The dependence of the FOG output signal on the test signal is plotted and the coefficient K_4 is determined as the slope of the line. The value of the test signal N_0, at which the FOG output signal is zero, is equivalent to an external zero angular rate. The range of the test signal variation should be such that the point N_0 is within this range and it is possible to determine the coefficient K_4.

At the second stage, taking into account the obtained coefficient K_4, the range of change of the test signal is set within $N_0 \pm \Delta N$ and the value Y of the FOG output signal is recorded. The N_0 magnitude sets the zero equivalent angular rate for the current attitude of the FOG sensitivity axis, and the value ΔN sets the range in which it is necessary to determine the dead band. The dependence of the FOG output signal on the test signal, converted into equivalent angular rate, is plotted by the formula

$$\Omega_{in} = K_4(N - N_0). \qquad (3)$$

Figure 2 shows the dead bands of the FOG obtained on the turntable and using a test signal.

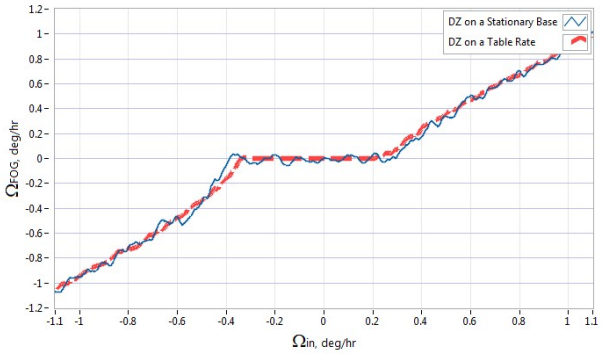


Fig. 2. Dead bands of the FOG obtained during test effect (blue) and on a turntable (red)

It can be seen that the dead bands obtained by the two methods are almost identical.

To test the method, the software was developed that ensures automatic testing of three FOGs simultaneously.

III. Conclusion

A method for determining the dead band of a FOG on a stationary base using a test impact is proposed and experimentally verified. The method allows us to abandon the use of a high-precision turntable, determine the dead band of a FOG in any position of the sensitivity axis relative to the Earth's rotation vector, and significantly reduce the testing time by determining the dead band of three FOGs of an inertial measurement unit simultaneously.

References

[1] Gu, Hong, Huan, Yunpeng, Wang, Ansu, and Luan, Jinwen, Real-time dynamic simulation of angular velocity and suppression of dead zone in IFOG, *Optical Review*, 2015, vol. 22, issue 1, pp. 39-45. DOI:10.1007/s10043-015-0046-0

[2] Pan, Xiong, Wang, Lei, Zhang, Shaobo, and Song, Ning-fang, Effect of digital-to-analog converter glitch on the modulation and demodulation of a digital closed-loop fiber-optic gyroscope, *Optical Fiber Technology,* 2019, vol. 47, pp. 43-50. DOI:10.1016/j.yofte.2018.11.024

[3] Jin, Jing, Tian, Hai-Ting, Pan, Xiong, and Song, Ning-Fang, Electrical crosstalk-coupling measurement and analysis for digital closed loop fiber optic gyro, *Chin. Phys. B,* 2010, vol. 19, no. 3. DOI:10.1088/1674-1056/19/3/030701

[4] Schelkunov, E.V. and Makarov, A.A., Problems of the influence of back reflection on the insensitivity zone of a fiber-optic gyroscope, *Proceedings of Intern. Conf. "Aerospace Instrumentation and Operational Technologies"*, 2023, pp. 160-163.

[5] Kurbatov, A.M., Kurbatov, R.A., and Goryachkin, A.M., Fiber-optic gyroscope accuracy improvement by suppressing the parasitic effects in integrated optic phase modulators, *Gyroscopy and Navigation,* 2019, vol. 10, pp. 256-267. DOI:10.1134/S2075108719040114.

[6] Chen, Chung-Jen, Interferometric fiber optic gyroscope dead band suppression, *Applied Physics Express*, 2008. DOI:10.1143/APEX.1.072501

[7] Chong, Kyoung-Ho, Choi, Woo-Seok, and Chong, Kil-To, Analysis of dead zone sources in a closed-loop fiber optic gyroscope, *Applied Optics,* 2016, vol. 55, no.1, pp. 165-170. DOI:10.1364/AO.55.000165

[8] Egorov, D.A., Olekhnovich, R.O., Untilov, A.A., Aleinik, A.S., Deineka, G.B., and Strigalev, V.E., Study on dead zones of fiber-optic gyros, *Gyroscopy and Navigation*, 2011, vol. 2, pp. 197-207. DOI:10.1134/S2075108711040080

[9] Hantos, G., Flynn, D., and Desmulliez, M.P.Y., Built-in self-test (BIST) methods for MEMS: A review, *Micromachines,* 2020, vol. 12(1):40. DOI:10.3390/mi12010040

[10] Pogorelaya, D.A., Smolovik, M.A., Volkovskiy, S.A., Mikheev, M.A., Aleinik, A.S., and Strigalev V.E., Adjustment of PID controller in fiber-optic gyro feedback loop, *Gyroscopy and Navig.*, 2017, vol. 8, pp. 235-239. DOI:10.1134/S2075108717030075

[11] Deb, N. and Blanton, R.D, Built-in test of CMOS-MEMS accelerometers, *Proc. of International Test Conference (ITC*), 2002, 1075-1084. DOI:10.1109/TEST.2002.1041864.

[12] Kurbatov, A.M., Method for eliminating the dead zone in a fiber-optic gyroscope, Patent RU No. 2441202, 2012.

[13] Zhang, Shaobo, et al. High-performance fully differential photodiode amplifier for miniature fiber-optic gyroscopes, *Optics Express, 2019,* vol. 27, no. 3, pp. 2125-2141. DOI:10.1364/OE.27.002125

[14] Xu, You Li and Na, Zhang, A new dynamic test method of fiber optic gyroscope, *Applied Mechanics and Materials,* 2012,vol. 157-158:245-249. DOI:10.4028/www.scientific.net/AMM.157-158.245

[15] Wang, Yan, and Zhang, Chunxi, Investigation of testing dynamic characteristics of digital closed-loop fiber optic gyroscope, *Journal of Beijing University of Aeronautics and Astronautics*, 2004, vol. 30, no.9.

Research on the Optimization of Double-Layer Magnetic Shielding for Fiber Optic Gyroscope

Zerun Zang
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
zangzerun2022@163.com

Pei Zhang
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
1184526231@qq.com

Shitao Ji
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
844395029@qq.com

Fengyuan Ma
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
mfy98123@163.com

Bin Ren
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
emailchina@163.com

Wei Hong
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
hongwei735@163.com

Yichen Wu
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
18829043344@163.com

Xudong Hu
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
huxd2020@163.com

Yu Liu
Xi'an Aerospace Precision Electromechanical Institute
Xi'an, China
410185073@qq.com

***Abstract*—As the core component of inertial navigation systems, the accuracy improvement of fiber optic gyroscope (FOG) has made magnetic field adaptability a critical factor limiting their performance. Currently, high-precision FOG employ a double-layer magnetic shielding scheme, but the structural design parameters remain unclear. This study combines the magnetic shielding requirements of high-precision FOG with finite element analysis using ANSYS to investigate the influence of interlayer gap, thickness ratio, and contact form on shielding effectiveness (SE). The results show that for the typical FOG structure described in this paper, when the interlayer gap is 1 mm and non-magnetic fasteners are used, the radial and axial SE can reach 102 dB and 74 dB, respectively, reducing the gyro bias error to the order of 10^{-5} °/h. The findings provide parametric guidance for the magnetic shielding of FOG and are of significant importance for enhancing their magnetic field adaptability.**

Keywords—fiber optic gyroscope, fiber coil, magnetic shielding, finite element analysis, bias error

I. INTRODUCTION

The fiber optic gyroscope (FOG) is an all-solid-state inertial instrument with advantages such as high reliability, long lifespan, fast startup, and wide dynamic range, making it a core component of inertial navigation systems. As the demand for accuracy continues to rise, magnetic field adaptability has gradually become a key factor limiting the performance of high-precision FOG. External magnetic fields induce non-reciprocal phase differences in the fiber coil through the Faraday effect, leading to errors in the gyroscope output [1]. Complex application environments in maritime, terrestrial, aerial, and space domains further exacerbate the impact of magnetic fields.

Currently, methods to improve the magnetic field adaptability of FOG mainly include magnetic shielding, magnetic error compensation, and the use of magnetically resistant optical fibers. Magnetic shielding involves installing a shield around the FOG to redirect and disperse external magnetic fields, thereby reducing their influence on the internal FOG [2–3]. Magnetic error compensation involves equipping the FOG with a magnetometer to measure the ambient magnetic field strength and using predefined software algorithms to compensate for magnetic-induced errors [4–5]. The application of magnetically resistant optical fibers relies on the development of new materials to reduce the impact of external magnetic fields on the fiber coil [6].

Given the complexities of software design and the cost and stability issues of new materials, the mainstream approach to improving the magnetic field adaptability of FOG is double-layer magnetic shielding. However, existing research primarily focuses on theoretical aspects, lacking systematic optimization of specific structural parameters for high-precision FOG. This results in practical challenges in balancing SE with engineering constraints such as volume and weight.

This study establishes a magnetic field error model for FOG based on the polarization state changes of light waves in optical fibers and the geometric relationship between FOG and magnetic fields. Subsequently, a three-dimensional model of double-layer magnetic shielding for high-precision FOG is constructed, and finite element analysis using ANSYS is conducted to determine the influence of interlayer gap, thickness ratio, and contact form on SE.

II. MAGNETIC FIELD ERROR MODEL FOR FOG

The fundamental principle of FOG is the Sagnac effect, expressed as:

$$\Phi_S = \frac{4\pi RL}{\lambda c}\Omega \tag{1}$$

where R is the radius of the fiber coil, $L = N \cdot 2\pi R$ is the total length of the fiber, Ω is the angular velocity, λ is the wavelength of light, and c is the speed of light.

By deriving the Jones matrix of polarization components in the FOG optical path multiplied by the Jones vector of incident light [7], the phase difference Φ induced by magnetic fields in the two light waves of the fiber coil is obtained as:

$$\Phi = \sum_{i=1}^{n} \frac{4\xi_i t_i}{\Delta\beta} dz \tag{2}$$

where z is the length of the i-th segment of the fiber in rad/m; ξ_i is the circular birefringence induced by the Faraday effect in rad/m; t_i is the circular birefringence caused by fiber twisting in rad/m; and $\Delta\beta$ is the intrinsic linear birefringence of the fiber.

The influence of magnetic fields on the fiber coil includes radial and axial components. The geometric relationship of the fiber coil under a radial magnetic field is shown in Figure 1, where B_R is the radial magnetic field at an angle θ to the X-axis. Point A on the fiber coil is located at a distance z from the starting point, the total fiber length is L, and the coil radius is R.

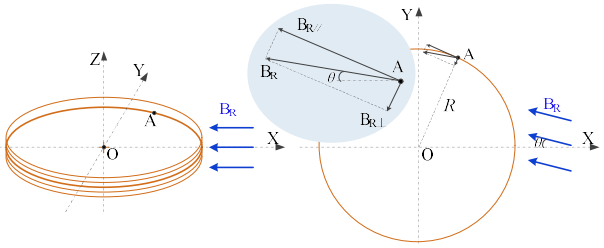


Fig. 1. Fiber coil under a radial magnetic field

The distribution of circular birefringence induced by the Faraday effect along the fiber is:

$$\xi(z) = VB_{R\parallel} = VB_R \sin(z/R + \theta) \tag{3}$$

where V is the Verdet constant.

Substituting Equation (3) into Equations (2) and (1), the angular velocity error is:

$$\Omega_R = \frac{\lambda c V B_R}{\pi R L \Delta\beta} \int_0^L t(z) \sin\left(\frac{z}{R} + \theta\right) dz \tag{4}$$

When $t(z) = t_0 \sin(z/R)$, the maximum error of the fiber coil under a radial magnetic field is:

$$\Omega_{R-MAX} = \frac{\lambda c V B_R t_0}{2\pi R \Delta\beta} \cos\theta \tag{5}$$

The geometric relationship of the fiber coil under an axial magnetic field is shown in Figure 2, where B_A is the axial magnetic field, and the angle between the cross-over sections of the fiber coil and the horizontal plane is α.

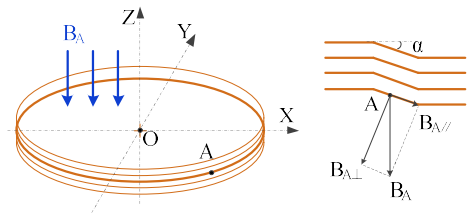


Fig. 2. Fiber coil under an axial magnetic field

The distribution of circular birefringence induced by the Faraday effect along the fiber is:

$$\xi(z) = \begin{cases} VB_A D(p) \sin\alpha, (\textbf{cross-over bending sections}) \\ 0, (\textbf{parallel-wound sections}) \end{cases} \tag{6}$$

where p is the number of layers in the fiber coil; $D(p)$ is a function of the layer number, representing the bending direction of the cross-over sections in each layer (1 for upward bending, -1 for downward bending).

Substituting Equation (6) into Equations (2) and (1), the angular velocity error is:

$$\Omega_A = \frac{\lambda c V B_A}{\pi R L \Delta\beta} \sum_{j=1}^{pq} \left[\int_{L_j}^{L_j+s} t(z) D(p) \sin\alpha dz \right] \tag{7}$$

where q is the number of turns per layer, $p \cdot q = N$, j represents the sequence number of the cross-over bending sections, L_j and $L_j + s$ are the start and end positions of the cross-over sections within the fiber coil, and s is the length of the cross-over sections.

When $t(z) = t_0 D(p)$, the maximum error of the fiber coil under an axial magnetic field is:

$$\Omega_{A-MAX} = \frac{\lambda c V B_A t_0 s}{2\pi^2 R^2 \Delta\beta} \sin\alpha = \frac{\lambda c V B_A t_0 D_{Fiber}}{2\pi^2 R^2 \Delta\beta} \tag{8}$$

Equations (5) and (8) indicate that the error induced in the gyroscope by magnetic fields is proportional to the magnetic field strength. Therefore, reducing the magnetic field strength through shielding is the most direct approach to error suppression.

III. Finite Element Simulation of Double-Layer Magnetic Shielding for FOG

In FOG, the optical and electronic components are typically separated, with external magnetic fields primarily affecting the polarization characteristics of the fiber coil. Thus, magnetic shielding is applied to the fiber coil. The general structure of the fiber coil assembly in FOG is shown in Figure 3.

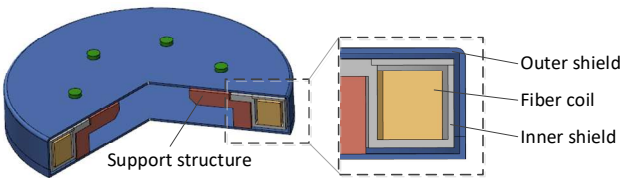


Fig. 3. Structure of the fiber coil assembly

This study takes a fiber coil assembly with an outer diameter of 120 mm and a height of 20 mm as an example. The outer and inner shields are made of high-permeability iron-nickel soft magnetic alloys and are fixed to the support structure using fasteners. Six interlayer gap configurations (0.25 mm, 0.5 mm, 0.75 mm, 1 mm, 1.25 mm, and 1.5 mm) are simulated, with a control model featuring only an inner shield and no outer shield.

The SE is calculated as:

$$SE = 20 \cdot \log_{10} \frac{B_e}{B_i} \tag{9}$$

where B_e is the ambient magnetic field strength, and B_i is the internal magnetic field strength after shielding. SE > 60 dB indicates that B_e is attenuated to less than one-thousandth.

The simulation sets the magnetic field strength to 0.054 mT and uses ANSYS finite element analysis to evaluate SE under radial and axial magnetic fields. The results are shown in Figure 4 and Figure 5.

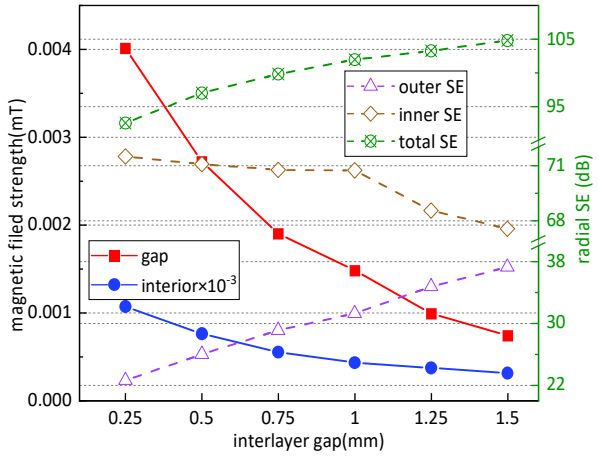


Fig. 4. Simulation results of radial magnetic field

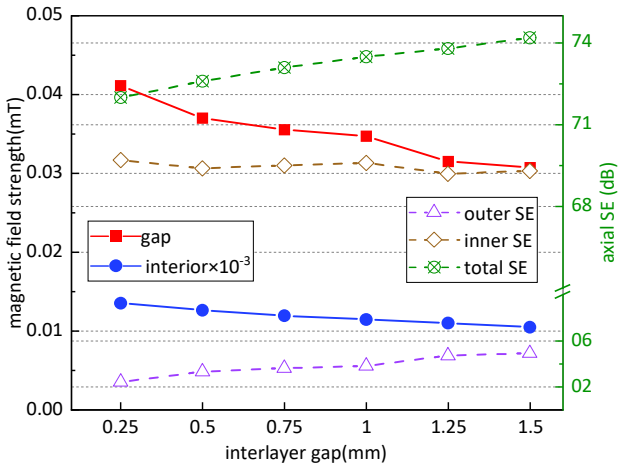


Fig. 5. Simulation results of axial magnetic field

Under a radial magnetic field, the internal magnetic field strength of the fiber coil assembly with only an inner shield is 4.90×10^{-5} mT, while the double-layer shielded assembly achieves 0.03×10^{-5} mT to 0.11×10^{-5} mT. Under an axial magnetic field, the internal magnetic field strength with only an inner shield is 2.78×10^{-5} mT, while the double-layer shielded assembly achieves 1.05×10^{-5}mT to 1.35×10^{-5} mT. These results demonstrate the significant advantage of double-layer shielding over single-layer shielding.

The simulation shows that as the interlayer gap increases, the SE of the outer layer improves, while the inner layer's effectiveness varies irregularly but remains negligible relative to the total SE. This is due to the soft magnetic alloy material of the shields, whose permeability changes with the magnetic field strength. Because the fiber coil assembly is a cylindrical structure with a low height-to-width ratio, the outer shield exhibits lower magnetic reluctance in radial fields, allowing it to redirect more magnetic flux. However, in axial fields, the outer shield's effectiveness is limited by the structural characteristics of the assembly. The interlayer gap disrupts the magnetic path between the two shields, so larger gaps generally improve overall SE. Considering volume constraints, an interlayer gap equal to the shield thickness (1 mm) achieves radial and axial SE of 102 dB and 74 dB, respectively.

When the fasteners are replaced with non-magnetic materials such as aluminum or zinc alloys, the magnetic path between the outer and inner shields is further isolated, enhancing SE.

Based on Equations (5) and (8) in Section 2, with a light wavelength $\lambda = 1550$nm, speed of light $c = 3.0\times10^{8}$m/s, Verdet constant $V = 6\times10^{-4}$rad/m/mT, fiber diameter 150μm, fiber twist amplitude 0.5rad/m, and polarization-maintaining fiber linear birefringence $\Delta\beta = 2200$rad/m, the error under a radial geomagnetic field is reduced from 1.87 °/h to 1.49×10^{-5} °/h, and the error under an axial geomagnetic field is reduced from 1.49×10^{-3} °/h to 3.16×10^{-7} °/h when using a 1 mm interlayer gap and non-magnetic fasteners.

IV. Conclusion

This study, based on the magnetic field error model and finite element simulation of FOG, clarifies the key parameters of double-layer magnetic shielding and their influence. The research shows that with an interlayer gap of 1 mm and non-magnetic fasteners, the radial and axial SE of double-layer shielding can reach 102 dB and 74 dB, respectively, reducing the gyro bias error to the order of 10^{-5} °/h. The results further indicate that SE is limited by the height-to-width ratio of the fiber coil and requires optimization of winding methods or structural design. The interlayer gap and shield thickness ratio must balance volume constraints and magnetic saturation suppression. This study provides parametric guidance for the magnetic shielding of high-precision FOG and is of significant importance for further improving their accuracy. The research primarily focuses on SE in static magnetic fields, and future work should validate its performance in dynamic magnetic environments.

References

[1] Wang W. High-Precision Fiber Optic Gyroscope Technology [M]. National Defense Industry Press, 2021.

[2] Smith J A, Brown K L. Optimization of Double-Layer Magnetic Shields for FOG Applications[J]. Journal of Applied Physics, 2022, 131(8): 084501.

[3] Roberts E, et al. Experimental Validation of Multi-Layer Shielding in Harsh Magnetic Environments[J]. IEEE Magnetics Letters, 2021, 12: 2500505.

[4] Zhou T, et al. Comparative Study of Magnetic Shielding Methods for FOGs[J]. IEEE Transactions on Industrial Electronics, 2022, 69(2): 1789-1798.

[5] Anderson B, et al. Adaptive Magnetic Compensation Algorithms for Fiber Optic Gyroscopes[J]. Sensors, 2021, 21(3): 872.

[6] Li Q, et al. Anti-Magnetic Fiber Materials: Progress and Challenges[J]. Advanced Materials, 2023, 35(7): 2104567.

[7] Liu P. Research on Mechanism Analysis and Suppression Measures of Magnetic Field Error in Fiber Optic Gyroscopes[D]. Harbin Engineering University, 2018.

Research on the Technology for Improving Long Term Performance of High Precision Fiber Optic Gyroscope by Hermetic Packing

Yichen Wu
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA
18829043344@163.com

Pei Zhang
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

Bo Xie
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

Yu Liu
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

Wei Hong
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

Bin Ren
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

Hui Cao
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

Zerun Zang
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

Yu Bai
XI'an Aerospace Precision Electromechanical Institute
China Aerospace Science and Technology Corporation
Xi'an, CHINA

***Abstract*—This article studies the factors affecting the long-term stability of fiber optic gyroscopes（FOGs） and proposes a method to improve the long-term stability of FOG through hermetic packaging technology.**

Based on the hermetic packaging scheme, a 3D model of the hermetic packaging FOG structure was created. By comparing different leakage detection methods, a helium mass spectrometry helium spray detection method for FOG leakage detection was proposed. A FOG prototype was built based on the hermetic packaging scheme, and a hermetic detection platform was established. The hermetic packaging FOG was tested using helium mass spectrometry helium spray detection method, and the leakage rate was less than 5×10^{-10}pa • m^3/s, meeting the hermetic standard. After the leak detection was completed, the FOG was finally packaged and filled with helium gas. And long-term performance comparison tests were conducted on the prototype of the FOG by hermetic packaging. The test results showed that the scale factor change of the FOG after hermetic packaging was only 30ppm within 6 months, which was more than 60% lower than the 80ppm scale factor change of the conventional packing scheme within 6 months. The long-term scale factor change of the FOG was significantly improved, and the hermetic packaging structure is of great significance for improving the long-term stability of the FOG.

Keywords—fiber optic gyroscope (FOG), hermetic packaging, leakage rate detection, long-term stability

I. INTRODUCTION

Fiber optic gyroscope(FOG) is an inertial device that uses fiber coils as sensitive components and Sagnac effect to measure the angular velocity of the optical path. Due to its advantages of all solid state, no moving parts, high reliability, and low power consumption, it is widely used in weapons, aviation, aerospace, navigation and other fields [1-2]. With the rapid development of submarines, weapon carriers, and aerospace fields, FOGs are facing increasingly complex external environments. FOGs are sensitive to environmental factors such as water vapor and pressure changes, and most FOGs use open optical path structures. It is difficult for FOGs to maintain long-term performance stability in these environments. Changes in external environmental pressure can cause changes in the stress field of FOGs, resulting in drift in zero bias and excessive changes in scale factor, affecting FOG performance. In addition, water vapor and salt spray in the external environment can cause microcracks to propagate on the surface of the fiber, degrade the performance of fiber coils and optical devices, and affect the long-term repeatability and reliability of fiber optic gyroscopes[3].Considering its service life and storage environment, it is necessary to carry out hermetic packaging design for FOGs. After the hermetic packaging design is completed, in order to maintain the stability of the internal environment of the hermetic packaging, inert gas needs to be filled inside the FOG to improve its long-term repeatability, reliability, and ability to resist external environmental interference.

This article is based on the hermetic packaging scheme of FOG, and establishes a 3D model of the hermetic packaging structure. By comparing various sealing detection methods, a helium mass spectrometry helium spray detection method for FOG leak detection is proposed. After leak detection is completed, the FOG is finally packed and filled with helium gas inside. A test platform is built to conduct comparative experiments and verify its long-term reliability by comparing the scale factor changes of FOG. This is of great significance for a comprehensive analysis of the long-term stability

improvement of FOG s through hermetic packaging, and provides important theoretical basis for improving the long-term environmental adaptability of FOGs in the future.

II. 1. Factors Affecting the Long-term Stability of Fiber Optic Gyroscopes in the External Environment

Factors affecting the long-term stability of FOGs due to changes in air pressure:

When the air pressure in the environment where the fiber optic gyroscope is located changes, it will cause a change in the internal stress field of the gyroscope, resulting in pressure differences at different positions of the gyroscope. When it acts on the fiber optic, it will cause the gyroscope to exhibit bounce light effect. When the air pressure decreases, the air convection in the gyroscope weakens and the thermal conductivity efficiency decreases, which can lead to changes in the internal temperature field of the gyroscope, and temperature changes can also cause the Shupe effect in the gyroscope. Both of these effects can affect the accuracy of gyroscopes and cause changes[4].

The error caused by pressure changes can be expressed as:

$$\Delta\Omega_P(t) = \frac{n}{DL}\int_0^L \left[A\dot{\xi}_z(z,t) - B\dot{\xi}_r(z,t) \right](L-2z)dz \tag{1}$$

where

n the Refractive index of optical fiber
D the Fiber coil diameter
L the Fiber coil length
$\dot{\xi}_z$ the Time derivative of axial strain within an optical fiber
$\dot{\xi}_r$ the Time derivative of radial strain within an optical fiber

$A = n(1-\frac{1}{2}n^2P_{12})$, $B = \frac{1}{2}n^3(p_{11}+p_{12})$

p_{11}、p_{12} is elasto-optic coefficient

The error caused by temperature changes can be expressed as:

$$\Delta\Omega_T = \frac{n}{LD}\frac{dn}{dT}\int_0^L \frac{dT}{dt}(z,t)(L-2z)dz \tag{2}$$

Therefore, the error caused by changes in air pressure can be expressed as:

$$\Delta\Omega(t) = \Delta\Omega_P(t) + \Delta\Omega_T(t) \tag{3}$$

It can be seen that in order to ensure the long-term performance of fiber optic gyroscopes, it is necessary to reduce the impact of pressure changes on them.

Factors affecting the long-term stability of FOGs due to changes in water vapor:

Fiber coil is the core component of FOG, and its performance directly determines the accuracy of the gyroscope. When the water vapor content in the environment where the fiber coil is located is too high, the active media in the water vapor, such as water, salt, etc will adhere to the surface of the fiber optic, causing microcracks to form on the fiber surface and slowly propagate, resulting in a decrease in the reliability of the fiber[5]. The fatigue factor f of optical fiber can be expressed as:

$$f = \frac{1}{-0.9 + 0.093\lg Z + RT/400} + 1 \tag{4}$$

where
Z the relative humidity
R the gas constant
D the ambient temperature

The service life of optical fiber can be expressed as:

$$T_s = B\sigma^{-f}S_i^{f-2} \tag{5}$$

Where

S_i the static tensile strength
σ the gas constant

It can be seen that the service life of light is only related to the fatigue factor f. In order to improve the long-term stability of FOGs, the service life of fibers is crucial, so it is necessary to eliminate the direct influence of water vapor on FOGs.

III. Design of Hermetic Packaging Structure

Establish a 3D structural model of FOG with hermetic packaging to address the influencing factors of external environment on FOG. The hermetic packing FOG is composed of ASE source, fiber coil module, digital demodulation circuit, shell, nozzle, and insulator in structural form. The ASE source adopts a base integrated structure, providing the installation reference and external installation interface for the entire FOG; The fiber coil module is installed on the central circular installation protrusion of the ASE source base; The digital demodulation circuit is installed on the fiber coil module through a pillar; The shell is equipped with a nozzle and insulator, which are installed in conjunction with the ASE source base. The hermetic packaging structure of the FOG is shown in Figure 1, and the cross-sectional view of the hermetic packaging structure of the fiber optic gyroscope is shown in Figure 2.

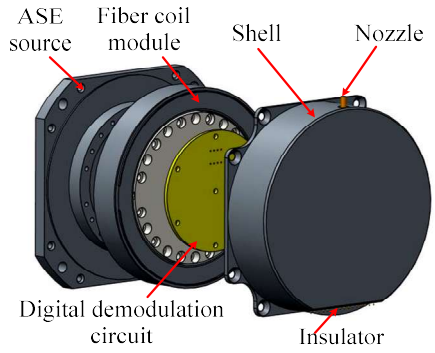


Fig. 1. hermetic packaging structure of the FOG

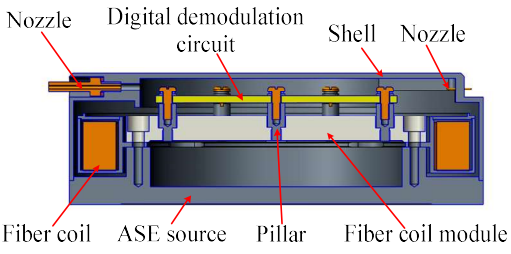


Fig. 2. the cross-sectional view of the hermetic packaging structure

IV. HERMETIC PACKAGING SCHEME AND LEAK DETECTION SCHEME

The hermetic packaging design scheme for the FOG is shown in Figure 3. The nozzle and insulator are fixed to the shell through sealant bonding. The nozzle is used for FOG leakage rate detection, and the insulator forms the external electrical interface of the FOG; The installation screws and overlapping positions of the shell and ASE source base are packed with sealant to complete the hermetic packaging design of the entire FOG.

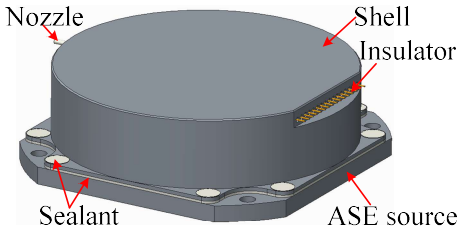


Fig. 3. The hermetic packaging design scheme for the FOG

The commonly used hermetic packaging detection methods include weighing and bubble detection, helium mass spectrometry back pressure detection, and helium mass spectrometry spray helium detection.

For the bubble detection method, due to the external electrical characteristics of FOGs and the presence of various optical devices and components inside, the use of weighing and bubble detection methods may damage the FOG. Therefore, the bubble detection method is not suitable for hermetic packaging detection of FOGs;

For the helium mass spectrometry back pressure detection method, due to the large surface area and complex structure of the FOG, it cannot effectively determine whether there is still helium gas on the external surface after being pressurized in a helium environment and left to stand in the air for a period of time. At this point, the FOG is placed in a leak detection box for leak detection. If helium gas still exists on the outer surface of the FOG, it will affect the leak detection results and make it difficult to determine the hermetic packaging performance of the FOG. Therefore, the helium mass spectrometry back pressure detection method is not suitable for hermetic packaging FOG detection;

For the helium mass spectrometry helium spray detection method, due to its high detection sensitivity and the ability to detect leakage rates for a certain leak point, it can accurately locate the leak hole position and be used for leak detection of complex structural products without causing damage to the FOG. After comparison, this leak detection method is suitable for leak detection of hermetic packing FOGs, so it was adopted for leak detection in subsequent experiments[6].

The specific leak detection process of helium mass spectrometry spray helium detection method is as follows: connect the FOG to the helium mass spectrometer leak detector through an nozzle, evacuate the inside of the FOG, and then use a spray gun to spray helium gas at possible leak locations such as the FOG insulator and sealant. The helium mass spectrometer vacuum leak detection method is used for leak rate testing. If there is a leak hole in the FOG, when helium gas is sprayed on the leak hole, it will be immediately sucked into the vacuum system due to pressure, and then diffuse into the mass spectrometer chamber, and the leak rate of the helium mass spectrometer will change. The principle of hermetic packing detection of FOG using helium mass spectrometry detection spray helium method is shown in Figure 4.

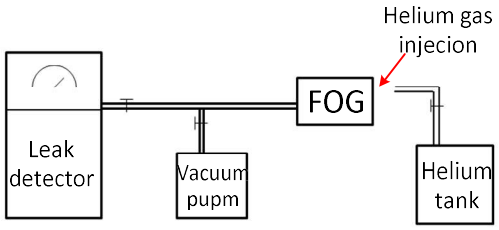


Fig. 4. Helium Mass Spectrometry Detection Principle of Spray Helium Method

The above-mentioned method was used to detect leakage points in the hermetic packaging FOG. After the detection was completed, the gyroscope leakage rate was less than5×10^{-10}pa • m^3/s, meeting the hermetic standard. The leakage rate detection is shown in Figure 5.



Fig. 5. Leakage rate detection

After the leakage rate detection is completed, fix the FOG on the vacuum and inflation device through the base and pressure plate, align the FOG nozzle with the welding device, and then use a hermetic packing cover for overall hermetic packing. Firstly, evacuate the FOG, then fill it with 1 atmosphere of helium, and finally use a welding device to weld and package the gyroscope inflation nozzle in an inert gas environment of 1 atmosphere, thus completing the helium packaging of the FOG, as shown in Figure 6.

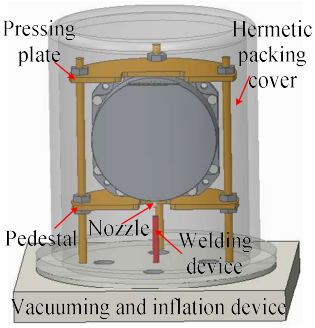


Fig. 6. Final hermetic packaging scheme for FOG

V. HERMETIC PACKAGING FIBER OPTIC GYROSCOPE TEST VERIFICATION

Scale factor is an important indicator of FOG, and its long-term performance is also a major factor affecting the performance and application of FOG. In order to eliminate the influence of external factors such as water vapor and pressure changes on FOGs, a comparative test platform was

designed to monitor the scale factor changes of FOGs in two packaging forms: hermetic packaging and conventional packaging. Two FOGs with different packing forms mounted on the same specification fiber coil module were selected for long-term performance comparison tests. By comparing the long-term scale factor changes, the ability of hermetic packaging gyroscopes to improve long-term performance was reflected.

Two FOGs use ASE light sources, and the conventional packing FOG is packed with rubber strips. The hermetic packing FOG is packed according to the method in this paper, and the leakage rate of the hermetic packing FOG is less than 5×10^{-10}pa • m^3/s. The FOG is filled with 1 atmosphere of helium gas inside.

The scale factor of the hermetic packing FOG has changed by 30ppm (3σ) in 6 months. By comparing the scale factor changes of 80ppm (3σ) in the conventional packing FOG during the same period, the scale factor change has decreased by 50ppm, a decrease of over 60%. Therefore, the hermetic packaging design of the FOG can effectively isolate the influence of environmental factors such as external water vapor and pressure changes, and significantly improve the long-term scale factor stability of the FOG. The test data are shown in Figure 7 and Figure 8.

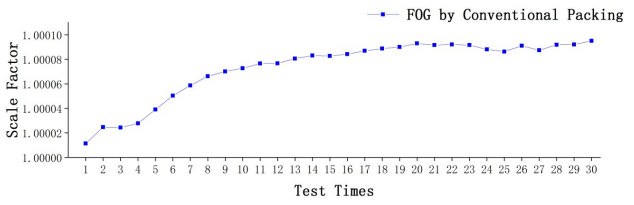


Fig. 7. Six month scale factor variation of the conventional packing FOG

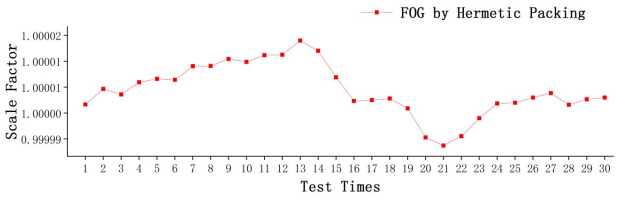


Fig. 8. Six month scale factor variation of the hermetic packing FOG

VI. CONCLUSION

Starting from improving the long-term stability of FOGs, this article optimizes the structure of FOGs and proposes an hermetic packaging FOG structure. The hermetic packaging scheme is analyzed in depth from theoretical analysis, scheme formulation, and experimental verification, and the hermetic packing and leak detection scheme of the hermetic packaging FOG is determined. After the packaging is completed, the leakage rate of the FOG is less than 5×10^{-10}pa • m^3/s. After the hermetic packaging is completed, it is treated with helium gas to effectively solve the environmental factors such as external water vapor and pressure changes. Through comparative experiments, it was found that the scale factor change of the hermetic packing FOG after 6 months was 30ppm, which is more than 60% lower than the scale factor change of 80ppm in the conventional packing form. The hermetic packing structure provides reliable guarantee for the operation of FOGs in complex environments and is of great significance for improving their ability to resist external environmental changes.

REFERENCES

[1] G.C.Zhang, Fiber Optic Gyroscope Principles and Technology, National Defense Industry Publishing, Beijing, 2008.

[2] W.Wang, Interferometric Fiber Optic Gyroscope Technology, China Aerospace Publishing, Beijing, 2010.

[3] W.Wang, Application and De ve lopme nt Te nde ncy of Fibe r Optic Gyroscope in Space，Navigation and Control，2020，vol.19，no.4/5，pp.18-28.

[4] Q.Li，H.C.Yu，D.X.Zhang，Research on Restrain Method of Effect on FOG Zero Bias from Variety of Atmospheric Pressure，Navigation and Control，2016，vol.15，no.1，pp.36-40.

[5] B.X.Xu，Z.Xiong，J.X.Huang，Influence of Hermetic Packaging on the Bias Error and Long-term Reliability of IFOG，Navigation and Control，2021，vol.20，no.1，pp. 97-102.

[6] B.Li，D.F.Hou，X.D.Wang, Study on the influence of the position of precise detection port on the performance of helium mass spectrometer leak detector，Chinese Journal of Scientific Instrument，2020，vol.41, no.11, pp. 58-65.

Analysis of Influence Factors on the Navigation Accuracy of Fiber Optic Gyro Inertial Measurement Unit in Vibration Environments

Xiao Jie Ma
College of Automation Engineering
Nanjing University of Aeronautics and Astronautics
Nanjing, China
ma_xj1989@126.com

Jizhou Lai
College of Automation Engineering
Nanjing University of Aeronautics and Astronautics
Nanjing, China
laijz@nuaa.edu.cn

Ji Xun Huan
Beijing Aerospace Times Optoelectronic Technology Co.Ltd
Beijing, China
jeff331@ns.dept3.buaa.edu.cn

Wen long Feng
Beijing Aerospace Times Optoelectronic Technology Co. Ltd
Beijing, China
fengwenlong@126.com

Wei Hou
Beijing Satellite Manufacturing Factory Co. Ltd
Beijing, China
houwei-bsc@sina.com

Xiang tao Meng
Beijing Aerospace Times Optoelectronic Technology Co.Ltd
Beijing, China
mengxiangtao1982@163.com

***Abstract*—To investigate the internal correlation mechanism between the vibration damping system of a Fiber Optic Gyro (FOG) inertial measurement unit and navigation accuracy, a multi-degree-of-freedom dynamic model of the FOG vibration damping system was established based on multi-body dynamics theory. Theoretical analysis indicates that the performance discrepancy of dampers and the deviation of the system's centroid are the primary causes of linear angular vibration coupling. Furthermore, the nonlinear dynamic interaction between these factors significantly affects the vibration accuracy of the FOG. In this study, using a three-point damper configuration, vibration precision tests were conducted under various vibration magnitudes. Experimental results demonstrate that under different excitation intensities, the impact of damper performance discrepancies and the migration of the system's center of mass on FOG accuracy varies significantly. Additionally, the contributions of these two factors to vibration accuracy differ markedly under varying vibration conditions. This research provides deep IMU into the mechanisms by which the vibration damping system of an FOG inertial measurement device influences accuracy during vibration, offering both theoretical support and experimental validation for the optimal design of vibration damping systems in diverse vibration scenarios. These findings hold significant engineering value for enhancing the environmental adaptability of high-precision inertial navigation systems.**

Keywords—Fiber optic gyroscope inertial measurement unit, Vibration Damping System, Zero Bias.

I. INTRODUCTION

Fiber Optic Gyroscope Inertial Measurement Unit Systems (IMU) constitute the core navigation equipment for spacecraft. The accuracy of these systems directly determines the precision of spacecraft autonomous navigation and attitude control, which is pivotal to the success or failure of space missions. During transportation and launch, the IMU must endure broadband random vibrations spanning from 5 to 2000 Hz, as well as transient shocks exceeding 1800g [1]. In the orbital operation phase, the system is exposed to prolonged low-magnitude micro-vibrations. Ensuring the reliability of FOG-based systems has therefore become a critical technical challenge in aerospace engineering that demands urgent attention [2].

In engineering applications, vibration damping system optimization is used to reduce the impact of external mechanical forces on optical fiber inertial systems. This involves adjusting stiffness and damping parameters to minimize vibration energy transmission[3]. However, improper design of the damping system can lead to line-angle vibration coupling, a major challenge for improving the adaptability of optical fiber inertial units. Discrepancies in the center of mass or asymmetric dynamics in the isolation unit can cause coupling effects under external excitation[4]. These effects disrupt motion signal sensing, causing measurement errors that affect navigation system accuracy and stability.

Currently, both domestic and international academic communities have conducted comprehensive theoretical and engineering investigations into the vibration damping system of fiber optic inertial units [5-7]. Dynamic modeling reveals that aligning the elastic center with the center of mass is crucial for achieving vibration decoupling. Principles for designing low-frequency linear resonance and high-frequency angular resonance have been established. Engineering comparisons confirm that an eight-point support structure outperforms traditional four-point schemes in suppressing angular coupling vibrations. These findings support shock absorber parameter optimization. However, current research focuses on static characteristics, while actual engineering scenarios involve significant dynamic behavior due to varying vibration excitation affecting line-angular coupling.

This study systematically investigates three key aspects: dynamic modeling, parameter impact assessment, and experimental validation. A comprehensive mechanical model of the optical fiber inertial system was developed by integrating theoretical analysis with numerical simulation, elucidating the coupling evolution of linear and angular vibrations under varying excitation conditions[8]. The sensitivity differences of vibration reduction parameters and centroid offsets were evaluated in a wide-field vibration

environment. A shaking table experiment encompassing both high- and low-magnitude spectra validated the model's effectiveness through comparative analysis of experimental data and theoretical predictions.

II. Design of IMU vibration damping system

The fiber optic IMU, a key component for spacecraft navigation and control, includes fiber optic gyroscopes, quartz-based accelerometers, and signal processing circuits. The gyroscopes measure angular velocity based on the Sagnac effect, while the quartz accelerometers rely on the electromechanical conversion properties of piezoelectric crystal materials. These precision inertial sensors are fabricated using advanced optical interference structures and micromachining technologies, but their design inherently limits mechanical strength and adaptability to vibration environments [9].

In the system-level design of fiber optic inertial arrays, the optimization of vibration damping schemes is critically important. The design must systematically incorporate dynamic response characteristics, structural material parameters, and three-dimensional spatial configurations to effectively suppress external excitations via high-performance vibration isolation devices. Furthermore, precise control of vibration transmission during dynamic coupling phenomena, especially in eliminating interactions between axial angles and vibrations/oscillations [10] , ensures the high-precision performance of the navigation and positioning system under complex vibration conditions.

A. Design of vibration damping system

In response to the technical requirements for lightweight and integrated development of optical fiber inertial units, the design of the vibration damping system must concurrently optimize structural compactness and vibration suppression performance. To achieve this, this study proposes an innovative comprehensive solution that integrates a T-shaped vibration reduction unit with a three-point vibration isolation architecture. By symmetrically arranging three T-shaped vibration modules along three axes, a compound vibration isolation system with spatial symmetry is constructed, which significantly enhances its ability to attenuate random vibrations and transient impact loads. Through precise control of the IMU angles and stiffness parameter matching for each vibration unit, the system achieves collaborative optimization of linear vibration transmissibility and angular displacement restraint. The spatial configuration topology of the vibration damping system is illustrated in Fig. 1.

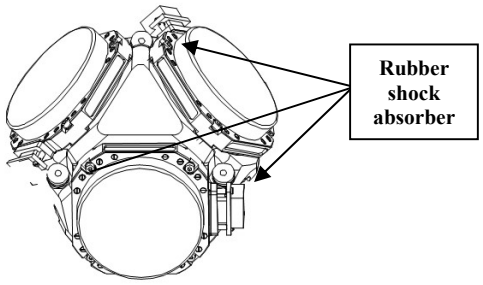


Fig. 1. Schematic diagram of the damper system layout

B. 2.2 Shock absorber design

To effectively control the line-angle vibration coupling in the optical fiber inertial damping system under multi-dimensional excitation, this study developed a compound vibration damping device by integrating a T-shaped rubber shock absorber with a multi-dimensional constraint mechanism. The device features a radial restraint sleeve and an axial rigid stop gasket to prevent function failure of the damping element due to excessive displacement. As shown in Fig.2, this structural integration not only balances vibration reduction efficiency with motion accuracy but also establishes a two-way protection mechanism for stable operation of the optical fiber inertial unit under extreme conditions.

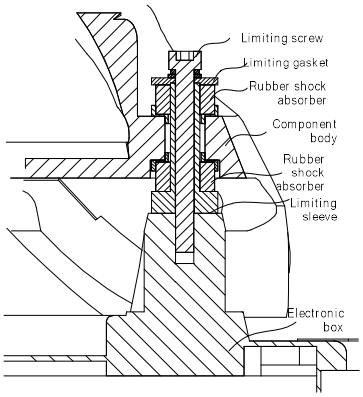


Fig. 2. Schematic diagram of shock absorber composition

III. Dynamic Model for Fiber Optic IMU

To accurately characterize the dynamic response characteristics of an optical fiber IMU in a complex mechanical environment, theoretical modeling research can be conducted based on the 6-degree-of-freedom dynamic equivalent model [11]. In this modeling approach, the IMU is abstracted as a rigid body structure. When the system is subjected to external mechanical excitation, linear vibration responses are induced along the three Cartesian axes and angular vibration modes are simultaneously excited around each coordinate axis. This coupled vibration mechanism effectively elucidates the dynamic correlation between translational displacement and rotational angular displacement during spatial motion, thereby comprehensively characterizing the dynamic behavior of the system in both translational and rotational dimensions.

The vibration damping device can be further simplified to include three orthogonal axial elastic damping elements, which respectively characterize the stiffness coefficient and damping characteristics in the translational motion degree of freedom, while three rotational elastic damping elements are configured to characterize the dynamic parameters in the angular vibration degree of freedom[12]. The simplified structure of the dynamic model is shown in Fig. 2.

Where X , Y , Z respectively represent the linear displacement of the IMU when it is translated along the three coordinate axes of the solving coordinate system θ_X , θ_Y , θ_Z respectively represent the angular displacement movement of the IMU when it is rotated along the three coordinate axes of the solving coordinate system.

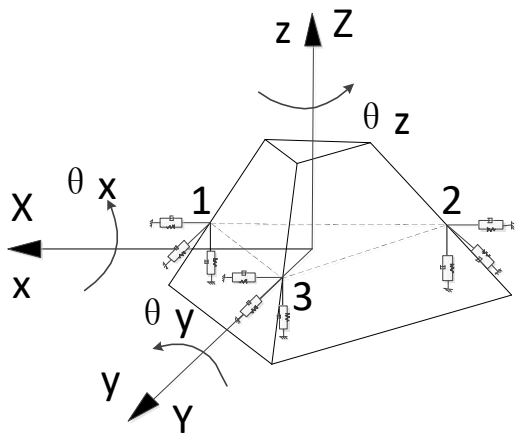


Fig. 3. Physical Modeling of the Vibration Damping System

If the structural stiffness of the IMU is high enough, the motion characteristics of the IMU can be approximately regarded as rigid motion. The 6-DOF equivalent model shown in Fig. 2 can be used to approximately describe the physical model of the inertial fiber array with three-point vibration damping structure. The Cartesian coordinate system oxyz was established at the centroid position of the equivalent model, and the direction of the coordinate axis was shown in Figure 1. x, y and z are defined to represent the linear displacement components of the IMU when it is translated along the three coordinate axes, φ_x , φ_y , φ_z are used to represent the angular displacement parameters of the IMU when it is rotated around the corresponding coordinate axes. Under the basic displacement excitation, when the IMU is in the small displacement vibration condition, based on the multi-body dynamics theory, the space dynamic differential equations of the system in the free vibration state can be approximated as:

$$\begin{Bmatrix} m\ddot{x}+K_x x+R_{xz}\varphi_y-R_{xy}\varphi_z=0 \\ m\ddot{y}+K_y y-R_{yz}\varphi_x+R_{yx}\varphi_z=0 \\ m\ddot{z}+K_z z+R_{zy}\varphi_x-R_{zx}\varphi_y=0 \\ I_{xx}\ddot{\varphi}_x-R_{yz}y+R_{zy}z+K_{xx}\varphi_x+I_{yx}\ddot{\varphi}_y-K_{xy}\varphi_y+I_{zx}\ddot{\varphi}_z-K_{xz}\varphi_z=0 \\ I_{yy}\ddot{\varphi}_y+R_{xz}x-R_{zx}z-K_{yx}\varphi_x+I_{xy}\ddot{\varphi}_x+K_{yy}\varphi_y+I_{zy}\ddot{\varphi}_z-K_{yz}\varphi_z=0 \\ I_{zz}\ddot{\varphi}_z-R_{xy}x+R_{yx}y-K_{zx}\varphi_x+I_{xz}\ddot{\varphi}_x-K_{zy}\varphi_y+I_{zy}\ddot{\varphi}_y+K_{zz}\varphi_z=0 \end{Bmatrix} \quad (1)$$

where:

$K_x=\sum_1^n k_{xi}$, $K_y=\sum_1^n k_{yi}$, $K_z=\sum_1^n k_{zi}$;

$R_{xy}=\sum_1^n k_{xi}l_{yi}$, $R_{yx}=\sum_1^n k_{yi}l_{xi}$, $R_{zx}=\sum_1^n k_{zi}l_{xi}$,

$R_{xz}=\sum_1^n k_{xi}l_{zi}$, $R_{yz}=\sum_1^n k_{yi}l_{zi}$, $R_{zy}=\sum_1^n k_{zi}l_{yi}$;

$K_{xx}=\sum_1^n (k_{yi}l^2{}_{yi}+k_{zi}l^2{}_{yi})$, $K_{yy}=\sum_1^n (k_{xi}l^2{}_{zi}+k_{zi}l^2{}_{xi})$,

$K_{zz}=\sum_1^n (k_{xi}l^2{}_{yi}+k_{yi}l^2{}_{xi})$; z $K_{xy}=K_{yx}=\sum_1^n k_{zi}l_{xi}l_{yi}$,

$K_{yz}=K_{zy}=\sum_1^n k_{xi}l_{yi}l_{zi}$, $K_{zx}=K_{xz}=\sum_1^n k_{yi}l_{xi}l_{zi}$;

$I_{xx}=\int (y^2+z^2)dm$, $I_{yy}=\int (x^2+z^2)dm$,

$I_{zz}=\int (x^2+y^2)dm$, $I_{xy}=I_{yx}=\int yzdm$,

$I_{yz}=I_{zy}=\int xzdm$, $I_{xz}=I_{zx}=\int xydm$;

When k_{xi} , k_{yi} , k_{zi} is the stiffness of the first shock absorber along the axis of x , y , z . n is the number of shock absorbers in the system, and l_{xi} , l_{yi} , l_{zi} is the coordinate value at the IMU Calculation point of the first shock absorber.

1) The center of mass of the component does not coincide with the elastic geometric center of the shock absorber

When a center of mass and damper elastic geometric center is not coincidence, the coordinates of the center of mass in the center of the elastic coordinates (u, v, w), if the stiffness of the damper with different features, at the same time each shock absorber in three axes stiffness parameters remain balanced, set three point within adjacent shock absorber damping layout spacing of 2, a single shock absorber along arbitrary axial stiffness coefficient is k_1 , k_2 ,, under this condition, The space coordinates of the three dampers established based on the centroid coordinate system can be expressed as(l_1 , l_2 , l_3).

where:

$$l_1=(u+l,v-\frac{\sqrt{3}}{3}l,w)\ l_2=(u-l,v-\frac{\sqrt{3}}{3}l,w)\ ,l_3=(u,v+\frac{2\sqrt{3}}{3}l,w)$$

When:

$$K_{xi}=K_{yi}=K_{zi}=k_i \quad (2)$$

$$R_{xy}=R_{zy}=-\frac{\sqrt{3}}{3}k_1 l-\frac{\sqrt{3}}{3}k_2 l+\frac{2\sqrt{3}}{3}k_3 l \quad (3)$$

$$R_{yx}=R_{zx}=k_1 l-k_2 l \quad (4)$$

$$R_{xz}=R_{yz}=0 \quad (5)$$

Substituting the above Eqs(2)~Eqs(5) into the Eqs(1),we obtain

$$\begin{Bmatrix} m\ddot{x}+3kx+3kw\phi_y-3kv\phi_z=0 \\ m\ddot{y}+3ky-3kw\phi_x+3ku\phi_z=0 \\ m\ddot{z}+3kz-3kv\phi_x-3ku\phi_y=0 \end{Bmatrix} \quad (6)$$

As shown in Eqs(6), the existence of eccentricity effect makes the system produce linear vibration in a certain axis, which will inevitably induce angular displacement in the other two orthogonal axes. This dynamic coupling effect results from the spatial offset between the center of mass and the geometric center, which leads to the cross transfer of vibration energy between the degrees of freedom in the Cartesian coordinate system.

2) The stiffness of the shock absorber is inconsistent

When the center of mass of the system is completely aligned with the elastic center of the vibration isolation device but the stiffness of each device is different, it is assumed that each vibration isolation unit has the same stiffness characteristics in the three axis directions, and the spacing between adjacent devices in the three-point vibration isolation system is set to 2 units of length. The stiffness parameters of each shock absorber in any coordinate axis are in sequence k_1 , k_2 , k_3 In the centroid coordinate system, the space coordinates of the three vibration isolation devices can be expressed as (l_1 , l_2 , l_3).

By taking the above parameters into the first three equalities of Eqs(1), it can be rewritten as:

$$\left\{\begin{array}{l} m\ddot{x}+(k_1+k_2+k_3)x-(-\frac{\sqrt{3}}{3}k_1l-\frac{\sqrt{3}}{3}k_2l+\frac{2\sqrt{3}}{3}k_3l)\phi_z=0 \\ m\ddot{y}+(k_1+k_2+k_3)y+(k_1l-k_2l)\phi_z=0 \\ m\ddot{z}+(k_1+k_2+k_3)z+(-\frac{\sqrt{3}}{3}k_1l-\frac{\sqrt{3}}{3}k_2l+\frac{2\sqrt{3}}{3}k_3l)\phi_x-(k_1l-k_2l)\phi_y=0 \end{array}\right\} \quad (7)$$

Based on formula (7) shows that the damper stiffness differences will lead to the dynamic coupling phenomenon, when the x-axis and the y-axis line vibration inevitably induce the z-axis angular motion, and when the z-axis line vibration occurs, will trigger a x-axis and y-axis Angle motion at the same time, and the coupling effect of multi degree of freedom system is mainly originated from asymmetric characteristics of stiffness matrix.

To summarize, the line-angle vibration coupling phenomenon in the optical fiber inertial system is primarily attributed to two core factors: center of mass displacement and damper stiffness heterogeneity. However, there are fundamental differences in the mechanisms and degrees of influence between these two factors. The displacement of the center of mass generates an additional inertial moment by introducing an eccentric distance, thereby promoting the interactive coupling between linear and angular vibrations. On the other hand, the stiffness heterogeneity of the dampers induces symmetry breaking in the vibration transmission path by reconfiguring the system dynamics, thus enhancing the coupling between linear and angular vibrations.

It is important to note that due to the inherent viscoelastic characteristics and geometric nonlinear properties of shock absorber materials, their dynamic mechanical responses exhibit significant nonlinear behaviors under random vibrations of varying intensities. When the system experiences low-magnitude vibrations, the dampers predominantly operate within the linear response range, where their stiffness and damping parameters remain relatively constant. Conversely, under high-intensity vibration excitation, molecular chain slippage within the material and nonlinear contact interface effects lead to stiffness degradation and time-varying damping phenomena. In this case, the material constitutive behavior becomes more pronounced, and the stiffness heterogeneity tends to diminish because all dampers are composed of the same material. This nonlinear evolution mechanism results in a dynamic adjustment of the contributions from center of mass displacement and stiffness heterogeneity to the vibration coupling, depending on the excitation intensity.

IV. Experimental verification

To verify the specific effects of the inconsistent stiffness of the shock absorber and the mismatch between the center of mass and the elastic center in the vibration damping system under different vibration magnitudes, this study designed and performed multiple sets of precision vibration tests.

A. Test Methods

A high-precision shaking table was used to construct a full-band mechanical environment simulation system. In accordance with the spacecraft dynamic environment test specification (GJB 1027A-2004), the vibration excitation was divided into three characteristic frequency bands: low frequency (5–50 Hz), mid-frequency (50–500 Hz), and high frequency (500–2000 Hz), with the total root mean square (RMS) acceleration serving as the evaluation index. Two typical operating conditions were defined: high magnitude (RMS = 8.10 g) and low magnitude (RMS = 1.73 g). The sinusoidal frequency sweep, random vibration, and composite excitation methods were utilized to simulate the impact vibration during the launch phase and the vibration environment during the orbit operation phase. Non-uniform stiffness conditions were constructed by adjusting the stiffness parameters of the vibration damping device, with a stiffness difference range of 4%–30% set based on engineering practice. The mathematical model of centroid coordinates was applied to achieve precise control of the centroid offset by attaching a mass block to key components of the optical fiber inertial set, thereby constructing experimental conditions for centroid offsets ranging from 0.2 mm to 10 mm.

B. Experimental results

The primary structure of the optical fiber inertial set remains unchanged, and stiffness configurations are adjusted by modifying the shock absorbers' height and diameter. Random vibration excitation was applied in three frequency bands—low (5–50 Hz), mid (50–500 Hz), and high (500–2000 Hz)—with an RMS acceleration of 8.10g for 1080 seconds. Test results show that stiffness configurations significantly affect system dynamics. Detailed data are in Table I.

TABLE I. Results of test condition 1

Type of operating conditions		Vibration magnitude	Zero change in FOG vibration (°/h)		
Stiffness difference	*Eccentricity (mm)*		*X*	*Y*	*Z*
4	0.2	Low magnitude	0.0221	0.0306	0.0228
		High magnitude	0.0432	0.0347	0.0461
10	0.2	Low magnitude	0.0659	0.08	0.0781
		High magnitude	0.2909	0.2915	0.2411
20	0.2	Low magnitude	0.6218	0.6797	0.6567
		High magnitude	0.7549	0.7485	0.7745
30	0.2	Low magnitude	1.1964	1.2252	1.1475
		High magnitude	1.538	1.3834	1.4257

In Test Condition 2, the optical fiber inertial unit uses a specific shock absorber configuration. The position deviation between the center of mass and the elastic center is adjusted by attaching a counterweight. The vibration test system applies two excitation conditions—high-magnitude (RMS = 8.10 g) and low-magnitude (RMS = 1.73 g)—across three frequency bands: low (5–50 Hz), mid (50–500 Hz), and high (500–2000 Hz). The random vibration test lasts for 1080 seconds. Table II shows detailed test data and system response characteristics under various vibration parameters.

The experimental data demonstrate that optimizing the static stiffness parameter differences of the dampers and precisely controlling the centroid offset of the Inertial Measurement Unit (IMU) significantly enhance the measurement accuracy of Fiber Optic Gyroscopes (FOG) in vibration environments. Through an in-depth analysis of two sets of control experiments, the specific outcomes are as follows:

In Experiment 1, the influence of damper stiffness dispersion on gyro measurement accuracy was systematically

investigated by fixing the eccentricity of the Inertial Measurement Unit (IMU). The experimental data indicate that when the stiffness consistency index is improved from 20% to 10%, the improvement rate of zero-offset accuracy in gyroscope vibration reaches 89% under low-intensity vibration conditions. In a high-intensity vibration environment, this parameter improves by 61%. Further comparative analysis reveals that when stiffness dispersion is optimized from 30% to 10%, the zero-bias accuracy increases by 95 percentage points under low-magnitude vibration conditions and achieves a 78% gain in high-magnitude vibration conditions. These results clearly demonstrate that enhancing the consistency of damper stiffness significantly improves the stability of Fiber Optic Gyroscopes (FOGs) in low-intensity vibration scenarios, with greater performance gains observed for higher initial stiffness dispersion levels.

TABLE II. RESULTS OF TEST CONDITION 2

Type of operating conditions		**Vibration magnitude**	**Zero change in FOG vibration (°/h)**		
Stiffness difference	***Eccentricity (mm)***		***X***	***Y***	***Z***
4	0.2	Low magnitude	0.0221	0.0306	0.0228
		High magnitude	0.0432	0.0347	0.0461
10	2	Low magnitude	0.1049	0.0806	0.0737
		High magnitude	0.1514	0.1231	0.1152
20	5	Low magnitude	0.3254	0.2254	0.2115
		High magnitude	0.5408	0.4407	0.6132
30	10	Low magnitude	0.8147	0.7075	0.6706
		High magnitude	1.8410	1.9439	1.7672

In Experiment 2, while maintaining constant damper stiffness, the study focuses on analyzing the effect of IMU centroid offset on gyroscope measurement accuracy. The data show that reducing the centroid offset from 5 mm to 2 mm improves zero-position accuracy by 67% under low-intensity vibration and by 81% under high-intensity vibration. When the offset is adjusted from 10 mm to 5 mm, the accuracy improvement rates are 60% and 78% in low- and high-intensity vibration environments, respectively. These experimental findings confirm that effectively controlling the IMU centroid offset significantly enhances gyroscope vibration accuracy under high dynamic load conditions, thereby validating the critical role of the centroid offset parameter in determining measurement accuracy during severe vibration.

V. CONCLUSION

This study focuses on the investigation of the vibration damping system for fiber optic IMU. The dynamic simulation model is established to carry out theoretical deduction, A dynamic simulation model is developed to conduct theoretical derivations, and vibration environment simulation experiments are performed to validate the effectiveness. The mechanisms by which damper performance discrepancies and system centroid offsets affect navigation accuracy are systematically analyzed. Theoretical derivations and experimental data confirm that the influence of centroid offset and stiffness heterogeneity on the coupling effect of linear-angle vibrations exhibits dynamic evolution characteristics under varying random vibration intensities. Under low-intensity vibration conditions, enhancing the uniformity of damper stiffness distribution effectively improves the vibration mitigation accuracy of FOG, with the degree of initial stiffness deviation positively correlated with the optimization effect. Under high-intensity vibration conditions, reducing the assembly eccentricity distance of the fiber optic inertial components yields a more significant improvement in navigation accuracy.

This research establishes a robust theoretical foundation for optimizing the vibration damping system of IMU and provides reliable data support under dynamic vibration conditions. In engineering applications, the key parameter configurations of the vibration damping system can be adjusted according to different vibration intensity scenarios to effectively enhance its environmental adaptability its and navigation accuracy.

REFERENCES

[1] Li X ,Xie Y ,Tian Y , et al. A Study on the Design and Implementation Technologies of EVA at the China Space Station [J]. Aerospace, 2024, 11 (4):22-24.

[2] Ayala L F, Carsten W, Vitali B. Analysis of Space Launch Vehicle Failures and Post-Mission Disposal Statistics [J]. Aerotecnica Missili & Spazio, 2022, 101(3): 243-256.

[3] YAO Jianjun. Contrast of different vibration isolation patterns used in strap-down inertial navigation system [J]. Structure & Environment Engineering, 2009, 36 (2) : 19-27.

[4] J.M. Leonard, F.G. Nievinski, and G.H. Born, “Gravity Error Compensation Using Second-Order Gauss-Markov Processes”, Journal of Spacecraft and Rockets, 2013, vol. 50, pp. 217-229.

[5] Banerjee K, Dan B, Majumdar K, et al. An improved dither-stripping scheme for strapdown ring laser gyroscopes[C]//TENCON 2004 IEEE Region 10 Conference. Thailand, November 24-24, 2004, 1: 689-692.

[6] Ren L, Du J, Wang M. Error analysis and compensation of size effect in IMU with IMU rotation[J]. Acta Aeronautica Et Astronautica Sinica, 2013, 34(6): 1424-1435.

[7] KLIMkOVICH B V. SIMU calibration in inertial mode. Combination of velocity and scalar methods[J]. Gyroscopy& Navigation, 2015, 6(1): 25-32.

[8] WANG Ting, YU Pei, LI Jing. Error suppression technique of platform inertial navigation system under condition of vibration [J]. Journal of Vibration and Shock, 2019, 38 (15) : 258-263.

[9] C. Hirt, B. Bürki, A. Somieski, G. Seeber, “Modern Determination of Vertical Deflections Using Digital Zenith Camera”, Journal of Surveying Engineering-Asce, 2010, vol. 136, pp.1-12.

[10] YAN Hongsong, LI Xin, et al. Relationship between parameters of vibration damping system and coupling angular motion of SIMU [J]. Journal of SIMU Chinese Inertial Technology, 2021, 29 (5) : 576-582.

[11] Wang Guang-yuan, Zheng Gang-tie, Han Chao. Dynamic modeling method of metallic rubbers in frequency domain [J]. Journal of Astronautics, 2008, 29(2):499-504.

[12] SUN Yi, LI Jiawei, JIANG Xiaoyan, et al. Research on three dimensional stiffness of wedge-shaped rubber shock absorber based on orthogonal experiment [J]. China Rubber Industry, 2022, 69 (4) : 268-273.

Study of the Characteristics of a Miniature Three-Axis Integrated Fiber-Optic Gyroscope Based on an Application-Specific Integrated Circuit (ASIC)

Fan Liu
OOO «CHONGQING TIANJI Inertial Technology»
Beijing, China
12121215121@163.com

Fanyu Sun
OOO «CHONGQING TIANJI Inertial Technology»
Beijing, China
12121215121@163.com

Yu Din
OOO «CHONGQING TIANJI Inertial Technology»
Beijing, China
12121215121@163.com

Binze Lin
OOO «CHONGQING TIANJI Inertial Technology»
Beijing, China
12121215121@163.com

Yuelan Li
OOO «CHONGQING TIANJI Inertial Technology»
Beijing, China
12121215121@163.com

Zihan Wang
Bauman Moscow State Technical University
Moscow, Russia
Szbrzhw@163.com

The paper presents the development of a specialized integrated circuit (ASIC, Application-Specific Integrated Circuit) for a miniature three-axis fiber-optic gyroscopic system designed for use in low-cost inertial systems. The ASIC implements key signal processing functions, including level conversion, sample-and-hold, amplification, and differential signal transformation. This solution is integrated with System-in-Package (SiP) technology, which combines digital processing, signal conversion, and a power management module into a single chip. The use of this circuit reduces size and power consumption while optimizing zero-bias stability, scale factor repeatability, and adaptability to external conditions compared to existing alternatives.

Keywords: fiber-optic gyroscope (FOG); Application-Specific Integrated Circuit (ASIC); System-in-Package (SiP); inertial navigation; integration.

I. INTRODUCTION

Fiber-optic gyroscopes (FOGs) offer advantages such as no mechanical moving parts, stable construction, high resistance to shock and vibration, and long service life. With technological advancements and expanding applications, FOGs are actively evolving toward integration and miniaturization.

The use of multiple discrete components in traditional FOG electronic systems increases equipment size, power consumption, and susceptibility to signal chain interference. These structural limitations significantly hinder practical applications in space-constrained and low-power environments. To overcome these barriers, integrated solutions based on ASICs and SiP technology are being adopted. Circuit integration not only achieves breakthroughs in miniaturization and energy efficiency but also fundamentally improves signal processing performance while optimizing production costs.

This technological transformation redefines FOG product formats, enabling new civilian applications and strengthening technological superiority in professional inertial navigation systems. This study implements deep circuit system integration based on ASIC and SiP technologies. Key achievements include the development of two critical integrated circuits:

- A hybrid analog front-end ASIC integrating key signal processing functions, voltage level conversion, sample-and-hold circuits, amplification stages, and differential transformation.
- A multi-chip SiP integration combining FPGA, DAC/ADC, and a power management module.

As a result, circuit dimensions are reduced to 60% of the previous design, and power consumption is cut by over 50% while maintaining excellent zero-bias stability and scale factor characteristics.

This work presents the development of a three-axis integrated fiber-optic gyroscope, the TG-40, using a specialized ASIC, along with an analysis of its performance. The gyroscope not only offers compact size and low power consumption but also demonstrates outstanding zero-bias stability, scale factor accuracy, and environmental resilience, supporting the creation of a new generation of miniaturized, integrated, energy-efficient, and high-performance FOGs.

II. DEVELOPMENT OF SPECIALIZED ASICS

2.1 Development of the Hybrid Analog Front-End ASIC

As shown in Figure 1, the developed hybrid analog front-end ASIC integrates a three-stage operational amplifier and a dual analog switch. The device measures 4 mm × 4 mm × 1.5 mm. Using ASIC technology, over ten discrete components—including a three-stage op-amp, a dual-channel analog switch, and a precision resistor network—are integrated into a single chip. This replaces traditional discrete solutions such as separate op-amp ICs, negative voltage supply modules, filtering circuits, and level conversion circuits.

Crystal-level integration reduces component count from over ten discrete elements to a single packaged solution. The ASIC directly processes weak signals in the range of -0.3 V to 5.3 V using a single-polarity rail-to-rail supply, achieving a differential output signal-to-noise ratio of 84 dB. Circuit area and power consumption are reduced by 70% and 50%, respectively.

As shown in Figure 2, the interface diagram of the hybrid analog front-end ASIC is presented. This IC provides a single-chip solution for analog signal conversion in gyroscopic systems, integrating a three-stage amplification structure.

First Stage: Processes input signals with negative components (-0.3 V to 5.3 V), converting them to a standard unipolar range (0.1 V to 4.9 V). This eliminates the need for traditional negative voltage generation circuits (e.g., charge pumps or DC-DC modules), simplifying power design, reducing energy consumption, and improving signal quality.

Second Stage (Buffer Driver): Features high input and low output impedance, ensuring stable signal transmission without significant attenuation or distortion.

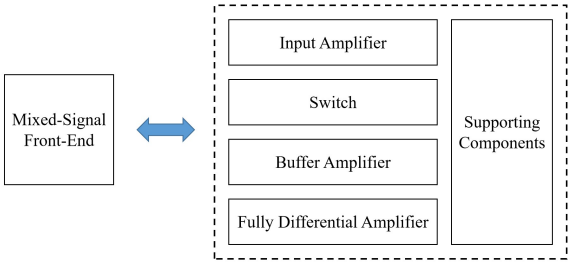


Fig. 1. Block diagram of the hybrid analog front-end ASIC.

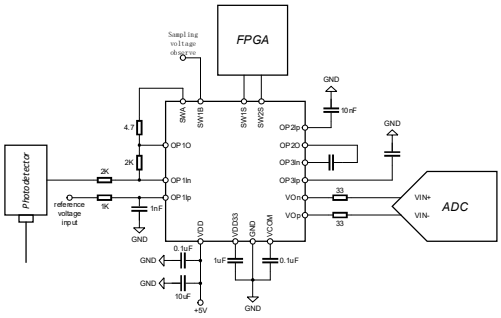


Fig. 2. Interface diagram of the hybrid analog front-end ASIC.

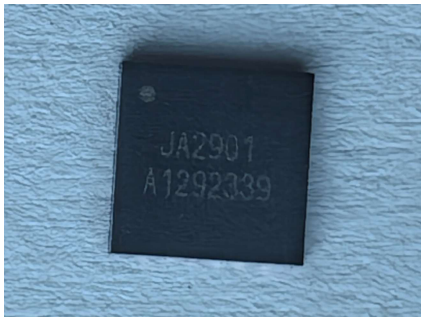


Fig. 3. Physical appearance of the ASIC.

Third Stage (Differential Converter): Transforms unipolar signals into differential signals, enhancing system noise immunity and creating an ideal interface for subsequent analog-to-digital conversion.

2.2 Development of the Hybrid Signal Processing IC (Section 2.2)

In the digital signal path of the fiber-optic gyroscope, the closed-loop digital control module plays a key role. This module performs real-time digitization and processing of the raw optical signal from the gyroscope output, extracts angular velocity data using specialized digital demodulation algorithms, and enables precise closed-loop control via a high-speed feedback circuit.

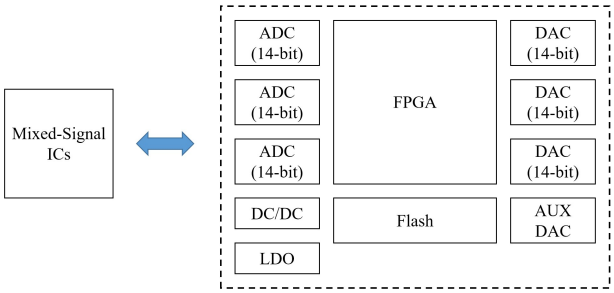


Fig. 4. Hybrid signal processing IC.

To enhance system data processing capabilities while reducing size and power consumption, this work presents a mixed-signal IC for FOG analog signal processing. As depicted in Figure 5, the IC integrates multiple previously discrete modules into a single BGA196-packaged die (12×12×1.5 mm). It includes:

- Three high-precision ADCs (analog-to-digital converters).
- Three parallel and three serial DACs (digital-to-analog converters).
- An FPGA (field-programmable gate array).
- A power management system.

This integration reduces external components and PCB layers while improving reliability and manufacturability, meeting the demands of high-performance inertial navigation systems.

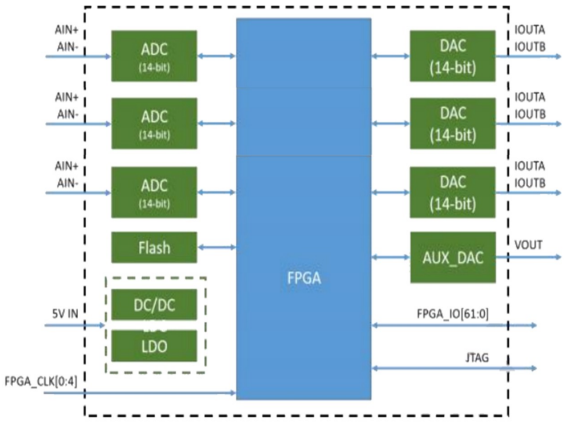


Fig. 5. Interface diagram of the hybrid signal processing IC.



Fig. 6. Physical appearance of the hybrid signal processing IC.

The key functions of this chip include three analog-to-digital converters (ADCs) for real-time collection and digitization of gyroscope signals, three parallel digital-to-analog converters (DACs) providing high-speed feedback control, as well as three serial DACs for auxiliary adjustment signals. The system's central processor utilizes a high-performance FPGA (Field-Programmable Gate Array), which supports flexible digital demodulation algorithms, closed-loop control, and temperature compensation. The power management module delivers stable 3.3V and 1.0V voltages to ensure reliable operation of all components, while SPI Flash memory stores system configurations and calibration data with capability for online updates and adjustments.

Through the integration of FPGA, DACs, ADCs and power management system into a single chip, this solution not only significantly reduces electronic circuit volume but also minimizes signal transmission delays while simultaneously lowering power consumption. Compared to conventional solutions, this integration has enabled a 50% reduction in power consumption and decreased processing delays to microsecond levels, ensuring high precision and stable operation of the fiber-optic gyroscope even under demanding environmental conditions.

2.3 Synthesis of the Two ICs

These hybrid signal processor chip designs, through carefully planned integration, successfully combine the entire fiber-optic gyroscope signal processing chain including acquisition, amplification, conditioning, digital processing and feedback control. This approach not only substantially reduces physical dimensions and power requirements but also enhances system performance and stability. Unlike traditional solutions, this architecture reduces peripheral circuit complexity and printed circuit board layers by minimizing discrete component count, allowing for significant reductions in gyroscope size and energy consumption while maintaining high reliability and performance. The high level of integration between analog input circuits and digital signal processors optimizes signal transmission paths, minimizes noise levels, and ensures excellent interference immunity along with real-time data processing capability even in challenging inertial navigation environments.

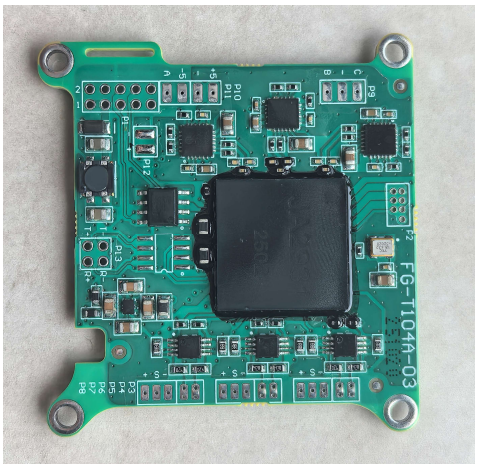


Fig. 7.

All functional modules of the chip interact efficiently through an optimized interconnect architecture, where an FPGA (Field-Programmable Gate Array) serves as the central processor, working in conjunction with high-speed data converters to meet all real-time signal processing requirements. The built-in power management system provides stable voltage to all modules, while SPI Flash memory stores configuration data. This integrated design significantly reduces system dimensions and power consumption while simultaneously enhancing reliability and coherence, making it an ideal solution for applications with stringent space constraints and energy efficiency requirements.

III. Design and Simulation of the Fiber-Optic Gyroscope Based on Specialized Integrated Circuits

3.1 Development of the Three-Axis Fiber-Optic Gyroscope TG-40 Using Specialized Integrated Circuits

As shown in Figure 8, the device schematic includes: a light source with driver and thermal stabilization that maintains constant pump current and temperature control; the optical signal is distributed across three axes via a 1×3 splitter, with each axis feeding into a Y-waveguide through a 1×2 splitter for phase modulation and separation into counter-propagating light beams. After passing through sensing coils, the optical beams return to the Y-waveguide for interference, are converted to analog signals by photodetectors, then digitized by ADCs in the ASIC and processed by the FPGA (Field-Programmable Gate Array). The results are then fed back through DACs to the Y-waveguide for phase modulation, forming a closed-loop control system. This design ensures high measurement accuracy with compact dimensions and resistance to external disturbances.

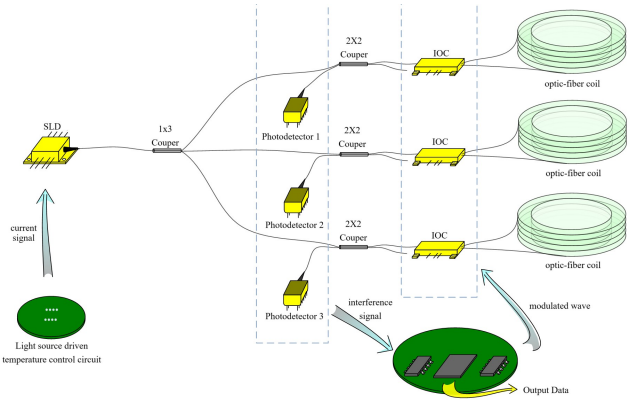


Fig. 8. Structure of the TG-40 Gyroscope with Integrated ASIC Control

The structure of the fiber-optic gyroscope TG-40 with integrated ASIC control, as shown in Figure 9, comprises three main components: mechanical assembly, optical scheme, and electronic circuits. The mechanical part consists of a base and magnetic shield, while the optical scheme integrates an SLD light source, optical splitters, Y-waveguide, fiber coil, and photodetector. The electronic section is implemented through a main control board, detection board, and driver board, with the specially developed ASIC highlighted in red on the schematic, emphasizing its crucial role in ensuring the device's compactness and functionality. This integration of all components into a unified system enables high measurement accuracy while maintaining minimal dimensions and power consumption.

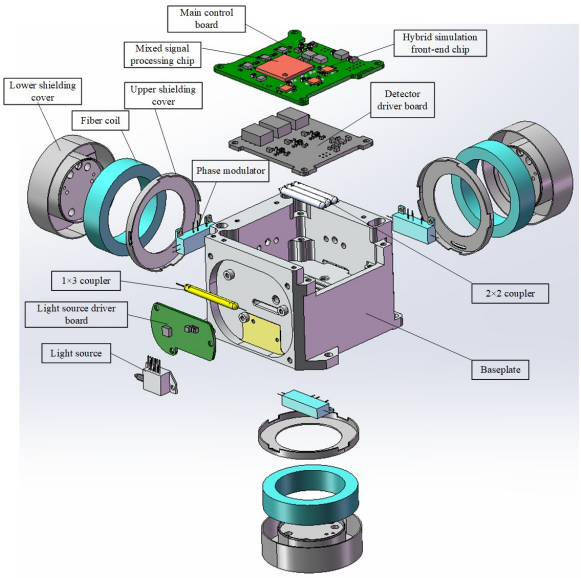


Fig. 9. Exploded View of the TG-40 Fiber-Optic Gyroscope with Integrated Circuit Control

The TG-40 fiber-optic gyroscope based on the specialized ASIC control chip (shown in Figure 10) features compact dimensions of 71.5×70.2×45 mm, with its printed circuit board measuring only 51.5×49.7 mm - 30% smaller than conventional solutions. This achieves significant reduction in mass-dimensional characteristics without compromising functionality or measurement accuracy, which is particularly important for modern inertial navigation systems with stringent size requirements for components.

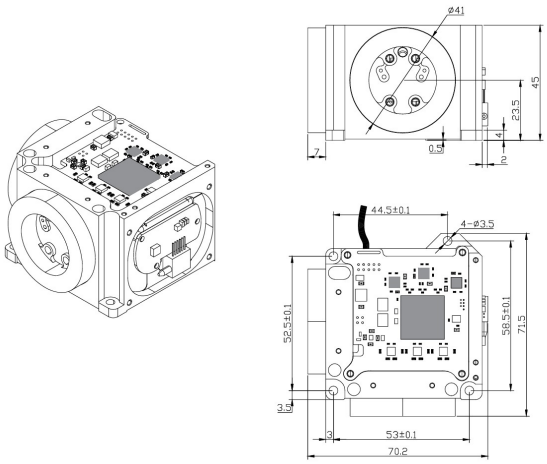


Fig. 10. Dimensional Drawing of the TG-40 Fiber-Optic Gyroscope with Specialized Integrated Circuit

3.2 Vibration Characteristics Simulation

Vibration simulation was conducted to precisely determine the operational characteristics of the TG-40 gyroscope with specialized IC control under various operating conditions. Initial modal analysis of the gyroscope structure with fixed mounting surface calculated frequency ranges of the system's first six natural vibration modes (results shown in Figure 11 and Table 1), enabling evaluation of dynamic stability under external mechanical impacts and optimization for harsh operating environments.

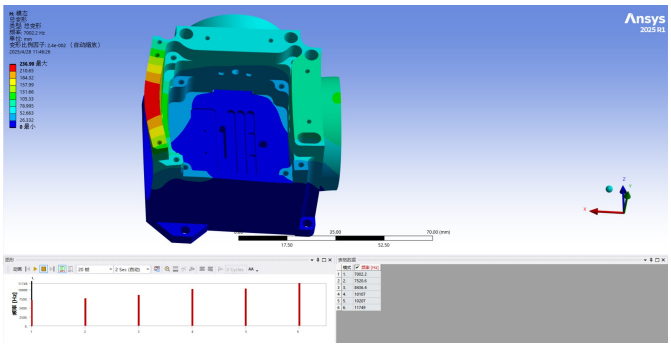


Fig. 11. Modal Analysis Results

TABLE I. FREQUENCY CHARACTERISTICS OF FIRST SIX MODAL FORMS

Mode	1st Mode	2nd Mode	3rd Mode	4th Mode	5th Mode	6th Mode
Frequency (Hz)	7002.2	7520.6	8436.4	10107	10207	11749

The primary structure demonstrates its first resonant frequency at 7002.2 Hz with complete absence of resonance points below 2000 Hz, indicating excellent mechanical resonance and vibration characteristics.

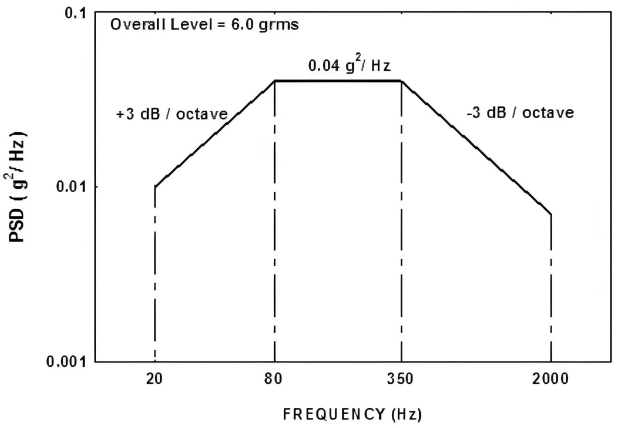


Fig. 12. Power Spectral Density of Random Vibration

Based on modal analysis and in accordance with GJB 150.16A-2009 standard, the power spectral density of random vibrations (shown in Figure 12) was determined, enabling simulation of random vibrations with calculation of deformations and equivalent stresses along X, Y, and Z axes. Results (Figures 13-16) visually demonstrate system behavior under various dynamic loads, confirming its stability and reliability under operational impacts.

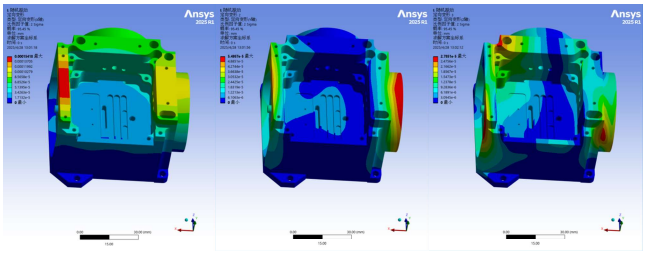


Fig. 13. Deformations Along X, Y, and Z Axes Under X-Axis Vibration.

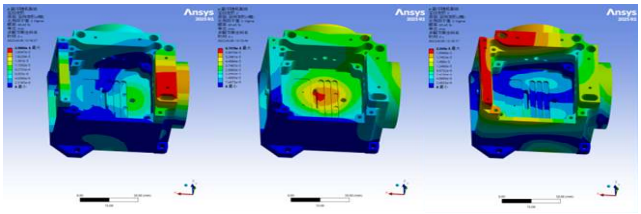


Fig. 14. Deformations Along X, Y, and Z Axes Under Y-Axis Vibration.

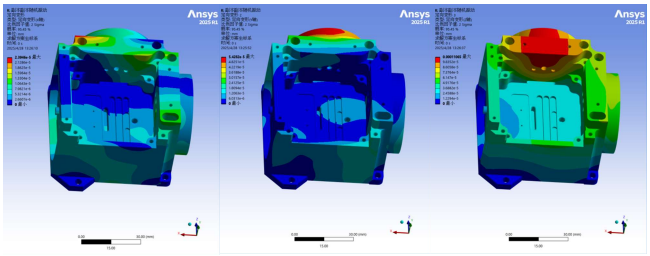


Fig. 15. Deformations Along X, Y, and Z Axes Under Z-Axis Vibration

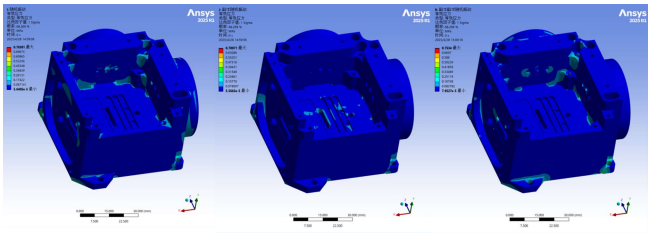


Fig. 16. Equivalent Stress of Gyroscope Under X, Y, and Z Axis Vibration

TABLE II. DEFORMATION AND EQUIVALENT STRESS RESULTS

Vibration Direction	Maximum Deformation (mm)			Equivalent Stress (MPa)
	X	Y	Z	
X	1.5×10-4	5.5×10-5	2.8×10-5	0.78
Y	2.1×10-5	6.7×10-5	2.2×10-5	0.71
Z	2.4×10-5	5.4×10-5	1.1×10-4	0.75

The TG-40 shows maximum deformation of 1.5×10^{-4} mm and minimum of 2.1×10^{-5} mm, with allowable deformation for precision instrument supports being 0.05 mm, confirming negligible deformations in all directions and exceptional vibration-resistant characteristics.

3.3 Thermal Simulation

Thermal simulation was performed under extreme temperatures (+85°C and -55°C). Results (Figures 17-18) show temperature field distributions: at +85°C, fiber coil reaches 85.7°C and board maximum 95.6°C; at -55°C, coil measures -54.4°C and board -44.2°C. The TG-40 design features low board power consumption, effective heat dissipation, and stable thermal characteristics, ensuring minimal temperature drift and reliable operation across wide temperature ranges.

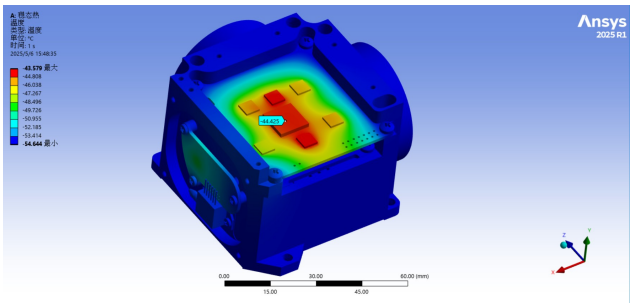


Fig. 17. Temperature Distribution Diagram at -55°C

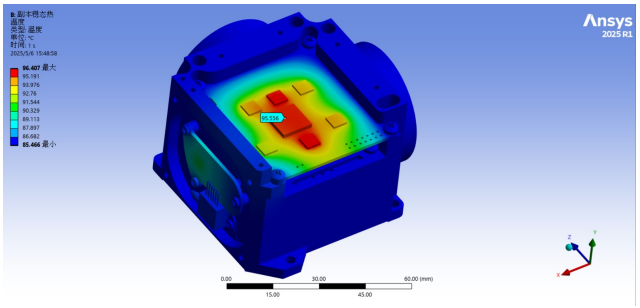


Fig. 18. Temperature Distribution Diagram at +85°C

TABLE III. AVERAGE TEMPERATURE AT FIBER RING BONDING SURFACE (WITH ASIC)

Operating Temperature (°C)	Average Temperature at Bonding Surface (°C)		
	X	Y	Z
-55	-54.433	-54.431	-53.921
+85	85.707	85.703	86.055

TABLE IV. AVERAGE TEMPERATURE AT FIBER RING BONDING SURFACE (WITHOUT ASIC)

Operating Temperature (°C)	Average Temperature at Bonding Surface (°C)		
	X	Y	Z
-55	-53.541	-53.354	-52.552
+85	86.214	86.316	87.263

IV. ANALYSIS OF PERFORMANCE CHARACTERISTICS

The fiber-optic gyroscope TG-40, developed based on a specialized ASIC integrated circuit, has completed the mass production phase. Experimental data from 500 manufactured devices (1,500 axes) were collected and analyzed, including a comprehensive evaluation of key parameters such as zero bias and scale factor.

4.1 Analysis of Zero Bias Stability

4.1.1 Zero Bias Stability at Room Temperature

This section presents the test results of the zero bias stability of the TG-40 gyroscope at room temperature over 60 minutes. The graphical representation, shown in Figure 16, demonstrates the high stability of the device's performance under normal operating conditions, achieved through the use of a specialized integrated circuit that ensures measurement accuracy and reproducibility in mass production.

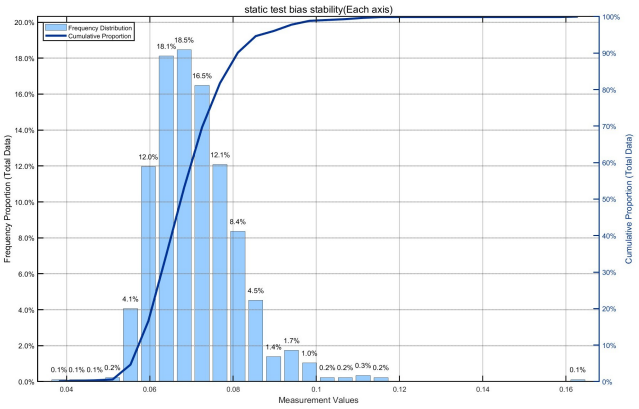


Fig. 19. Distribution of Zero Bias Stability at Room Temperature

The analysis in Figure 19 reveals that the average zero bias for the test samples is 0.0708°/h with a standard deviation of 0.0103°/h. When applying the stability criterion of zero bias not exceeding 0.1°/h over the full temperature range, the overall compliance rate reaches 98.72%.

4.1.2 Zero Bias Stability Over the Full Temperature Range

Testing was conducted across a temperature range of -55°C to +85°C with a gradient of 1°C/min and a 2-hour hold at extreme temperatures (+85°C and -55°C). The statistical data for zero bias stability over the full temperature range are displayed in Figure 20, confirming the device's reliability under extreme operating conditions due to its optimized design and specialized control IC.

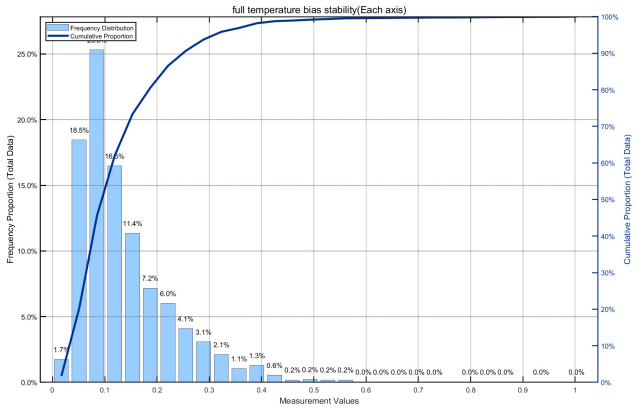


Fig. 20. Distribution of Zero Bias Stability Over the Full Temperature Range

As shown in Figure 20, the average zero bias stability for the test samples is 0.139°/h with a standard deviation of 0.0985°/h. Using an acceptability criterion of no more than 0.5°/h, the overall compliance rate reaches 99.08%, with approximately 90% of samples exhibiting zero bias stability better than 0.1°/h. This confirms the system's high thermal stability.

4.1.3 Zero Bias Instability Over the Full Temperature Range

A study was conducted on the difference between the maximum and minimum output signals of the TG-40, developed with a specialized integrated circuit, during full temperature testing. This allowed for a quantitative assessment of the fiber-optic gyroscope's thermal drift and validated the effectiveness of the technical solutions implemented to minimize temperature-induced measurement errors.

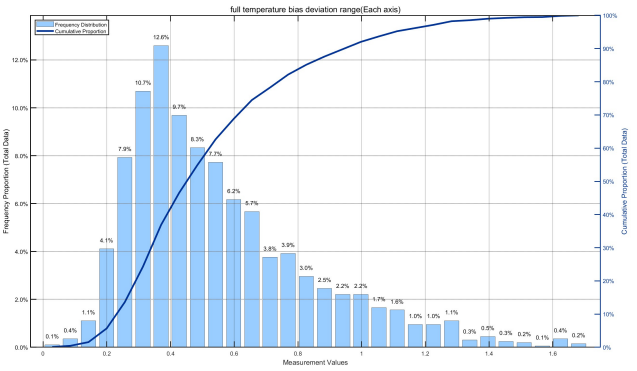


Fig. 21. Distribution of Zero Bias Instability Over the Full Temperature Range

Figure 21 shows that the average zero bias instability over the full temperature range is 0.552°/h with a standard deviation of 0.286°/h, indicating minimal thermal drift in real-world testing. Additionally, 91.01% of samples exhibit instability below 1°/h.

4.2 Scale Factor

The scale factor describes the proportional relationship between the fiber-optic gyroscope's digital output signal and the actual angular velocity. Nonlinearity of the scale factor characterizes the consistency of readings across different input values, while repeatability reflects the stability of the coefficient under repeated testing. Statistical analysis was performed on data from 500 manufactured devices (1,500 axes) of the ASIC-controlled TG-40 at room temperature, across a dynamic range of ±500°/s, with measurement control points at 0°/s, ±10°/s, ±50°/s, ±100°/s, ±200°/s, ±300°/s, ±400°/s, and ±500°/s.

4.2.1 Nonlinearity of the Scale Factor

The results of the scale factor nonlinearity analysis (shown in Figure 19) reveal an average value of 2.243 ppm with a standard deviation of 2.416 ppm and a high degree of data concentration. With an acceptability criterion of no more than 20 ppm, the overall compliance rate reaches 99.78%, confirming the device's exceptional accuracy and stability.

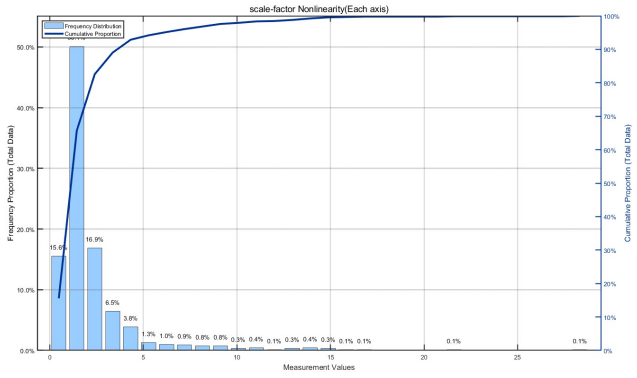


Fig. 22. Distribution of Scale Factor Nonlinearity

4.2.2 Repeatability of the Scale Factor at Room Temperature

The results of the scale factor repeatability analysis for the TG-40 fiber-optic gyroscope, presented in Figure 23, show an average value of 2.948 ppm with a standard deviation of 3.292 ppm and a high degree of data concentration. Using an acceptability criterion of no more than 20 ppm, the overall compliance rate reaches 99.65%, demonstrating exceptional stability and reproducibility in mass production. This is achieved through the use of a specialized ASIC, which ensures high measurement accuracy and minimal parameter variation between individual gyroscopes.

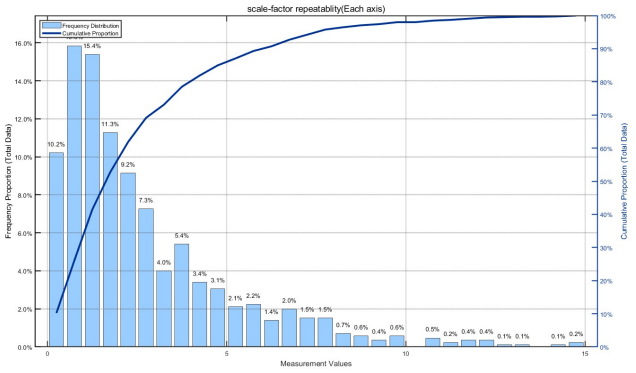


Fig. 23. Distribution of Scale Factor Repeatability

4.2.3 Repeatability of the Scale Factor Over the Full Temperature Range

To evaluate the thermal drift of the scale factor, tests were conducted at various temperature points, comparing relative deviations of the coefficient during heating and cooling phases.

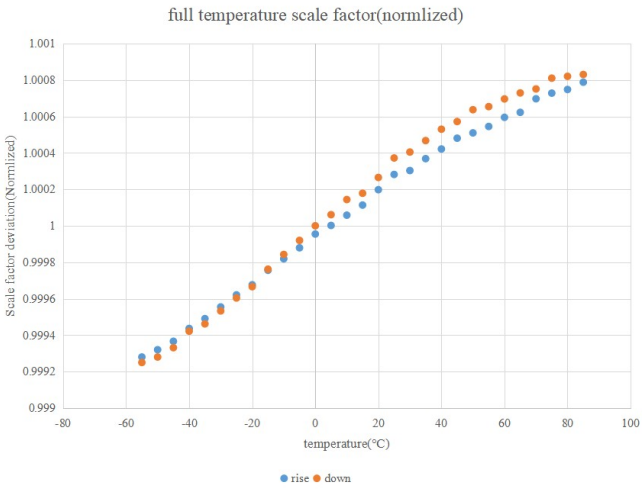


Fig. 24. Scale Factor Drift Over the Full Temperature Range

As shown in Figure 24, statistical analysis of data from 150 axes of test samples was performed relative to the baseline scale factor value at room temperature. Measurements in a thermal chamber at a controlled temperature change rate determined the relationship between the coefficient value at each temperature point and the reference value.

Figure 24 indicates that the scale factor drift of the TG-40 fiber-optic gyroscope is minimally affected by temperature, exhibiting good linearity and high compensability under thermal variations. In the low-temperature range, the scale factor shows lower sensitivity to temperature fluctuations, with high consistency between heating and cooling processes.

Overall, the difference in scale factor drift between heating and cooling modes is negligible, indicating excellent symmetry in response to temperature changes. The high stability of the scale factor across varying thermal conditions confirms the TG-40's exceptional resistance to temperature variations.

The developed three-axis integrated fiber-optic gyroscope TG-40, based on specialized integrated circuits, features compact dimensions of just 71.5 mm × 70.2 mm × 45 mm and a power consumption of only 3.3 W. Combining the advantages of miniaturization and low power consumption, the device demonstrates superior operational performance and adaptability to environmental conditions.

According to theoretical analysis and experimental data statistics:

- Zero bias stability at room temperature: <0.1°/h
- Random walk coefficient: <0.01°/h$^{1/2}$
- Average zero bias stability over the full temperature range: 0.139°/h
- Average zero bias drift over the temperature range: 0.613°/h
- Average scale factor nonlinearity: 2.243 ppm
- Average scale factor repeatability: 2.948 ppm

The scale factor exhibits excellent linear dependence and high compensability under thermal variations across the entire operating range. Additionally, the gyroscope demonstrates exceptional vibration resistance and maintains stable performance under extreme temperatures.

V. CONCLUSION

This study presents a specialized integrated circuit that combines key signal processing functions: level conversion, sample-and-hold circuitry, amplification, and differential conversion. By employing system-level layout technologies, a high degree of integration has been achieved within a single chip, encompassing digital processing, signal conversion modules, and power management. Leveraging the advantages of miniaturization and low power consumption, the circuit design of the gyroscope has been significantly reduced in size, while energy consumption has been minimized.

Testing was conducted on the performance characteristics of the developed ASIC-controlled fiber-optic gyroscope TG-40. The results demonstrate that the use of the specialized integrated circuit has enabled the TG-40 gyroscope to achieve exceptional performance in zero-bias stability, scale factor accuracy, and environmental resilience. Compared to similar solutions, this development offers distinct advantages and is well-suited for various civilian applications requiring multifunctional capabilities.

REFERENCES

[1] Liu Han. Research on Key Technologies of Fiber-Optic Gyroscope Integration [Dissertation]. China Academy of Launch Vehicle Technology, 2023. DOI:10.27096/d.cnki.ghtdy.2023.000034.

[2] Xu Yuxin. Design of FOG-Specific Integrated Circuits Based on IP Cores [Dissertation]. Guizhou University, 2024. DOI:10.27047/d.cnki.ggudu.2024.003166.

[3] Liu Qingyang. Development and Optimization of ASICs for Digital Control of FOGs [Dissertation]. Guizhou University, 2023. DOI:10.27047/d.cnki.ggudu.2023.002107.

[4] Suo Xinxin, Yu Haicheng, Feng Wenshuai, et al. Study on Interferometric FOGs Based on Multifunctional Integrated Optical Chips. Navigation and Control, 2025, vol. 24, no. 2, pp. 111–118.

[5] Li Dong. Research and Implementation of Signal Processing Circuits for FOGs [Dissertation]. Xidian University, 2023. DOI:10.27389/d.cnki.gxadu.2023.003718.

[6] Wu Shilan. Development of a High-Precision Closed-Loop FOG Based on FPGA [Dissertation]. University of Electronic Science and Technology of China, 2024. DOI:10.27005/d.cnki.gdzku.2024.005039.LEFÈVRE H C. The fiber-optic gyroscope: Challenges to become the ultimate rotation-sensing technology [J] . Opti cal Fiber Technology, 2013, 19(6) : 828-832.

[7] TRAN M A, KOMLJENOVIC T, HULME J C, et al. In tegrated optical driver for interferometric optical gyroscopes [J] . Optics Express,2017,25(4):3826-3840.

[8] RICKMAN A. The commercialization of silicon photonics [J] . Nature Photonics, 2014, 8: 579-582.

[9] YI X S, WEN X. Y-integrated optic chip (Y-IOC) applied in fiber optic gyro[C].13th International Conference on Advanced Laser Technologies, Tianjin, China, 2005: 63440U.

[10] STOPINSKI S, JUSZA A, PIRAMIDOWICZ R. An inter ferometric fiber-optic gyroscope system based on an appli cation specific photonic integrated circuit[C] . 2017Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/ Europe-EQEC) , Munich, Germany, 25-29 June, 2017: 8086879.

[11] GUO H J, LIU H F, LEI M, et al. Research progress of integrated optical gyroscope[J] . Chinese Optics Letters, 2024, 22(3) : 031302.

[12] WANG L M, HALSTEAD D R, D MONTE T, et al. Low cost, high-end tactical-grade fiber optic gyroscope based on photonic integrated circuit[C].2019 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL) , Naples, USA, 1-5 April, 2019: 8739700.

[13] Wang Zinan, Yang Yi, Li Yongxiao, Yu Xiaoqi, Zhang Zhenrong, Li Zhengbin. Quadrature demodulation with synchronous difference for interferometric fiber optic gyroscopes.[J]. Optics express, 2012, 20(23).

[14] Bo L, Tang X, Huang X. Research on Fiber Optic Gyroscope Adaptive Filtering Algorithm Based on FPGA[C], International Conference on Computational Intelligence & Communication Networks. IEEE, 2016: 16248239.

[15] Gomathi Swaminathan, G. Murugesan, S. Sasikala,L. Murali. A novel implementation of combined systolic and folded architectures for adaptive filters in FPGA[J]. Microprocessors and Microsystems, 2020, 74(prepublish).

[16] Huang Chunfu, Li An, Qin Fangjun, Wang Zhi. Fiber optic gyro signal denoising based on variational mode decomposition and adaptive noise cancellation[J]. Journal of Physics: Conference Series, 2020, 1449.

Neural Network Models for Thermal Drift Compensation of the Wave Solid-State Gyroscope

Mikhail Basarab

Dept. of Informatics and Control Systems
Bauman Moscow State Technical University
Moscow, Russia
basarab@bmstu.ru

Ivan Stroganov

Dept. of Informatics and Control Systems
Bauman Moscow State Technical University
Moscow, Russia
aug_4-4@mail.ru

Boris Lunin

Dept. of Chemistry,
Lomonosov Moscow State University
Moscow, Russia
luninboris@yandex.ru

Evgenii Chumankin

JSC ANPP
«TEMP-AVIA»
Arzamas, Russia
che54@mail.ru

***Abstract*—Thermal drift compensation of a wave solid-state gyroscope (WSG), based on linear and nonlinear autoregression models and artificial neural networks (ANN) forecast, is considered. The advantages of the approach based on machine learning neural network architectures are shown in comparison with analytical solutions obtained by the least squares method (LSM). Analysis of the results of the numerical experiment allows us to conclude that it is advisable to embed nonlinear regression forecast models based on ANN into the WSG control loop in order to compensate for thermal drift.**

Keywords— wave solid-state gyroscope, thermal drift, artificial neural networks, autoregression.

I. Introduction

The influence of temperature on the dynamics of the WSG and methods of compensation for thermal drift to improve the temperature stability of devices were considered in a number of works [1-3]. Software and algorithmic compensation is an economical and effective way to control the thermal drift of the WSG. Such compensation is convenient for design, economical from the point of view of equipment and does not require long procedures associated with its adjustment. Compensating coefficients of the polynomial regression thermal model can be estimated by the LSM. When modeling time series data of sensors, the algorithms of ordinary linear autoregression (AR) and nonlinear autoregression and moving average (ARMA) are also popular [4]. Compared with the traditional polynomial approximation of temperature, the ARMA model is capable of describing both steady-state and transient processes of gyro drift. Recently, in the practice of designing various types of navigation devices, an approach based on machine learning and ANN technologies has been actively used to solve temperature compensation problems [5-7]. This paper considers an approach to temperature compensation based on neural network implementations of the AR and ARMA models, as well as a nonlinear autoregressive neural network with external input data (NARX) [8, 9]. A conclusion is made about the advisability of embedding nonlinear regression forecast models based on ANN into the WSG control loop in order to compensate for the thermal drift of the device.

II. ARMA Model

Due to its simplicity and efficiency, the ARMA model is often used to forecast future values of time series based on historical data. A set of historical data is used for modeling, and then the ARMA components describing the relationship between the zero signal of the WSG and the corresponding value (or increment) of temperature are used to estimate the drift of the next future data (forecast). In the model, the input signal $u(t)$ and the output signal $y(t)$ are discretized in steps Δt:

$$u[k]=u(t_k),\quad y[k]=y(t_k),\quad t_k=k\Delta t,\quad k=0,1,2,\ldots \quad (1)$$

The relationship between the specified signals is expressed by means of a difference equation

$$y[k]=\sum_{i=1}^{p}a_i y[k-i]+\sum_{j=0}^{q}b_j u[k-j], \quad (2)$$

in which the uncertain coefficients a_i, b_j are found using the least squares method (LSM).

For a set of n dimensions, we introduce the following notations for vectors and matrices:

$$\mathbf{Y}=\begin{pmatrix} y[k] \\ y[k+1] \\ \vdots \\ y[k-n-2] \\ y[k-n-1] \end{pmatrix},\quad \mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix},\quad \mathbf{b}=\begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_q \end{pmatrix},$$

$$\mathbf{V}=\begin{pmatrix} y[k-1] & y[k-2] & \cdots & y[k-p] \\ y[k] & y[k-1] & \cdots & y[k+1-p] \\ \cdots & \cdots & \cdots & \cdots \\ y[k+n-2] & y[k+n-3] & \cdots & y[k+n-1-p] \end{pmatrix}, \quad (3)$$

$$\mathbf{U}=\begin{pmatrix} u[k] & u[k-1] & \cdots & u[k-q] \\ u[k+1] & u[k] & \cdots & u[k+1-q] \\ \cdots & \cdots & \cdots & \cdots \\ u[k+n-2] & u[k+n-1] & \cdots & u[k+n-1-q] \end{pmatrix}.$$

To determine the unknown coefficients, we obtain a system of linear algebraic equations (SLAE)

$$\mathbf{Y} = \underbrace{\left[\mathbf{V} \mid \mathbf{U}\right]}_{\mathbf{X}} \underbrace{\left[\frac{\mathbf{a}}{\mathbf{b}}\right]}_{\mathbf{W}}.$$

This SLAE can also be written in the form

$$\mathbf{Y} = \mathbf{X}\mathbf{W} \begin{cases} \mathbf{Y} \in \mathbb{R}^{n\times 1} \\ \mathbf{X} \in \mathbb{R}^{n\times m} \quad n > m. \\ \mathbf{W} \in \mathbb{R}^{m\times 1} \end{cases} \tag{4}$$

According to the LSM, the optimal solution of SLAE (4) is obtained from the solution of the minimization problem (in the Euclidean norm of space l_2)

$$\min \left\| \mathbf{Y} - \mathbf{X}\mathbf{W} \right\|_2^2. \tag{5}$$

The solution to equation (5) has the form

$$\mathbf{W}_{\text{opt}} = \mathbf{X}^{+}\mathbf{Y}, \tag{6}$$

where $\mathbf{X}^{+} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}$ is the pseudo-inverse matrix.

To select the order of the regression model (window size parameters p and q) and assess the quality of the forecast, one can use the root mean square deviation (RMSE) or the percentage deviation of the variance (VAF) indicators:

$$\text{VAF}(y,\hat{y}) = \left(1 - \frac{E\left\{\left[\left(y(t) - \hat{y}(t)\right) - \left(\langle y\rangle - \langle \hat{y}\rangle\right)\right]^2\right\}}{E\left\{\left[y(t) - \langle y\rangle\right]^2\right\}}\right)\%, \tag{7}$$

where $y(t)$ is the measured (reference) data; $\hat{y}(t)$ is the predicted data; $E\{\cdot\}$ is the mathematical expectation; $\langle\cdot\rangle$ is the averaging operator.

Solution (6) is obtained based on the assumption that the initial data are noise-free. However, errors inevitably arise during the measurement process, and instead of ideal undistorted vector **Y** and matrix **A**, we have to deal with their perturbed versions:

$$\mathbf{Y}_\delta = \mathbf{Y} + \delta\mathbf{Y}, \quad \mathbf{A}_\delta = \mathbf{A} + \delta\mathbf{A}.$$

Then, instead of SLAE (4), we will have a new system with respect to the optimal solution for noisy data, $\mathbf{X}_\delta$, and a model of the system with uncertainty:

$$\underbrace{\mathbf{Y} + \delta\mathbf{Y}}_{\mathbf{Y}_\delta} = \underbrace{\mathbf{A} + \delta\mathbf{A}}_{\mathbf{A}_\delta}\underbrace{\mathbf{X} + \delta\mathbf{X}}_{\mathbf{X}_\delta}. \tag{8}$$

The solution obtained using the LSM is

$$\mathbf{X}_\delta = \left(\mathbf{A}_\delta^{\mathrm{T}}\mathbf{A}_\delta\right)^{-1}\mathbf{A}_\delta^{\mathrm{T}}\mathbf{Y}_\delta. \tag{9}$$

It can be shown that a reduction in the forecast error $\mathbf{E} = \left\|\delta\mathbf{X}\right\| / \left\|\mathbf{X}\right\|$ is possible only with a simultaneous reduction in the perturbed terms $\delta\mathbf{Y}, \delta\mathbf{A}$.

For practical measurements, it is important to note that the drift of the WSG is distorted by noise, so the drift curves will be superimposed by the signal $\delta\mathbf{Y}$. In this case, not only the thermal drift of the WSG cannot be estimated explicitly, but also the accuracy and reliability of the calculation of the optimal parameters calculated by the LSM are deteriorated. To solve this problem, the primary task is to develop an appropriate filter that reduces the noise component.

To suppress noise components, it is necessary to use a suitable low-pass filter (LPF). However, the LPF inevitably causes a certain phase delay. This drawback can be avoided by using a minimum difference filter (MDF) obtained on the basis of Tikhonov regularization. The target function of the MDF filter is

$$\Phi = \sum_{i=1}^{\upsilon}\left(y[i] - \tilde{y}[i]\right)^2 + \sum_{i=2}^{\upsilon} v_{i-1}\left(\tilde{y}[i] - \tilde{y}[i-1]\right)^2, \tag{10}$$

where $y[i]$ is the original noisy data; $\tilde{y}[i]$ is the filtered data; v_{i-1} are weighting coefficients of the difference between the i-th and $(i-1)$-th samples of filtered data.

Using the LSM, we obtain the following optimal solution:

$$\tilde{\mathbf{y}}_{\text{opt}} = \left(\mathbf{I} + \mathbf{F}^{\mathrm{T}}\mathbf{V}\mathbf{F}\right)^{-1}\mathbf{y}, \tag{11}$$

where $\mathbf{y} = \left(y[1]\ y[2]\ \ldots\ y[\upsilon]\right)^{\mathrm{T}} \in \mathbb{R}^{\upsilon}$ is the vector of initial noisy data; $\tilde{\mathbf{y}}_{\text{opt}} = \left(\tilde{y}[1]\ \tilde{y}[2]\ \ldots\ \tilde{y}[\upsilon]\right)^{\mathrm{T}} \in \mathbb{R}^{\upsilon}$ is the filtered data vector; $\mathbf{I} \in \mathbb{R}^{\upsilon\times\upsilon}$ is the identity matrix; **V** is the weight matrix; **F** is the circulant matrix of the form

$$\mathbf{F} = \begin{pmatrix} -1 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -1 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

III. Neural Network Models

The procedure for calculating the pseudo-inverse matrix in (6) is numerically unstable and computationally expensive. These problems can be solved using machine learning methods based on ANN. As an example, we consider the simplest autoregressive model (AR) (the ARMA model can be implemented similarly), as well as a nonlinear autoregressive network with exogenous inputs (NARX).

A. AR Model

In the AR model, the predicted value of a time series is expressed only through its known values at previous points in time (in the ARMA model, $b_j = 0$):

$$y[k] = \sum_{i=1}^{p} a_i y[k-i] + a_0. \tag{12}$$

where $a_0 = \langle y\rangle$.

In the ANN training mode, elements of the row vectors of the training sample are fed to the network input sequentially, from epoch to epoch

$$\begin{pmatrix} y[k-1] & y[k-2] & \cdots & y[k-p] \\ y[k] & y[k-1] & \cdots & y[k+1-p] \\ \cdots & \cdots & \cdots & \cdots \\ y[k+n-2] & y[k+n-3] & \cdots & y[k+n-1-p] \end{pmatrix}, \quad (13)$$

and the output accordingly yields predicted values

$$\hat{y}[k],\ \hat{y}[k+1],\ \dots,\ \hat{y}[k+n-1].$$

These values are compared with the real ones

$$y[k],\ y[k+1],\ \dots,\ y[k+n-1]$$

and the error is estimated:

$$\varepsilon = \sqrt{\sum_k \left(y[k]-\hat{y}[k]\right)^2}. \quad (14)$$

The weights are adjusted at each epoch according to the Widrow-Hoff rule:

$$\Delta a_i = \eta \delta y[i], \quad i = 0,1,...,p, \quad (15)$$

where $\eta \in (0,1]$ is the learning rate.

In the working mode, the forecast is made by shifting the window further to the right by 1, 2, 3, etc. readings relative to the last column in (13). In this case, starting with a shift of 2, the forecast will take on values that were previously predicted, which can lead to an increase in the error of the "long-term" forecast.

B. NARX Model

Nonlinear autoregressive network with exogenous inputs (NARX) is a commonly used recurrent dynamic network designed for modeling time series and nonlinear systems [8, 9]. NARX networks are much better than conventional ANNs at detecting long-term dependencies. There are two modes of NARX network. One is a serial-parallel architecture (also called open-loop), in which the desired output signal is used in the feedback loop instead of an estimate:

$$\hat{y}[k+1] = f\begin{pmatrix} y[k],\ y[k-1],\ \dots,\ y[k-d_y]; \\ u[k+1],\ u[k],\ \dots,\ u[k-d_u] \end{pmatrix}, \quad (16)$$

where $\mathbf{y} = \left(y[k],\ y[k-1],\ \dots,\ y[k-d_y]\right)$ are true prior time series data; $f(\cdot)$ is a transfer function; $\mathbf{u} = \left(u[k+1],\ u[k],\ \dots,\ u[k-d_u]\right)$ is an input signal NARX; d_y is the output signal delay; d_u is the input signal delay.

Another type of NARX is a parallel architecture (also called closed-loop or open-loop), in which the previous computational output is fed back to the input of a feedforward neural network. This architecture is described by the expression

$$\hat{y}[k+1] = f\begin{pmatrix} \hat{y}[k],\ \hat{y}[k-1],\ \dots,\ \hat{y}[k-d_y]; \\ u[k+1],\ u[k],\ \dots,\ u[k-d_u] \end{pmatrix}, \quad (17)$$

where $\hat{\mathbf{y}} = \left(\hat{y}[k],\ \hat{y}[k-1],\ \dots,\ \hat{y}[k-d_y]\right)$ are previously evaluated output signal data.

It is advisable to use a serial-parallel architecture at the stage of ANN training, and a parallel one at the stage of its testing.

The detailed architecture of NARX is described by the expression

$$y(t) = \mathbf{B}_0 + \mathbf{W}_0 \tanh\left[\mathbf{B}_i + \mathbf{W}_i \mathbf{X}_d(t) + \mathbf{W}_l \mathbf{Y}_d(t)\right], \quad (18)$$

where $\mathbf{B}_0 \in \mathbb{R}^{1\times 1}$, $\mathbf{B}_i \in \mathbb{R}^{N\times 1}$ are bias weights; $\mathbf{W}_0 \in \mathbb{R}^{1\times N}$, $\mathbf{W}_i \in \mathbb{R}^{N\times p}$, $\mathbf{W}_l \in \mathbb{R}^{N\times p}$ are synaptic weights.

IV. Numerical Experiment

Let us consider the forecast of the zero signal of the WSG depending on the temperature change (Fig. 1). Here the step is $\Delta t = 1\,s$. We will limit ourselves to the time interval from 2380 to 2980 s and pre-process the signal with the MDF filter. The matrix of weights **V** in (10) is diagonal with elements $v_{ii} = 50$.

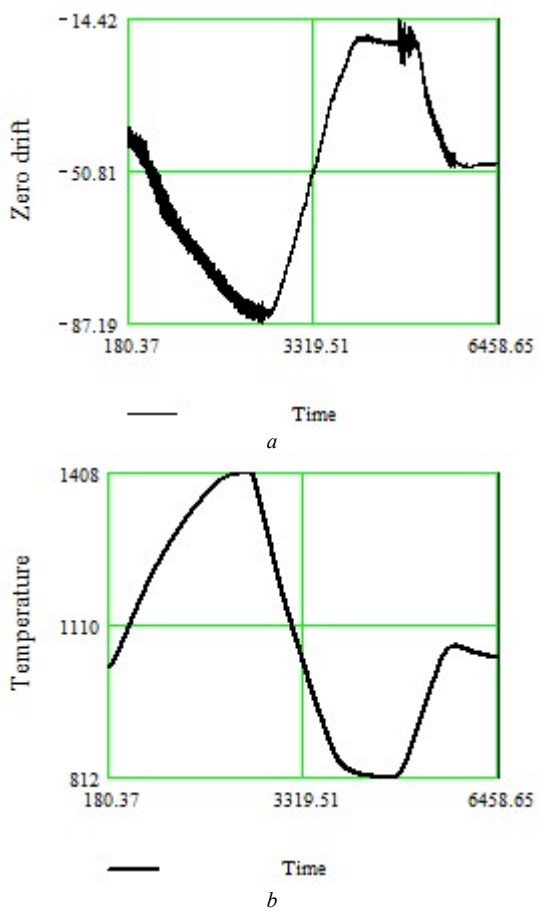


Fig. 1. Dependence of the zero signal (a) and temperature (b) on time

Figure 2 shows the training and forecast results of the ARMA and AR models. The following parameters were selected in (2): $p = 4, q = 0$. Training interval: from 2384 to 2480 s; Forecast interval: from 2480 to 2576 s. The ARMA coefficients found were

$$a_1 = 0.559,\ a_2 = 0.504,\ a_3 = -0.234,\ a_4 = 0.156;$$
$$b_0 = -9.241\cdot 10^{-4}.$$

The ANN training was carried out over 50 000 epochs. The calculated weighting coefficients were:

$w_1 = 0.216$, $w_2 = 0.273$, $w_3 = 0.658$, $w_4 = -1.416 \cdot 10^{-4}$, $w_0 = -0.305$.

Forecast metrics for both models:

$RMSE_{AR} = 0.044$, $VAF_{AR} = 92.09$ %,
$RMSE_{ARMA} = 0.026$, $VAF_{ARMA} = 94.52$ %.

Thus, the forecast quality of the ARMA model even with one input for the temperature signal is slightly higher than that of the conventional AR model based on ANN. At the same time, the neural network model has higher robustness and can be improved by switching to the NARX architecture.

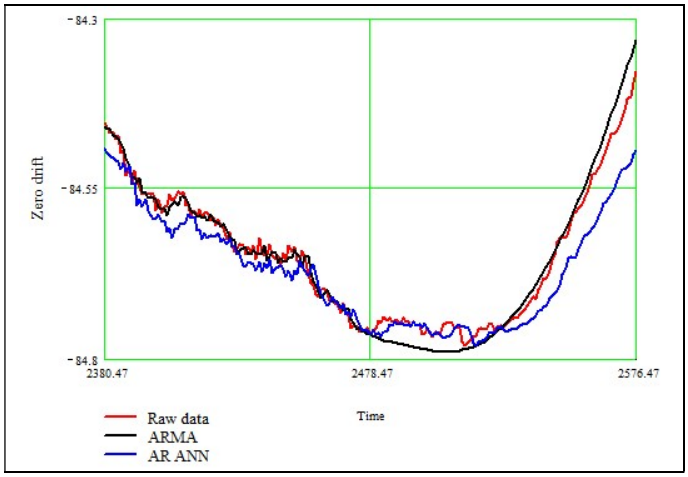


Fig. 2. The result of training and predicting the zero signal

CONCLUSION

The results of numerical experiments show wide possibilities for using ANN for constructing autoregressive (AR and ARMA) models of forecasting the WSG signal in the presence of noise depending on the external dynamic temperature effect. Using real data as an example, a comparison of AR (based on ANN) and ARMA models was carried out. Both models showed satisfactory forecast quality (after preliminary MDF filtering). It can be concluded that it is expedient to embed AR and NARX ANN models in the WSG control loop in order to compensate for thermal drift.

REFERENCES

[1] X. Wang, W. Wu, Z. Fang, B. Luo, Y. Li, and Q. Jiang, "Temperature drift compensation for hemispherical resonator gyro based on natural frequency," Sensors, vol. 12, no. 5, pp. 6434–6446, May 2012.

[2] Z. Xu, G. Yi, and W. Zhu, "An accurate thermoelastic model and thermal output error analysis of a hemispherical resonator gyroscope under varying temperatures," Mechanical Systems and Signal Processing, 2022, vol. 170, 108760.

[3] L. Xin-san, L. Can, S. Qiang, et al., "A multiple regression based method for indirect compensation of hemispherical resonator gyro temperature error," Scientific Reports, 2023, vol. 13, 6019.

[4] S.-J. Huang and K.-R. Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," IEEE Trans. Power Syst., vol. 18, no. 2, pp. 673–679, May 2003.

[5] G. Araghi and R. J. Landry, "Temperature compensation model of MEMS inertial sensors based on neural network," Proc. IEEE/ION Position, Location Navigat. Symp. (PLANS), pp. 301–309, April 2018.

[6] S. Chong et al., "Temperature drift modeling of MEMS gyroscope based on genetic-Elman neural network," Mech. Syst. Signal Process., vol. 72, pp. 897–905, May 2016.

[7] H. Cao, Y. Zhang, C. Shen, Y. Liu, and X. Wang, "Temperature energy influence compensation for MEMS vibration gyroscope based on RBF NN-GA-KF method," Shock Vib., vol. 2018, pp. 1–10, December 2018.

[8] H. Xie, H. Tang, and Y.-H. Liao, "Time series prediction based on NARX neural networks: An advanced approach," Proc. Int. Conf. Mach. Learn. Cybern., vol. 3, pp. 1275–1279, July 2009.

[9] E. Diaconescu, "The use of NARX neural networks to predict chaotic time series," WSEAS Trans. Comput. Res., vol. 3, no. 3, pp. 182-191, March 2008.

Effect of Assembly Errors on the Static Capacitance of the Hemispherical Resonator Gyroscope

Y. Xie
College of Advanced Interdisciplinary Studies, National University of Defense Technology,
Nanhu Laser Laboratory, National University of Defense Technology,
Changsha, China
xieyuxuan24a@163.com

W. Wu
College of Advanced Interdisciplinary Studies, National University of Defense Technology,
Nanhu Laser Laboratory, National University of Defense Technology,
Changsha, China
wuweihiter@163.com

J. Li
College of Advanced Interdisciplinary Studies, National University of Defense Technology,
Nanhu Laser Laboratory, National University of Defense Technology,
Changsha, China
lijingyu13c@nudt.edu.cn

Y. Pan
College of Advanced Interdisciplinary Studies, National University of Defense Technology,
Nanhu Laser Laboratory, National University of Defense Technology,
Changsha, China
panyao08@nudt.edu.cn

Y. Tao
College of Advanced Interdisciplinary Studies, National University of Defense Technology,
Nanhu Laser Laboratory, National University of Defense Technology,
Changsha, China
taoyunfengcn@foxmail.com

H. Luo
College of Advanced Interdisciplinary Studies, National University of Defense Technology,
Nanhu Laser Laboratory, National University of Defense Technology,
Changsha, China
luohui.luo@163.com

Abstract*—Enhancing the assembly precision of the resonator and flat electrode is essential for ensuring optimal performance of the hemispherical resonator gyroscope. Recent studies have demonstrated that identifying assembly errors through the measurement of static capacitance between the resonator and the flat electrode is a viable approach. However, the majority of these studies are based on an equivalent mathematical model of the capacitor formed by the resonator's end surface and the flat electrode, thereby neglecting the influence of the resonator's internal spherical surface. In this study, theoretical analysis and finite element method (FEM) analysis are employed to investigate the impact of assembly errors on static capacitance. First, a simplified theoretical model is derived to establish a relationship between static capacitance and assembly errors, then, a finite element model is developed that incorporates both the resonator's end surface and its internal spherical surface to accurately calculate the static capacitance affected by assembly errors, finally, the capacitance variation under different assembly errors is analyzed using these two models. This study focuses on the Type 20 resonator and identifies the assembly inclination angle error and direction angle error as the primary factor affecting the static capacitance value. The mathematical model that solely considers the capacitance directly opposite the end surface underestimated results, failing to accurately reflect the actual changes in static capacitance. Under the condition of an average gap of 70-90 μm, the difference in capacitance values obtained from finite element simulations that include the inner surface area compared to those that only consider the area directly opposite the end surface is approximately 0.15-0.16 pF. Moreover, this difference is found to be insensitive to inclination angle errors. The results of this study are of significant importance for optimizing assembly error identification methods based on static capacitance and for improving the accuracy of assembly error detection.

Keywords—hemispherical resonator gyroscope, assembly errors, static capacitance, finite element simulation

This work was funded by the National Natural Science Foundation of China (NSFC) (52305562).

I. INTRODUCTION

The hemispherical resonator gyroscope (HRG) is a solid-state vibratory gyroscope that measures angular velocity or angle based on the Coriolis effect. It offers several advantages, including high precision, a simple structure, good reliability, strong shock resistance, and an extended service life[1]. The assembly of the resonator with the flat electrode represents a critical manufacturing step for the HRG, necessitating exceptionally high standards of assembly precision.

Although the structure of the widely used two-piece HRG is relatively simple[2-4], comprising only two core components—the resonator and flat electrode—it imposes extremely stringent requirements on assembly precision. The gap between the end surface of the resonator and the flat electrode must be controlled at a level of tens of micrometers, while the circumferential gap non-uniformity must remain below 2%. This assembly precision not only influences the consistency of excitation gain for the flat electrode and detection gain for the resonator[5] but also affects the vibrational performance of the resonator[6, 7]. Consequently, these factors ultimately impact the overall performance of the gyroscope.

The identification of assembly errors constitutes a critical step in the assembly process, as it directly influences both assembly accuracy and the final performance of the product[8]. Theoretical models of the relationship between capacitance and assembly errors have been widely studied. In practical assembly processes for HRG, the occurrence of assembly errors is unavoidable. The current mainstream methods for detecting these errors primarily include optical measurement and static capacitance-based measurement. Among these techniques, static capacitance-based measurement has gained significant attention due to its efficiency and practicality, positioning it as a key area of research. Analyzing the effect of assembly errors on static capacitance is essential for optimizing related technologies.

Wei[9] developed an equivalent planar capacitor model incorporating assembly error parameters and introduced an equivalent electrode angle to simplify the model. By analyzing the impact of assembly errors on HRG capacitance, the study successfully bridged the gap between the practical system and theoretical framework. Ruan[10] proposed an assembly error identification method based on static capacitance, and this method can identify assembly inclination and direction angle errors by combining an ideal parallel plate capacitor model with genetic and Nelder-Mead algorithms. Yuan[11] presented a static capacitance-based assembly gap measurement scheme and analyzed the impact of assembly errors on static capacitance. Yu[12] proposed an assembly pose error identification method for planar electrode HRG. Using a support vector regression (SVR) model, they further constructed a nonlinear relationship between equivalent capacitance and assembly pose errors, enabling efficient identification of assembly pose errors. Existing research is mostly limited to parallel plate capacitor mathematical models, treating the resonator and flat electrode as an equivalent parallel plate capacitor and processing it with mathematical approximations. The discrepancy between mathematical and actual models reduces the accuracy of assembly error identification, ultimately affecting the HRG's precision. Thus, accurately analyzing the impact of assembly errors on static capacitance has become a critical challenge.

Based on the parallel plate capacitor model, this paper examines how inclination and direction angle errors affect static capacitance, obtaining results on how static capacitance varies with these errors. Additionally, this study combines finite element simulation to establish a finite element model that includes the internal spherical coating area of the resonator, analyzing the impact of assembly direction angle error, and assembly inclination error on static capacitance. This study facilitates a precise analysis of the relationship between assembly errors and the static capacitance of the HRG, thereby enhancing the accuracy of identifying assembly errors in HRG.

II. The Effect Of Assembly Errors On The Static Capacitance Based On The Theoretical Model

The excitation and detection of the two-piece HRG primarily rely on flat electrodes, which are composed of eight individual electrodes arranged on a flat electrode disk. The coated resonator is positioned at the center of this disk, with its edge forming a parallel plate capacitor structure in conjunction with the flat electrode. According to the principles governing parallel plate capacitors, any alteration in the relative positioning of these electrodes leads to charging and discharging events, resulting in variations in capacitance.

During the assembly process involving both the resonator and the flat electrode, discrepancies may arise due to assembly errors. Once stabilized, these errors manifest as differences in static capacitance values among the eight parallel plate capacitors. By analyzing these capacitance value discrepancies, potential assembly errors associated with the resonator can be effectively identified. The parallel plate capacitor formed by the metal film layer on the resonator's end surface and the flat electrode, free from assembly errors, can be described as Eq. 1:

$$C = \frac{\varepsilon_0 \varepsilon_r S}{d_0} \tag{1}$$

where ε_0 is the vacuum permittivity, ε_r is the relative permittivity, S is the facing area of the electrode plates, and d_0 is the initial gap.

In the actual assembly process of the HRG, due to the influence of factors such as machining errors, clamping precision, and measurement errors, the assembly results of the resonator and the flat electrode inevitably deviate from the design objectives. The primary deviations include the inclination angle error γ, the direction angle error $\theta_{\phi 0}$, and the average gap error Δd, as shown in Fig. 1. The inclination angle error γ refers to the angle between the resonator's rotational axis and the flat electrode's rotational axis during assembly. The direction angle error $\theta_{\phi 0}$ denotes the angle between the projection of the resonator's rotational axis onto the upper surface of the flat electrode and the x-axis. The mean gap error Δd represents the difference between the actual and designed mean gaps.

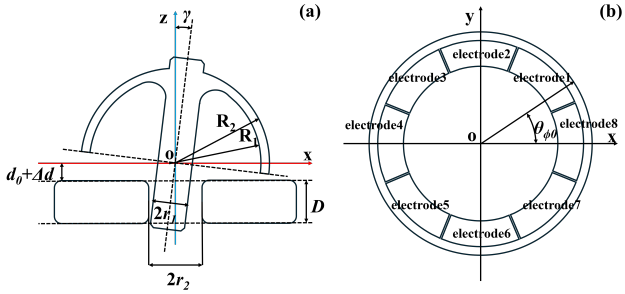


Fig. 1. Schematic of assembly errors in HRG (a) side view (b) top view

The main factors affecting the end surface capacitance of the HRG are the inclination angle error, direction angle error, and average gap of the end surface. According to the assembly relationship between the resonator and the flat electrode, the inclination angle error is limited by the dimensions of the HRG, as demonstrated in Eq. 2.

$$\begin{cases} \sin \gamma_{MAX} \le \dfrac{d}{R_2} \\ \sin \gamma_{MAX} \le \dfrac{r_2 - r_1}{D} \end{cases} \tag{2}$$

where γ_{MAX} is the maximum value of the inclination angle error, d is the actual average gap, R_2 is the outer radius of the resonator, r_1 is the radius of the support column, D is the thickness of the flat electrode, r_2 is the radius of the fitting hole in the resonator support column.

The Type 20 HRG, a typical two-piece structure, has the following parameters: the inner radius of the resonator, the outer radius, the average gap, the radius of the support column, the thickness of the flat electrode, and the radius of the fitting hole in the resonator support column. The size parameters of Type 20 HRG are shown in TABLE I.

According to the Eq. 2 and HRG dimensional parameters as presented in TABLE I, the assembly inclination angle error ranges from [0, 15] arcmin. Due to the circumferential symmetry between the resonator and the flat electrode, the direction angle error is simplified to a range of [-22.5, 22.5] degrees. Set the electrode plate where the direction angle error

is located as Electrode 1, and number the electrodes clockwise in sequence as Electrode 1 to 8, as shown in Fig. 1(b).

TABLE I. PARAMETER OF TYPE 20 HRG

Parameter	Value
Inner radius	10 mm
Outer radius	11 mm
Average gap	70-90 μm
Radius of the support column	2.5 mm
Thickness of the flat electrode	5 mm
Radius of the fitting hole	2.6 mm

Based on the parallel plate capacitor model, a capacitance theoretical model for the flat electrode HRG can be established by considering assembly inclination angle error, direction angle error and average gap error, as shown in Eq. 3.

$$C = \int_{R_1}^{R_2} \int_{\theta_{\phi_1}}^{\theta_{\phi_2}} \frac{\varepsilon r}{d_0 + \Delta d - r\cos\left(\theta_\phi - \theta_{\phi_0}\right)\sin\gamma} \mathrm{d}r\,\mathrm{d}\theta_{\phi_0} \quad (3)$$

Within the specified ranges of assembly errors, the inclination angle error primarily affects the gap between the electrode and the resonator end face, thereby influencing the capacitance value. The direction angle error, on the other hand, mainly affects the distribution of the gap around the entire circumference, thus impacting the distribution of capacitance values. Using the theoretical model established in Eq. 2, we calculated the changes in capacitance for the eight electrodes by introducing various errors in assembly inclination angles and direction angles, with an average gap of 70 micrometers. The range of assembly inclination angle error is defined as [0, 15] arcmins, while the direction angle error is simplified to a range of [0, 22.5] degrees, as illustrated in Fig. 2.

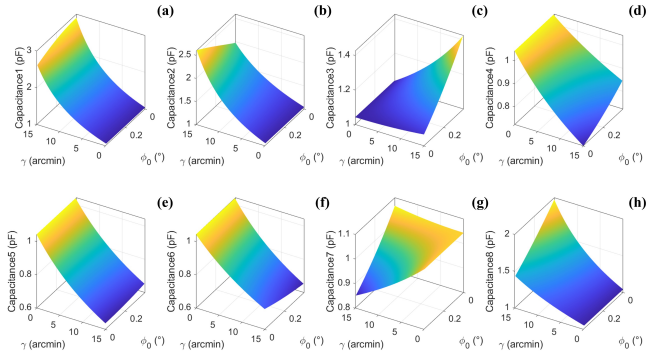


Fig. 2. The Impact of Assembly Errors on Static Capacitance at 70 μm Gap based on the theoretical model (a)Electrode 1 (b) Electrode 2 (c) Electrode 3 (d) Electrode 4 (e) Electrode 5 (f) Electrode 6 (g) Electrode 7 (h) Electrode 8

The capacitance of electrode plates facing the direction angle is most sensitive to inclination angle changes. By comparing electrode 1 and 5, 2 and 8, 3 and 7, and 4 and 6, the capacitances of the eight electrodes are symmetrically distributed with respect to the direction angle error, with the line connecting electrodes 1 and 5 serving as the axis of symmetry. The capacitance of the electrode where the direction angle error is located is most sensitive to the inclination angle error. When the inclination angle error changes from 0 to 15 arcmin, Electrode 1's capacitance changes from 1.04 pF to 2.89 pF. Capacitance changes caused by the direction angle error are less than 0.5 pF, indicating that the resonator's inclination angle error affects electrode capacitance more significantly than the direction angle error.

III. THE EFFECT OF ASSEMBLY ERRORS ON THE STATIC CAPACITANCE BASED ON THE FINITE ELEMENT MODEL

This study employs a finite element model (FEM) to analyze the capacitance of the parallel plate capacitor formed between the resonator and the excitation/detection electrode plate. The electrostatic module of COMSOL Multiphysics is employed to determine the steady-state charge distribution.

A FEM of the resonator and the flat electrode of Type 20 HRG is built in COMSOL Multiphysics, as illustrated in Fig. 3. The mesh is refined to ensure circumferential symmetry and uniformity, with an increased mesh density at the details of the electrode to enhance calculation precision as Fig. 3 (a). Assembly errors are simulated by adjusting the positional relationship between the resonator and flat electrode. Boundary conditions are applied to the end face of the resonator, its internally coated spherical area, and the surface of the support column to replicate actual coating conditions. The eight electrodes on the flat electrode are designated as terminal potentials, while a ground boundary condition is assigned to the grounded section, as illustrated in Fig. 1. The static capacitance of each electrode is calculated by analyzing terminal charges alongside spatial potential distribution.

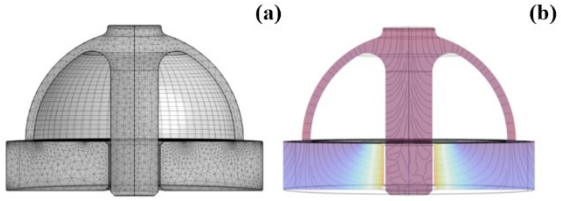


Fig. 3. FEM of HRG (a) Meshing (b) Electric Potential and Charge Simulation Instance

The parameters used in the FEM are shown in the TABLE II.

TABLE II. SIMULATION PARAMETER USED IN THE FEM

Parameter	Value
Resonator radius	10 mm
Average resonator gap	70 μm
Flat electrode spanning angle	44 degrees

The study investigates the impact of assembly inclination angle error and direction angle error, on the capacitance of eight electrodes. Fig. 4 illustrates the differences in capacitance between simulated and calculated values for these eight electrodes under an average gap of 70 μm, with the assembly inclination angle error varying from [0, 15] arcmins. The results for each electrode demonstrate the discrepancies between FE simulation values and calculated capacitances influenced by both assembly inclination angle error and direction angle error. For simplicity, due to symmetry considerations, the direction angle error is constrained to a range of [0, 22.5] degrees. The findings indicate that variations in capacitance are particularly sensitive to changes in inclination angle errors.

The capacitance differences among the eight electrodes, influenced by inclination and direction angle errors, exhibit distinct values. These differences align with the trends predicted by the mathematical model. Overall, the FE simulation results are consistent with the mathematical model, but the capacitance values from the simulations are larger than those from the calculations, with differences ranging from 0.3 to 0.5 pF Given the measurement error of approximately 0.01 pF in practical operations, the capacitance differences between the calculation model and the simulation model cannot be ignored.

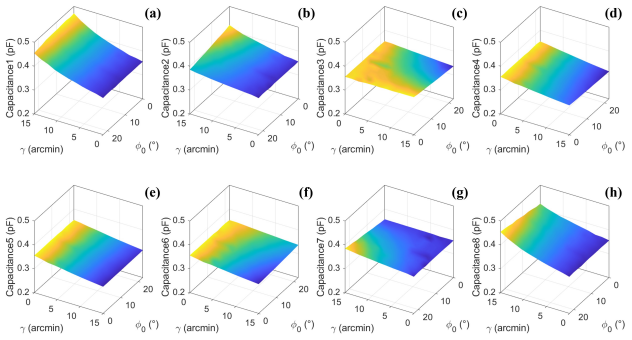


Fig. 4. The Impact of Assembly Errors on Capacitance Differences Between FEM and Mathematical models for Resonator End Face at 70 μm Gap (a)Electrode 1 (b) Electrode 2 (c) Electrode 3 (d) Electrode 4 (e) Electrode 5 (f) Electrode 6 (g) Electrode 7 (h) Electrode 8

The comparison between FE simulation results and mathematical model results shows that the capacitance values obtained from FE simulations are generally larger than those from calculations. According to electrostatic field theory, the electric field distribution of an ideal parallel plate capacitor can be derived from Maxwell's equations. The mathematical model for ideal capacitance presumes infinite plate dimensions and disregards fringing effects, resulting in a uniform electric field established between the plates. However, in reality, the electric field of a finite-sized plate capacitor is significantly distorted due to fringing effects. At the plate edges, electric field lines diverge outward, resulting in a larger capacitance value than theoretical predictions[13]. Many theoretical models ignore fringing effects, such as Eq. 3, have notable errors in practical applications.

The theoretical model that considers only the capacitance directly opposite the end surface tends to underestimate the actual capacitance and does not accurately capture its true variations. Calculating the capacitance associated with the internal spherical coated area of the resonator presents significant challenges when employing conventional capacitance models. FE simulations were therefore used to analyze this region. When the average gap is 70 μm, and the assembly inclination angle error is in the range of [0, 15] arcmins and the azimuth angle error is in the range of [0, 22.5] degrees, the capacitance difference between the FE models with and without the inner surface is analyzed to obtain the inner surface capacitance results. Fig.5 shows the results of how the inner surface capacitance values of the eight electrodes are affected by inclination and direction angle errors. The capacitance variation in the inner surface region is approximately 0.15–0.16 pF which is insensitive to changes in inclination and azimuth angle errors.

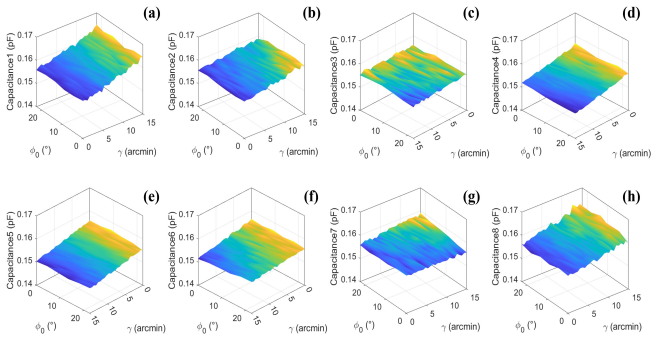


Fig. 5. The Impact of Assembly Errors on Capacitance for Resonator Inner Surface at 70μm Gap (a)Electrode 1 (b) Electrode 2 (c) Electrode 3 (d) Electrode 4 (e) Electrode 5 (f) Electrode 6 (g) Electrode 7 (h) Electrode 8

FE analysis was used to investigate the effect of different average gaps on the inner surface capacitance. Fig.6 shows the average differences in capacitance values for eight electrodes with and without the inner spherical coating, under average gaps of 70, 80, and 90 μm. The capacitance variation in the inner coating region is approximately 0.15 – 0.16 pF and is insensitive to inclination angle errors. As the average gap increases, the average capacitance of this inner region decreases slightly but remains non-negligible in practical applications.

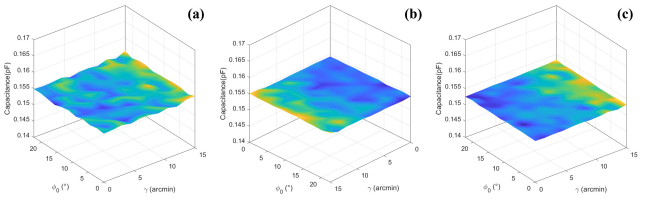


Fig. 6. Average Capacitance of Eight Electrodes Impact by Resonator Inner Surface at (a) 70μm (b) 80μm (c) 90μm Gap

IV. CONCLUSION

This study combines theoretical method and Finite element method to investigate how assembly errors affect static capacitance. It finds that inclination angle error is the key factor influencing capacitance. The theoretical model can partially reflect how assembly errors affect capacitance. Capacitance values from FE simulations are larger than those from the theoretical model, indicating that the theoretical model is not precise enough and not rigorous enough for accurately identifying assembly errors.

For the Type 20 HRG with an average gap of 70 – 90 μm, FE results show that capacitance values are 0.3–0.5 pF larger than those from the mathematical model. FE simulations show a capacitance difference of approximately 0.15 – 0.16 pF between models with and without the internal spherical coating. This indicates significant deviations in theoretical models. Additionally, this difference demonstrates insensitivity to errors in inclination angle, underscoring the advantages of FE simulation in precisely evaluating the effects of assembly errors on capacitance.

REFERENCES

[1] Y. Pan, T. Qu, K. Yang, and H. Luo, " Hemispherical Resonator Gyroscope: Current Advances and Future Trends," Navigation, Positioning & Timing, vol. 4, no. 2, pp. 9-13, 2017.

[2] H. Peng, Z. Fang, W. Tan, H. Fang, and L. Yan, " The Technical Characteristics of Hemispherical Resonator Gyro Development," Navigation, Positioning & Timing, vol. 6, no. 04, pp. 108-114, 2019.

[3] A. Jeanroy, A. Bouvet, and G. Remillieux, "HRG and marine applications," Gyroscopy and Navigation, vol. .5, no. .2, pp. 67-74, 2014.

[4] F. Delhaye, HRG by SAFRAN: the game-changing technology. Safran Electronics & Defense, Boulogne-Billancourt, France, 2018.

[5] Z. Ruan, X. Ding, Z. Qin, H. Li, and IEEE, "Compensation of Assembly Eccentricity Error of Micro Hemispherical Resonator Gyroscope," in IEEE International Symposium on Robotic and Sensors Environments (ROSE), 2021.

[6] Z. Ruan, X. Ding, Y. Gao, Z. Qin, and H. Li, "Analysis and Compensation of Bias Drift of Force-to-Rebalanced Micro-Hemispherical Resonator Gyroscope Caused by Assembly Eccentricity Error," *Journal of Microelectromechanical Systems,* vol. Vol.32, no. No.1, pp. 16-28, 2023

[7] K. Deng et al., "Separation of mass and gap errors of hemispherical resonator gyroscopes by varying bias voltage," Measurement, vol. 228, p. 114308, 2024.

[8] X. Jin et al., " Development Status and Trend of Hemispherical Resonator Gyroscope Assembly Technology," Flight Control & Detection, vol. 4, no. 01, pp. 1-10, 2021.

[9] Z. Wei, G. Yi, Y. Huo, Z. Qi, and Z. Xu, "The Synthesis Model of Flat-Electrode Hemispherical Resonator Gyro," Sensors, vol. 19, no. 7, Apr 1 2019.

[10] Z. Ruan, X. Ding, Z. Qin, J. Jia, and H. Li, "Modeling and Compensation of Assembly Inclination Error of Micro Hemispherical Resonator Gyroscope Under Force-to-Rebalance Mode," IEEE Sensors Journal, vol. 21, no. 13, pp. 14726-14738, Jul 1, 2021.

[11] L. Yuan, Q. Zeng, C. Wang, S. Ren, Z. Wei, and Y. Huo, "High Precision Assembly and Welding Method for Flat-Electrode Hemispherical Resonator Gyro," IEEE Transactions on Instrumentation and Measurement, vol. 72, p. 1, 2023.

[12] H. Yu et al., "Identification method for assembly pose errors of flat electrode hemispherical resonator gyro," IEEE Sensors Journal, vol. 24, no. 4, p. 1, 2023.

[13] H. Niu et al., "Fringing-Effect-Based Capacitive Proximity Sensors," Advanced Functional Materials, vol. 34, no. 51, p. 1, 2024.

Dynamic Modeling and Mass Balancing of Hemispherical Resonators with Variable Thickness

Wei Cheng
School of Astronautics
Harbin Institute of Technology
Harbin, China
21b904034@stu.hit.edu.cn

Shunqing Ren
School of Astronautics
Harbin Institute of Technology
Harbin, China
renshunqing@hit.edu.cn

Boqi Xi
School of Astronautics
Harbin Institute of Technology
Harbin, China
xibq@hit.edu.cn

Zhen Tian
School of Astronautics
Harbin Institute of Technology
Harbin, China
22b904048@stu.hit.edu.cn

Youhuan Ning
School of Astronautics
Harbin Institute of Technology
Harbin, China
20b904047@stu.hit.edu.cn

Changhong Wang
School of Astronautics
Harbin Institute of Technology
Harbin, China
cwang@hit.edu.cn

***Abstract*—This paper presents a novel dynamic model for the hemispherical resonator with variable thickness, which comprehensively considers the uneven mass distribution in both circumferential and meridional directions. By analyzing the impact of mass distribution on the vibrational performance of the resonator, a model for identifying vibration parameters is proposed. Furthermore, a high-precision mass balancing experiment is designed to eliminate unbalanced mass utilizing ion beam etching technology. Through multiple etching experiments, the frequency split of the resonator sample is reduced by 99.71%, which validates the accuracy of the theoretical model and the effectiveness of the mass balancing scheme. This study lays the foundation for the fabrication of high-precision hemispherical resonator gyroscopes.**

Keywords—Hemispherical Resonator, Dynamic Model, Variable Thickness, Frequency Split, Ion Beam Etching

I. Introduction

Hemispherical Resonator Gyroscopes (HRGs) represent a new type of inertial sensor, which have advantages such as long lifespan, low power consumption, radiation resistance, and small size, thereby garnering significant interest within the global inertial navigation community [1,2]. As the core component of the HRG, the vibration characteristics of the hemispherical resonator crucially impacts output accuracy of the HRG [3]. This paper presents a novel dynamic model for the hemispherical resonator with variable thickness and designs an innovative mass balancing scheme. First, based on the thin shell theory, the equation of motion for the hemispherical resonator is derived with Lagrange equation. Then, by solving the characteristic equation, the natural frequencies of the resonator in the second-order resonance state are obtained and the frequency split is calculated. Utilizing the polar decomposition method, the stiffness axis is determined. Based on vibration envelope analysis and nonlinear optimization method, we develop an identification model for vibration parameters. Finally, to reduce the impact of unbalanced mass on the frequency split, we design an innovative mass balancing scheme that integrates harmonic etching with point etching.

II. Dynamic Model of Resonator

A. Variable Thickness

First, taking into account the nonuniformity of the mass distribution in both circumferential and meridional directions, its thickness can be mathematically described as

$$h(\alpha,\beta) = h_0 + \sin\alpha \sum_{n=1}^{\infty} h_n \cos[n(\beta-\beta_n)], \tag{1}$$

where h_0 is the average thickness, h_n is the n^{th} Fourier harmonic amplitude of the thickness, and β_n is the initial phase of the n^{th} Fourier harmonic.

As shown in Fig. 1, along the meridional direction, when the parameter α is 0, the wall thickness of the resonator attains its minimum value; conversely, when α is $\pi/2$ (at the edge of the resonator), the wall thickness reaches its maximum value.

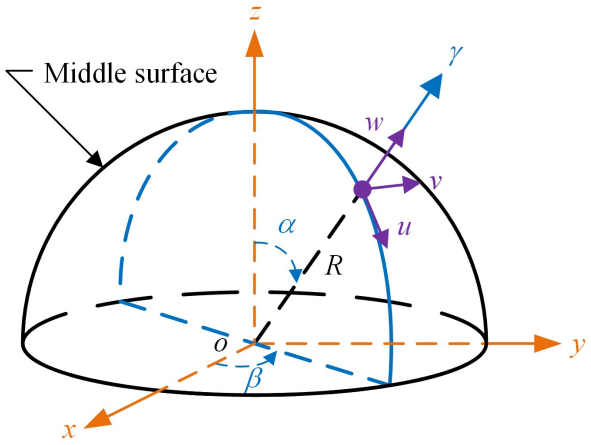


Fig. 1. Mathematical model of the hemispherical resonator. *o-xyz* is the inertial coordinate system.

B. Kinetic Energy and Strain Energy

As shown in Fig. 1, the displacement of any point on the middle surface can be expressed as [4]

$$\begin{cases} u(\alpha,\beta,t)=\sin\alpha\tan^2\dfrac{\alpha}{2}[p(t)\cos 2\beta+q(t)\sin 2\beta], \\ v(\alpha,\beta,t)=\sin\alpha\tan^2\dfrac{\alpha}{2}[p(t)\sin 2\beta-q(t)\cos 2\beta], \\ w(\alpha,\beta,t)=(2+\cos\alpha)\tan^2\dfrac{\alpha}{2}[p(t)\cos 2\beta+q(t)\sin 2\beta], \end{cases} \tag{2}$$

where $p(t)$ and $q(t)$ are the physical coordinates that represent the second-order operating mode of the resonator.

Then, the kinetic energy of the hemispherical resonator is written as

$$E_K=\int_{-\frac{h}{2}}^{\frac{h}{2}}\int_0^{2\pi}\int_0^{\frac{\pi}{2}}\frac{1}{2}(\dot{u}^2+\dot{v}^2+\dot{w}^2)\rho R^2\sin\alpha\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma, \tag{3}$$

where R represents the radius of the middle surface and ρ represents the density of the resonator.

The geometric equation for the hemispherical resonator can be expressed as [5]

$$\begin{cases} e_1=\dfrac{1}{R}(\dfrac{\partial u}{\partial\alpha}+w)+(-\dfrac{1}{R^2}\dfrac{\partial^2 w}{\partial\alpha^2}-\dfrac{w}{R^2})\gamma, \\ e_2=\dfrac{1}{R\sin\alpha}(\dfrac{\partial v}{\partial\beta}+u\cos\alpha+w\sin\alpha) \\ \quad+(-\dfrac{1}{R^2\sin^2\alpha}\dfrac{\partial^2 w}{\partial\beta^2}-\dfrac{\cot\alpha}{R^2}\dfrac{\partial w}{\partial\alpha}-\dfrac{w}{R^2})\gamma, \\ e_{12}=\dfrac{1}{R\sin\alpha}(\dfrac{\partial u}{\partial\beta}+\dfrac{\partial v}{\partial\alpha}\sin\alpha-v\cos\alpha) \\ \quad+2(-\dfrac{1}{R^2\sin\alpha}(\dfrac{\partial^2 w}{\partial\alpha\partial\beta}-\cot\alpha\dfrac{\partial w}{\partial\beta}))\gamma, \end{cases} \tag{4}$$

where e_1, e_2, and e_{12} are the normal and shear strains in the directions of α and β, respectively. The physics equation for the hemispherical resonator can be expressed as

$$\begin{cases} \sigma_1=\dfrac{E}{1-\mu^2}(e_1+\mu e_2), \\ \sigma_2=\dfrac{E}{1-\mu^2}(e_2+\mu e_1), \\ \tau_{12}=\dfrac{E}{2(1+\mu)}e_{12}, \end{cases} \tag{5}$$

where E represents the Young's modulus, and μ represents the Poisson's ratio. σ_1 and σ_2 are the normal stresses in the directions of α and β respectively. τ_{12} is the shear stress.

Thus, the strain energy of the resonator is expressed as

$$E_S=\int_{-\frac{h}{2}}^{\frac{h}{2}}\int_0^{2\pi}\int_0^{\frac{\pi}{2}}\frac{1}{2}(\sigma_1 e_1+\sigma_2 e_2+\tau_{12}e_{12})R^2\sin\alpha\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma. \tag{6}$$

C. Equation of Motion

Based on the kinetic energy and strain energy of the resonator, the Lagrange equation is defined as

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}(\dfrac{\partial L}{\partial\dot{p}})-\dfrac{\partial L}{\partial p}=0, \\ \dfrac{\mathrm{d}}{\mathrm{d}t}(\dfrac{\partial L}{\partial\dot{q}})-\dfrac{\partial L}{\partial q}=0, \end{cases} \tag{7}$$

where L= E_K-E_S. Substituting (3) and (6) into (7), for simplification, the equation of motion for the resonator are obtained as

$$\begin{cases} 2(a_1+a_2)\ddot{p}+a_3\ddot{q}+2(c_1+c_2)p+c_3 q=0, \\ 2(a_1-a_2)\ddot{q}+a_3\ddot{p}+2(c_1-c_2)q+c_3 p=0. \end{cases} \tag{8}$$

The coefficients a_1, a_2, a_3, c_1, c_2, and c_3 are integral constant values, which are complex but do not affect the subsequent analysis, so their specific expressions are not provided here.

Traditional methods, such as the point mass method and variable density method, have only addressed the impact of unbalanced mass on the kinetic energy. In contrast, this study employs a Fourier series expansion of the thickness, which not only considers the effects of unbalanced mass on the kinetic energy but also fully incorporates its influence on the potential energy.

To facilitate the following analysis, (8) is rewritten as a matrix

$$\boldsymbol{M}\ddot{\boldsymbol{X}}+\boldsymbol{K}\boldsymbol{X}=\boldsymbol{0}, \tag{9}$$

where the physical coordinates vector $\boldsymbol{X}$, mass matrix $\boldsymbol{M}$ and stiffness matrix $\boldsymbol{K}$ are

$$\boldsymbol{M}=\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}, \boldsymbol{K}=\begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}, \boldsymbol{X}=\begin{bmatrix} p \\ q \end{bmatrix}. \tag{10}$$

III. Frequency Split and Stiffness Axis

The characteristic equation for the resonator model is $|\boldsymbol{K}-\omega^2\boldsymbol{M}|=0$, where ω is the natural frequency of the resonator. Then, we can obtain frequencies ω_1 and ω_2, as follows

$$\begin{cases} \omega_1=\sqrt{\dfrac{k_{11}m_{22}-2k_{12}m_{12}+k_{22}m_{11}+c_c}{2(m_{11}m_{22}-m_{12}^2)}}, \\ \omega_2=\sqrt{\dfrac{k_{11}m_{22}-2k_{12}m_{12}+k_{22}m_{11}-c_c}{2(m_{11}m_{22}-m_{12}^2)}}. \end{cases} \tag{11}$$

The coefficient c_c in (11) is

$$\begin{aligned} c_c=&[(k_{11}m_{22}-k_{22}m_{11})^2 \\ &+4(k_{11}m_{22}-k_{22}m_{12})(k_{12}m_{11}-k_{11}m_{12})]^{\frac{1}{2}}. \end{aligned} \tag{12}$$

Therefore, the frequency split of the resonator is

$$\Delta f=\frac{|\Delta\omega|}{2\pi}=\frac{|\omega_1-\omega_2|}{2\pi}. \tag{13}$$

The eigenvectors form a modal matrix $\boldsymbol{C}$, which is the coordinate transformation matrix between the physical coordinates and the principal coordinates.

$$\boldsymbol{C}=\begin{bmatrix}\dfrac{k_{22}m_{11}-k_{11}m_{22}-c_c}{2k_{11}m_{12}-2k_{12}m_{11}} & \dfrac{k_{22}m_{11}-k_{11}m_{22}+c_c}{2k_{11}m_{12}-2k_{12}m_{11}}\\ 1 & 1\end{bmatrix}. \quad (14)$$

Here, the eigenvectors in the modal matrix contain information about the azimuth of stiffness axis. By using polar decomposition method [6], a pure rotation matrix $\boldsymbol{R}$ can be extract from the modal matrix $\boldsymbol{C}$.

$$\boldsymbol{C}=\boldsymbol{R}\boldsymbol{S}, \quad (15)$$

where $\boldsymbol{S}$ represents the scaling matrix. The rotation matrix $\boldsymbol{R}$ can be calculated by continuously averaging the transpose of the inverse of matrix $\boldsymbol{C}$ until convergence.

$$\boldsymbol{C}_{i+1}=\frac{1}{2}[\boldsymbol{C}_i+(\boldsymbol{C}_i^{\mathrm{T}})^{-1}]. \quad (16)$$

Then, the form of the rotation matrix $\boldsymbol{R}$ is

$$\boldsymbol{R}=\begin{bmatrix}R_{11} & R_{12}\\ R_{21} & R_{22}\end{bmatrix}=\begin{bmatrix}\cos 2\theta & -\sin 2\theta\\ \sin 2\theta & \cos 2\theta\end{bmatrix}. \quad (17)$$

The azimuth of stiffness axis can be obtained as

$$\theta=\frac{1}{2}\arctan\frac{-R_{12}}{R_{11}}\pm n\times\frac{\pi}{4},\quad n=0,1,2,\cdots. \quad (18)$$

IV. Identification Model

Considering the energy dissipation of resonator, the radial vibration velocity of its lip edge is expressed as [7]

$$\begin{aligned}vel_w=&-A_0e^{-t/\tau}[\omega_1\cos^2(2\theta)\sin(\omega_1 t)\\&+\omega_2\sin^2(2\theta)\sin(\omega_2 t)],\end{aligned} \quad (19)$$

where A_0 is the initial amplitude of the resonator and τ is the time constant of free decay.

Equation (19) undergoes squaring operations and is subsequently filtered through a low-pass filter. The remaining portion after filtering is

$$\begin{aligned}\Gamma(t)=&\frac{1}{4}A_0^2(\omega_1^2+\omega_2^2)e^{-2t/\tau}\{3+\cos(8\theta)\\&+[1-\cos(8\theta)]\cos 2\pi\Delta ft\}.\end{aligned} \quad (20)$$

The envelope of the vibration velocity signal of the resonator is

$$S_E(t)=\sqrt{2\Gamma(t)}. \quad (21)$$

Given that the non-uniformity of the quality factor of the resonator can induce frequency shifts of the principal standing wave and the phenomenon of self-locking [7], we correct (20) and reformulated it in a discrete form.

$$\begin{aligned}\Gamma(n)=&\frac{1}{4}A_0^2(\omega_1^2+\omega_2^2)e^{-2nT_s/\tau}\{3+\cos(8\theta+\lambda nT_s)\\&+\cos(2\pi\Delta fnT_s+\varphi_0)[1-\cos(8\theta+\lambda nT_s)]\}\\&+S_N(n),\end{aligned} \quad (22)$$

where φ_0 is the initial phase of free vibration envelope signal, n is the serial number of the sampling point. T_s is the sampling period and $S_N(n)$ is the set of zero-mean interference and noise signals. λ is the equivalent drift rate of the principal standing wave.

The objective function is

$$\begin{aligned}F_O(n)=&\sum_{n=1}\{\Gamma(n)-\frac{1}{4}\hat{A}_0^2(\omega_1^2+\omega_2^2)e^{-2nT_s/\tau}\\&\{3+\cos(8\hat{\theta}+\hat{\lambda}nT_s)+[1-\cos(8\hat{\theta}+\hat{\lambda}nT_s)]\\&\cos(2\pi\Delta\hat{f}nT_s+\hat{\varphi}_0)\}\}^2.\end{aligned} \quad (23)$$

Based on the Nelder-Mead Simplex Method [8], the optimal parameters $\hat{A}_0$, $\hat{\theta}$, $\hat{\lambda}$, $\Delta\hat{f}$ and $\hat{\varphi}_0$ can be identified when the objective function $F_O(n)$ reaches its minimum value. ω_1 and ω_2 can be obtained by frequency analysis of the original vibration signal of the resonator.

Furthermore, the quality factor of the resonator can be obtained as

$$Q=\frac{\tau}{2}\sqrt{\frac{\omega_1^2+\omega_2^2}{2}}. \quad (24)$$

V. Experiment and Results

A. Experimental design

The ion beam etching system is shown in Fig. 2. It primarily consists of a vacuum box, an ion beam source, a control cabinet, and an isolation platform.

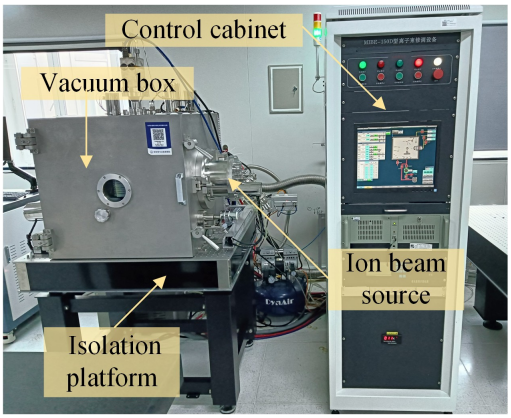


Fig. 2. Ion beam etching system.

The function of the ion beam source is to produce ions and accelerate these ions to high velocities. They are ejected from the diaphragm to impact the outer surface of the resonator, thereby achieving atomic-level mass. The control cabinet is utilized to adjust various parameters of the ion beam and to ensure that the pressure within the vacuum box is maintained at levels below 5×10^{-3} Pa. The isolation platform can effectively suppresses the influence of external vibrations on the experimental results, ensuring the accuracy and reliability of the data obtained.

As the frequency split of the resonator decreases, its sensitivity to mass removal significantly increases. The harmonic etching offers higher efficiency, it is relatively lower in precision, whereas point etching provides higher precision but at a lower efficiency. To address this, we designed an innovative mass balancing scheme that integrates harmonic etching with point etching. Specifically, harmonic etching is first employed to reduce the frequency

split of the resonator to 0.003 Hz, followed by point etching to further lower the resonator frequency split to below 0.0005 Hz.

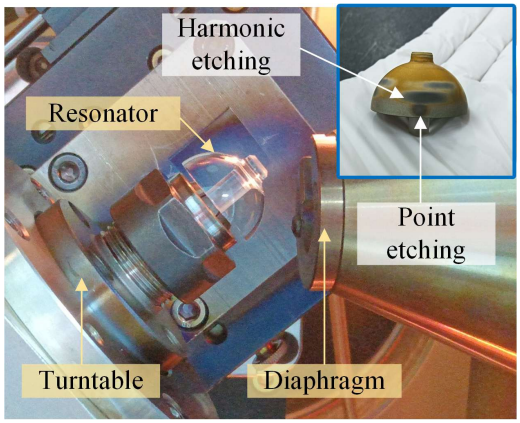


Fig. 3. Etching attitude of the hemispherical resonator.

As shown in Fig. 3, by controlling the turntable in the vacuum box, harmonic etching refers to the continuous mass removal where the ion beam follows a fourth-order harmonic distribution along the direction of β on the surface of the resonator. In contrast, point etching involves mass removal at four specific points that are equally spaced along the direction of β, which is a discrete etching process.

Before mass balancing of the hemispherical resonator, a layer of 280 nm thick gold film was deposited on the outer surface of the resonator using magnetron sputtering technology. As shown in Fig. 3, the morphological changes of the gold film upon ion beam etching were observed to calibrate the mass removal effects of both point and harmonic ion beam etching.

B. *Experimental results*

To verify the correctness of theoretical model and the effectiveness of mass balancing scheme, experiments were conducted on a hemispherical resonator sample. The experimental results are shown in Fig. 4.

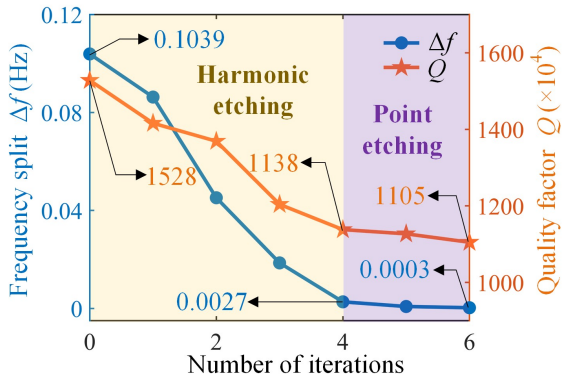


Fig. 4. Experimental results of multiple iterations.

After four iterations of harmonic etching, the frequency split of the hemispherical resonator sample decreased from 0.1039 Hz to 0.0027 Hz, as shown in Fig. 4. Subsequently, following two iterations of point etching, the frequency split was further reduced to 0.0003 Hz, ultimately achieving a 99.71% reduction in frequency split. This validated the accuracy of the theoretical model proposed in this study and demonstrate the effectiveness of its mass balancing scheme.

In addition, after multiple iterations, the quality factor of the resonator decreased from 15.28 million to 11.05 million, reduced by 27.68%. The quality factor of the hemispherical resonator is primarily influenced by anchor loss and surface loss. During the mass balancing process, ion beam etching induces varying degrees of contamination of the resonator surface, subsequently leading to a deterioration in the quality factor.

Finally, future research will focus on the specific mechanisms of post-etch surface treatment to gain deeper insights into these variations and to seek effective strategies for enhancing the quality factor of the resonator.

VI. Conclusions

In this paper, a novel dynamic model is developed, incorporating the effects of uneven mass distribution in both circumferential and meridional directions. Through multiple iterations of ion beam etching, the frequency split of the resonator sample was reduced by 99.71%, from 0.1039 Hz to 0.0003 Hz, demonstrating the high precision and effectiveness of the balancing scheme proposed in this study. This significant reduction in frequency split validates the accuracy of the theoretical model. However, it was observed that the quality factor decreased by 27.68% during the etching process, primarily due to surface contamination induced by ion beam etching.

This study provides a theoretical foundation and an effective mass balancing methodology for the fabrication of high-precision hemispherical resonators, which can facilitate the development of HRGs with improved accuracy and reliability.

References

[1] K. Yan, X. Wang, K. Zou et al., Self-Excitation Enabled Decoupling, Calibration, and Compensation of Errors for Whole-Angle Hemispherical Resonator Gyroscope, IEEE Transactions on Instrumentation and Measurement, 2024, vol. 73, 1001613.

[2] A. Maslov, D. Maslov, I. Ninalalov, I. Merkuryev, Hemispherical resonator gyros (an overview of publications), Gyroscopy and Navigation, 2023, vol. 14, no. 1.

[3] D. M. Rozelle, The Hemispherical Resonator Gyro: From wineglass to the planets, Advances in the Astronautical Sciences, 2009, pp. 1157-1178.

[4] W. Cheng, S. Ren, B. Xi et al., The Influence of Nonlinear High-Intensity Dynamic Processes on the Standing Wave Precession of a Non-Ideal Hemispherical Resonator, 2024, Sensors, vol. 24, no. 9, 2709.

[5] Y. Ning, C. Wang, W. Cheng et al., Anchor loss improvement in hemispherical resonators with ion beams, International Journal of Mechanical Sciences, 2024, vol. 277, 109413.

[6] K. Shoemake, T. Duff, and C. C. S. Canadian Human, MATRIX ANIMATION AND POLAR DECOMPOSITION, Proceedings - Canadian Information Processing Society, 1992, pp. 258-264.

[7] Y. Huo, Z. Wei, S. Ren et al., High Precision Mass Balancing Method for the Fourth Harmonic of Mass Defect of Fused Quartz Hemispherical Resonator Based on Ion Beam Etching Process, IEEE Transactions on Industrial Electronics, 2023, vol. 70, no. 9, pp. 9601-9613.

[8] J. C. Lagarias, J. A. Reeds, M. H. Wright et al., Convergence properties of the Nelder-Mead simplex method in low dimensions, Siam Journal on Optimization, 1998, vol. 9, no. 1, pp. 112-147.

The Implementation of the Hilbert Transform of the Primary Signal of a Wave Solid-State Gyro

S.E. Perelyaev
Ishlinsky Institute of Problems of Mechanics of the Russian Academy of Sciences, Moscow, Russia,
e-mail: *sergei-perelyaev@mail.ru*

A.V. Alekhin
JSC "Inertial technologies of technocomplex", Ramenskoye, Russia,
e-mail: *aalehin@itts.ru*

Abstract. **This report addresses the solution to a relevant practical problem – the formation of a quadrature signal when processing the primary signal of a wave solid-state gyroscope (WSG). It is known that the accuracy of quadrature signal formation determines the main errors and dynamics of the wave gyroscope under various operating conditions. The report proposes a new method for generating the quadrature signal using the Hilbert transform. One implementation of the digital Hilbert transforms for solving this problem is examined. The proposed methodology allows for the realization of the synthesized algorithm in a modern programmable system on a chip (FPGA).**

Keywords – wave solid-state gyroscope, WSG, quadrature signal, state vector.

I. INTRODUCTION

Today, all promising control systems for wave solid-state gyroscopes (WSG) are based on digital signal processing (DSP). DSP methods enhance the accuracy of measurement systems and enable the implementation of more complex error models of instruments. Signal processing schemes using DSP are free from these drawbacks, and the realization of DSP algorithms does not pose significant challenges. The foundation of DSP methods includes spectral analysis, Fourier transform, Hilbert transform, Kalman filter, and wavelet transform. The Hilbert transform is a traditional approach to signal processing in radio and high-frequency techniques, but it is rarely applied in mechanics. However, results obtained from test trials of WSG have demonstrated the effectiveness and broad capabilities of this method for identifying the quadrature component. The control loop of WSG requires the formation of the derivative of the primary signal. DSP allows this task to be solved in many ways. There are numerical differentiation techniques; however, they are accompanied by substantial noise in the output signal. An algorithm based on the optimal Kalman filter (OKF) has proven to be very effective in many modern and promising developments. It is worth noting that a large number of experimental studies in recent years indicate that the control loop is most effectively implemented on the platform of a modern system on chip - field-programmable gate array (FPGA). The Kalman filter, despite all its advantages, requires a sufficiently complex mathematical signal model of the dynamic system and a large number of floating-point operations for implementation. The OKF is implemented on a microcontroller, however, it is very challenging to implement it on domestic FPGAs due to a number of computational resource limitations. The advantage of FPGAs lies in their high speed and parallel processing of primary signals. However, for floating-point operations, domestic systems on a chip typically lack sufficient resources. In such conditions, the best choice is the implementation of the digital Hilbert transform. The algorithm is relatively straightforward to implement on any digital platform, is resistant to interference, and has low computational complexity. However, it is known that the classical Hilbert integral in the time domain has a singular point when $t-\tau\to 0$, where the principal value of the Cauchy integral is used for computation.

II. PROBLEM STATEMENT

Let's consider the solution to a relevant applied problem – the formation of a quadrature signal during the processing of the primary signal from a WSG. We will examine one implementation of the digital Hilbert transform to address this task. Let's denote the primary signal of the device as $x(t)$. The samples of this signal take the form of discrete data stored in the memory of a digital computer. It is required to form a signal that will be phase-shifted by exactly 90 degrees relative to the signal. All operations must be conducted in real-time. The solution to the problem is provided by the Hilbert transform of the primary signal. However, in its classical form, this transformation is difficult to implement in real-time. Therefore, it is proposed to apply the Fourier transform to realize the digital Hilbert transform. The Fourier transform allows for a relatively straightforward implementation in real-time, including systems on FPGA/PLIS chips. This paper addresses the important task of processing the primary signal of the WSG – the formation of the quadrature component of the signal, which enables the assessment of the necessary parameters of the gyroscope's state vector.

III. SOLUTION OF THE PROBLEM

To solve the posed problem, it is useful to introduce the concept of an analytical signal. Let us now define the concept of an analytical signal. An analytical signal is defined as a complex signal of the form:

$$z(t) = x(t) + iy(t)$$

This is a complex signal, where the real part coincides with our primary signal, and the imaginary part is the quadrature signal. The analytical signal can be geometrically represented as follows. Let's consider a spatial curve defined in parametric form:

$$x(t) = sin(t),\ y(t) = cos(t),\ z(t) = t \quad (1)$$

The curve defined by equations (1) describes a known spiral in space (fig. 1), which can be associated with the analytical signal. We will construct a graph of the analytical signal in a three-dimensional coordinate system. Along the x-axis, we will represent our original signal $x(t)$, along the y-axis the imaginary part $iy(t)$, and along the z-axis the analytical signal $z(t)$.

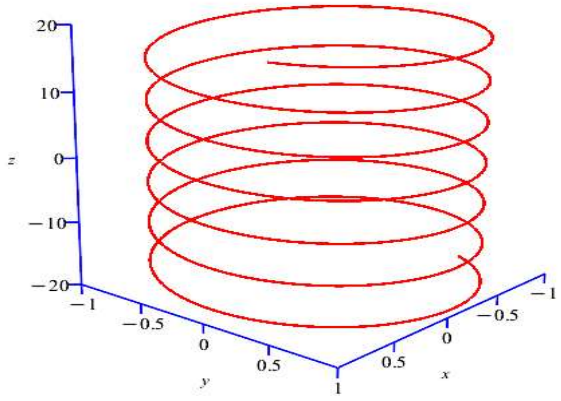


Fig. 1. Representation of the analytical signal.

In the considered analytical signal $z(t)$, the real part is known to us, and we will find the imaginary part using the Hilbert transform (the input of this transform is the signal $x(t)$, and the output is the analytical signal $y(t)$).

By definition, the Hilbert transform is an ideal wideband phase shifter of the original signal by 90 degrees [1]. All harmonics of the primary signal $x(t)$ are phase-shifted by 90 degrees. It is important to note that, in addition to the phase rotation, the Hilbert filter removes the DC component of the signal. Using this definition, we will construct the amplitude-frequency response ic (AFR) and the phase-frequency response (PFR) of the Hilbert transform (see fig. 2).

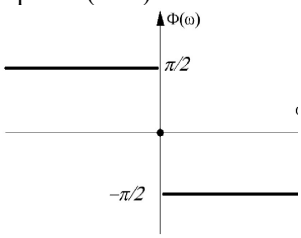


Fig. 2a. PFR

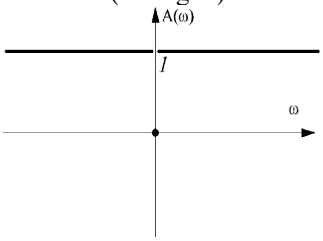


Fig. 2b. AFR

The constant component is eliminated by the Hilbert transform, while the amplitude of the output signal remains unchanged. Based on the form of the amplitude and phase frequency characteristics, we can write the formula for the complex spectral density $H(\omega)$ of the Hilbert transform:

$$H(\omega) = \begin{cases} e^{\frac{i\pi}{2}}, \omega < 0 \\ e^{-\frac{i\pi}{2}}, \omega > 0 \\ 0, \omega = 0 \end{cases} \quad (2)$$

Let's assume that the spectral density of the signal $y(t)$ is equal $Y(\omega)$, and the spectral density of the input signal $x(t)$ is equal $X(\omega)$. Then, obviously the following formula

$$Y(\omega) = H(\omega)X(\omega) \quad (3)$$

We remind you that $X(\omega), Y(\omega)$ - complex functions. Let's assume that the spectral density of the analytical signal $z(t)$ is $Z(\omega)$. Then, the following analytical equality is true:

$$Z(\omega) = X(\omega) + iY(\omega) \quad (4)$$

By virtue of the validity of relation (3), we obtain:

$$Z(\omega) = X(\omega)(1 + iH(\omega)) \quad (5)$$

Let's introduce a new function $S(\omega)$:

$$S(\omega) = 1 + iH(\omega) \quad (6)$$

We see from [1] that the function $S(\omega)$, is the spectral density of the transformation connecting the spectral densities of the input signal $X(\omega)$ and the analytical signal $Z(\omega)$:

$$Z(\omega) = X(\omega)S(\omega) \quad (7)$$

Using the analytical expression (2), we will immediately find the formula for the function $S(\omega)$:

$$S(\omega) = \begin{cases} 0, \omega < 0 \\ 2X(\omega), \omega > 0 \\ X(0), \omega = 0 \end{cases} \quad (8)$$

Let's note that the analytical signal $Z(\omega)$ we obtained does not have negative frequencies in its spectrum. Thus, using the formulas (7) and (8), we have linked the spectral densities of the input signal $x(t)$ and the desired analytical signal $z(t)$. However, the spectral density is defined as the result of the Discrete Fourier Transform (DFT). If we compute the DFT of the input signal $x(t)$, we can unambiguously find the analytical signal by using the Inverse Discrete Fourier Transform (IDFT) of the function $S(\omega)$, thereby obtaining the signal $z(t)$. By extracting the imaginary part of the analytical signal $z(t)$, we obtain the desired quadrature signal. Let's take samples of the real signal and demonstrate all the actions we described above with a specific example. Fig. 3 shows a discrete signal for control testing, which has approximately 23 discrete samples per period.

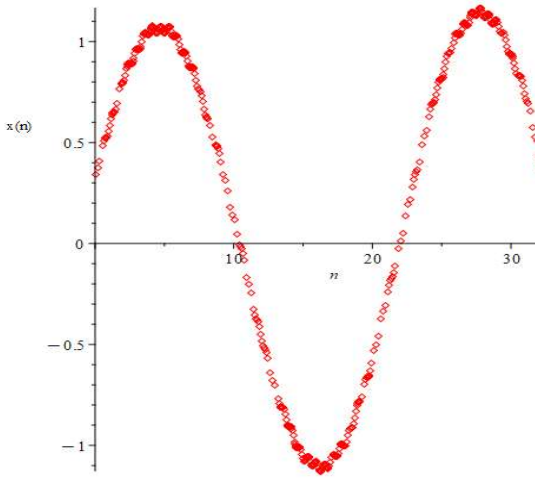


Fig. 3. Test signal.

We determine the DFT of the signal in fig.3 using the formula:

$$S_k = \sum_{n=0}^{N-1} 2x_n e^{-\frac{2i\pi kn}{N}}, k > 0 \quad (9)$$

$$S_0 = \sum_{n=0}^{N-1} x_n$$

where N is the period of the signal (in our case $N = 23$).

The multiplier of two in the analytical expression (9) appeared from the formula (8) synthesized above. We determine the IDFT of the signal S_k using the following formula:

$$z_k = \frac{1}{N}\sum_{n=0}^{N-1} S_n e^{\frac{2i\pi kn}{N}} \quad (10)$$

After performing calculations using numerical methods, we obtain an approximation of the analytical signal z_k:

$$\begin{aligned} z_k = {} & 0.0098\,sin(3.4k) + sin(0.27k) \\ & +0.32\,cos(0.26k) + 0.017\,cos(10.63k) \\ & +i(-0.0098\,cos(3.4k) - cos(0.27k) \\ & +0.32\,sin(0.26k) + 0.017\,sin(10.63k)) \end{aligned} \quad (11)$$

Accordingly, based on the analytical expression (11) obtained by us, we have

$$\begin{aligned} x_k &= Re(z_k) \\ y_k &= Im(z_k) \end{aligned} \quad (12)$$

In formula (12), x_k is the input signal, y_k is the quadrature signal. For example, we will calculate the amplitude of the input signal using the formula:

$$a_k = \sqrt{x_k^2 + y_k^2} \quad (13)$$

The remaining parameters of the state vector can also be easily found [2], knowing the quadrature signal, but this is a topic for a separate article. Fig. 4 shows the signals -x_k input signal, y_k - quadrature signal, a_k - amplitude.

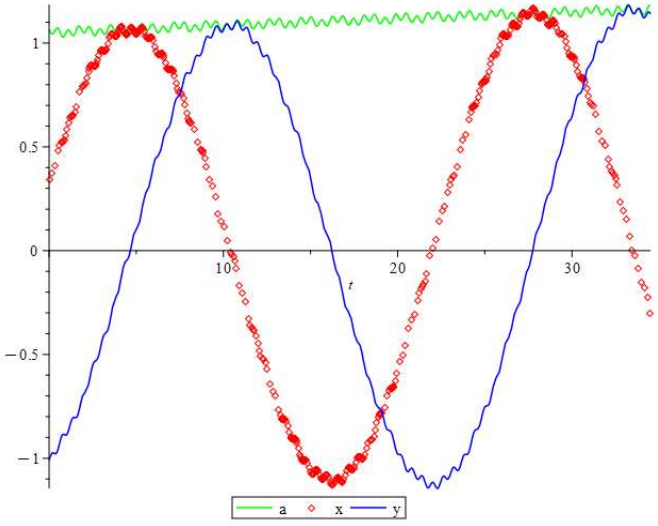


Fig. 4. Result of the Hilbert transform.

IV. RESULTS

The report presents one of the methods for implementing the Hilbert transform using the discrete Fourier transform method, the computation algorithm of which is programmed on a modern FPGA (Field-Programmable Gate Array) system. The approach is based on finding the transfer function of the Hilbert transform followed by processing using the Fourier method. Modern FPGA systems operate with parallel processes and pipelining of computations. A significant number of tests have been conducted using mathematical modeling packages such as Matlab and Maple, and all critical modules of the synthesized algorithm have been debugged. Experimental verification on several specific samples of the WSG (Quartz Resonator) was carried out in real-time using the hardware implementation of the synthesized algorithm based on FPGA technology. These experiments demonstrated that the error in the phase shift of the quadrature signal does not exceed a few percent, allowing for the effective application of the developed digital signal differentiation method for implementing an efficient control loop for the hemispherical quartz resonator of the WSG under various real-world conditions.

V. CONCLUSION

The article presents the implementation of the Hilbert transform through the Fourier transform for the purpose of generating a quadrature signal. An example of a specific implementation for processing the primary signal of a vibrating gyroscope is provided, along with the calculation of the signal amplitude (as one of the parameters in the state vector). The approach proposed in this publication can be utilized in the development of various control systems for vibrating gyroscopes, as well as in any tasks that involve digital signal processing and require the generation of a quadrature signal. It should be noted that the spectral density according to formula (8) is implemented on FPGA, and this methodology is proposed and elaborated in detail in reference [3].

REFERENCES

1. L. Rabiner, B. Gold. Theory and Application of Digital Signal Processing. - M.: Mir, 1978.
2. Klimov D.M., Zhuravlev V.F., Zhbanov Yu.K. Quartz hemispherical resonator (solid-state wave gyro). - M.: Kim L.A., 2017.
3. Chan Van Ngia. Application of complex filters for reducing the peak-to-average power ratio of OFDM-modulated signals and their implementation on FPGAs. "Digital Signal Processing," No. 1, 2018.

Features of Digital WSG Algorithm Implementation on Modern System-on-Chip Platforms

S.E. Perelyaev
A.U. Ishlinsky Institute of Problems in Mechanics,
Russian Academy of Sciences,
Moscow, Russia
Emails: sergei-perelyaev@mail.ru

S.T. Lesnoy
A.U. Ishlinsky Institute of Problems in Mechanics,
Russian Academy of Sciences,
Moscow, Russia
Emails: lesnoist@yandex.ru

Abstract: The paper discusses the principles of designing a control system for a wave solid-state gyroscope (WSG) based on a hemispherical quartz resonator, with widest possible use of digital technologies. The development and implementation of a promising VTG digital algorithm based on a modern system on a chip — a field-programmable gate array (FPGA) of the seventh Kintex-7 generation is presented. The key implementation features of digital WSG control systems using contemporary FPGA platforms are examined. These systems offer significant advantages over traditional ASIC solutions.

A description of the digital control system is provided. The results of tests on an automated test bench of the BTG industrial design of the navigation accuracy class are presented, fully confirming the operability of the developed digital control system.

Keywords – wave solid-state gyroscope, WSG signal processing channels, digital control system, digital signal processing, field-programmable gate array (FPGA).

I. INTRODUCTION

The evolution of microelectronics and digital device technologies enables the realization of increasingly complex digital control systems. Applied to WSG, this allows for reduced physical dimensions while simultaneously improving measurement accuracy and minimizing susceptibility to external disturbances.

In modern digital control systems, there is a growing trend toward using system-on-chip (SoC) architectures, especially those based on field-programmable gate arrays (FPGA) which can gradually displace application-specific integrated circuits (ASICs) from the market. Highly integrated SOC systems of modern FPGA of the Kintex-7 series allow developers to provide excellent bandwidth in processing primary analog signals of WSG with the possibility of digital programming while meeting the requirements for cost and energy consumption. The unprecedented power of the digital signal processor (DSP) makes universal FPGAs of the Kintex-7 type an excellent implementation option for a wide variety of applications. The Kintex-7 series FPGAs provide peak sequential throughput of 800 Gbit/s and include CPRI/OBSAI IP cores (9.8 Gbit/s) optimized for modern distributed architectures (baseband). Kintex-7 programmable devices can be reconfigured quite easily.

II. BLOCK DIAGRAM OF A SYSTEM ON A CHIP OF A GYROSCOPE CONTROLLER MODULE

The block diagram of the VTG digital primary analog signal processing module includes the following main components: an ADC, the core is a system-on-chip (SoC), and a DAC. High-speed closed-loop ADCs and DACs are permanently connected to the WSG sensor and provide continuous conversion of the primary analog cosine and sine WSG signals into digital form. The Kintex-7 on-chip system processes the output codes of two ADCs with an LVDS interface and generates control signal codes for high-speed DACs. ADC controllers control the serial LVDS interface of the ADC and read data at a rate of about 3 MSPS (millions of samples per second). The data from the two ADCs is stored in a buffer (buf3) for 3 periods of the sensor reference frequency. Such an array of data is necessary for calculating averages and derivatives of the components of increments of physical quantities. The ADC controller provides two-way exchange over the LVDS serial interface with a 16-bit AD7626 type ADC. The serial LVDS interface provides high throughput transmission of ADC output codes with small overall characteristics of the chips. The performance of the Kintex-7 FPGA and the MCG-2 module should provide parallel exchange over six LVDS interfaces of six ADC chips. To correct the amplitude in the fast closed loop of the MCG-2 module, a digital signal processor dsp_K is used, and structural devices with PhaseX,Y are used for phase correction. To calculate the coefficients of a closed loop and calculate information about angle increments, a Micro-Blaze software processor core has been implemented based on the selected on-chip system. A processor device PD based on the Micro-Blaze (uBlaze) software processor core from Xilinx IP-Core. The developed PD should perform the following functions of a top-level SoC processor: controlling the reception of output codes from ADC controllers; exchanging data with DSP coprocessors for gyroscope parameter tracking circuits; issuing output parameters for generating amplitude-modulated voltages Xin,Yin of a 64-bit floating-point calculator for calculating gyroscope output parameters; two-way exchange via a serial interface UART with an external navigation calculator or computer. The uBlaze is 32-bit processor with floating-point. It calculates the angle increments at a rate that provides the required refresh rate for the module's output parameters when output is sent to the top-level calculator. Using a separate serial interface controller with UART, the output parameters of the module are transmitted to the asynchronous RS-422 serial interface. The circuit's output DAC control circuit contains multiplexers and digital DDS frequency synthesizers. Immediately after powering the module, DDS is connected to one of the DACS. It "accelerates" the sensor at a frequency close to the rated value of this device. By iterating through frequencies in increments of up to 0.05 Hz, the uBlaze processor core controls the DDS and selects the optimal reference frequency to be used later for phase correction. After that, the DACs of the sine and cosine channels are connected to the closed gyroscope control loop with automatic correction of physical quantities.

In the previous version of the Artix-7-based MCG-1 digital signal processing module, a DAC8822 dual-channel multiplying 16-bit bipolar DAC was used to generate control signals with a setting time of two output signals of 1 microsecond. With the main resonant frequency of the

gyroscope of 5000 Hz and a sampling rate of 1 microsecond, the frequency resolution was more than 12.5 Hz. The search for a resonant frequency was problematic because of this. The reference voltage of the DAC in the MCG-1 module set a bipolar amplitude quantization scale and was constant. The instantaneous value of the generated control signal at the digital 15-bit input of the DAC contained both amplitude and frequency components. In fact, the update rate of the DAC was determined by the duration of the cycle of processing input and output signals in the Micro-Blaze module. In this case, it is not possible to build a parallel tracking circuit for the reference frequency on the PLL. The electronic module is designed to operate in closed control circuits of a wave solid-state gyroscope (WSG) and must provide calculations and transmission of information about angle increments via the RS-422 serial interface to an external computer or navigation computer. For the operation of the fully digital PLL of the new MCG-2 module, fine-tuning the resonant frequency of the gyroscope, the minimum possible sampling step of the control signals generated from the output of high-speed DACs is required. The MCG-2 module implements a step of updating the carrier frequency of each control signal in a time of no more than 50 ns.

For this purpose, when implementing the digital DDS module, fully digitally controlled generators (NCO based on the new Kintex-7 series FPGA) and a high-speed dual-channel multiplying bipolar DAC of the AD5447 type were used. As a result, the frequency resolution is more than 0.68 Hz at the main resonant frequency of operation of the VTG gyroscope of 5000 ... 6000 Hz. The reference voltage for amplitude quantization of the multiplying 12-bit DAC is also generated on a new 16-bit dual-channel DAC of the AD5689R type. The time to set the output voltage of the new precision DAC is no more than 8 microseconds. The refresh rate of the two amplitude and quadrature tracking circuits, including PID_E, PID_Q, a 2Ω electric angle meter based on the new Kintex-7 FPGA programmable logic, as well as a control signal generator implemented on Micro-Blaze, is at least 1250 Hz required by the customer.

Regarding the Kintex-7 series FPGA system specifically selected for the implementation of the MCG-2 module, this is the seventh generation FPGA family, which allows achieving maximum price/performance. At the same time, the Kintex-7 SOC system itself has significantly better parameters.

III. IMPLEMENTATION OF THE CONTROL ALGORITHM

The system on a chip (SOC) of the WSG controller module of the MKG-2 type was developed on the basis of the modern FPGA of the seventh Kintex-7 series from Xilinx. The main difference between the MCG-2 conversion method and the previous version of the MCG-1 is the presence in the MCG-2 of three parallel high-speed feedback loops for tracking phase, amplitude, and quadrature. In the MCG-1 module, calculations of all parameters were based on a relatively slow top-level serial processor Micro-Blaze FPGA Artix-7, which has 2-3 times lower performance than the Kintex-7 type Micro-Blaze FPGA used in the MCG-2 digital module. In the MCG-2, the oscillations of the WSG resonator are triggered and maintained at a resonant frequency by a phase-locked loop system (PLL), which includes synchronous demodulators of WSG signals, digitized ADCs, numerically controlled oscillators (NCO) and high-speed multiplying DACs. After finding a resonance, a constant oscillation amplitude is maintained using the full energy PID controllers and the resonator oscillation quadrature. Without this, temperature drift leads to degradation of the accuracy characteristics of all instrument measurements. In the system-on-chip (SOC) of the MCG-2 module, three-loop feedback is closed at the lower level of DSP cells, VHDL modules, and HLS IP cores. The calculation of the full response of the PID controller at the DSP level of the Kintex-7 FPGA system takes 2-3 microseconds. This roughly corresponds to the duration of one multiplication of a 32-bit floating-point number of the MicroBlaze FPGA Artix-7 in the MCG-1 module. The most important component of all feedback loops is a synchronous demodulator, which performs narrow-band filtering of the envelopes of amplitude-modulated signals from the WSG outputs at the carrier frequency from the output of the NCO module. The demodulation results are used to calculate the phase error, energy, quadrature, and output angle of the WSG from signals with a significant noise component. Low-pass filters (LPF) IIR average 1250 samples of high-speed ADCs over three periods of the carrier components of the WSG signals to isolate the envelopes of the common-mode and quadrature components of the primary signals Sx,Cx,Sy,Cy. Calculations for PID controllers are also performed at the DSP module level in IP cores using the formulas given below. The variables E,Q,R,S,L of the WSG parameter vector are calculated after synchronous demodulation of the common-mode and quadrature components Sx,Cx,Sy, and Cy of the oscillations, relative to the reference phase, along the X and Y axes of the inertial sensor, based on the formulas:

$$E = Cx^2 + Sx^2 + Cy^2 + Sy^2$$
$$Q = 2(Cx*Sy - Cy*Sx)$$
$$R = Cx^2 + Sx^2 - Cy^2 - Sy^2$$
$$S = 2(Cx*Cy + Sx*Sy)$$
$$L = 2(Cx*Sx + Cy*Sy)$$
$$Angle(\Theta) = 1/4 atan(S/R)$$

The variables E,Q,R,S,L are evaluated and used as input data for two control loops.

The MCG-2 module uses high-speed digital outputs with sampling T=50ns of the input signal carrier for the WSG. The response time of the digital PLL to a change in the phase shift between the carrier of the WSG output signals and the carrier from the output of the NCO reference digital generator does not exceed 640 microseconds.

After the appearance of a sign of readiness for the formation of the reaction of the PID controllers over the last three periods of the WSG output carrier, Micro-Blaze reads the parameters from the PID controllers and generates input effects for WSG from them.:

$$Xin = PID_E*\cos(2\Theta)*Cos(wt) - PID_Q*\sin(2\Theta)*Sin(wt)$$
$$Yin = PID_E*\sin(2\Theta)*Cos(wt) + PID_Q*\cos(2\Theta)*Sin(wt)$$

Next, MicroBlaze outputs a data packet to a serial communication line with a computer or product processor. Thus, the operation cycle of the MicroBlaze top-level processor ends on the 4th period of the carrier frequency. At a carrier frequency of 5 kHz, the cycle duration is 800 microseconds. MicroBlaze in the MCG-2 SoC is used only for receiving PID controller constants and transmitting output parameters via a bidirectional interface with a computer or an external processor. MicroBlaze procedures take less than one carrier period, so MicroBlaze MCG-2 does not add any delay.

When implementing asynchronous demodulation of WSG signals in the MCG-1 module, arrays of ADC samples accumulate over three carrier periods. Unlike the MCG-2 digital synchronous demodulation, which is a simple logical condition, further complex mathematical processing of arrays using time-consuming algorithms begins in MCG-1. For 5-6

whole periods of the MicroBlaze signal carrier, it is possible to isolate the amplitude and phase of the envelope of the input signals with an appropriate delay. Therefore, the fact that about a dozen DSP coprocessors are performed in parallel in the MCG-2 module requires the use of a fast computing core in the MCG-1 module algorithm based on Artix-7 with asynchronous demodulation of primary signals. Several fast parallel computing cores are required to achieve the performance of the MCG-2. Moreover, in both cases (MCG-1 and MCG-2), it is necessary to additionally use a set of logical functional nodes from the FPGA. Another important difference between the MCG-1 module is the absence of PID regulators, which in the MCG—2 provide constant feedback (with appropriate adjustment ranges) to monitor changes in external environmental conditions. Instead of continuous monitoring, the MCG-1 module provides only a single calibration (selection of coefficients) for certain temperature conditions at the stage of joint production of the converter with the gyroscope. If the external operating conditions change to the side not provided for by calibration, then an object with a navigation system based on WSG and MCG-1 will not be able to navigate independently in space. Thus, to process the signals of the WSG sensor, an FPGA with a sufficient number of high-speed logic and DSP cells is needed. In addition, for the operation of the MCG-2 digital PLL module and fine-tuning of the gyroscope's resonant frequency, the minimum possible sampling step of the control signals from the DAC output is urgently needed. The MCG-2 module implements a step of updating the carrier frequency of each control signal in time T<50ns. For this purpose, the new digital DDS module uses digitally controlled generators (NCO in Kintex-7 FPGA) and a high-speed dual-channel multiplying bipolar DAC of the AD5447 type.

The refresh rate of the tracking circuits of the total energy (amplitude) of the resonator oscillations, suppressing the quadrature components of the oscillations, is 1250 Hz. Both WSG control circuits are implemented on the basis of classical proportional–integral–differential (PID) regulators – PID_E, PID_Q, a 2Θ wave angle meter based on FPGA programmable logic, and a Micro-Blaze control signal generator. In the new MCG-2 module for generating a control signal, the type

$$Xin=PID_E*\cos(2\Theta)*Cos(wt)–PID_Q*\sin(2\Theta)*Sin(wt)$$

the PID_E*cos(2Θ) signal must be written to the amplitude quantization DAC AD5689R - putDAC1R, and the required frequency code in the sampling DAC AD5447 - putDDS1F (after finding the main resonant frequency of the quartz resonator WSG can support it automatically when implementing a fully digital PLL system).

IV. CONCLUSION

The accuracy parameters of modern WSGs based on quartz hemispherical resonators are largely determined by measuring and controlling functional electronics. A fully digital WSGs control system operating in the mode of an integrating gyroscope (full angle) has been developed. The system includes a fully digital phase-locked frequency loop (PLL), a loop for stabilizing the total oscillation energy of the resonator, and a quadrature control loop for suppressing the quadrature signal of the inertial sensor. The PLL is used to bind the reference phase to the oscillation phase at any arbitrary azimuth of the wave angle for synchronous modulation and demodulation of the primary and control signals. Meanwhile, energy stabilization and quadrature control circuits are used to compensate for energy dissipation and anisoelasticity effects. Fast signal processing, including modulation, demodulation, and phase-locked loop, is implemented in the hardware of the modern Kintex-7 series FPGA. Laboratory tests of the new MCG-2 digital module have shown that its main parameters meet the requirements of the technical specifications of a potential customer.

REFERENCES

[1] Matveev V.A., Lipatnikov V.I., Alekhin A.V. Development of a Wave Solid-State Gyroscope. Bauman Moscow State Technical University, 1998, 167 pages.

[2] Asadian, M.H., Wang, Y., Shkel, A.M. “Development of 3D Fused Quartz Hemi-Toroidal Shells for High-Q Resonators and Gyroscopes.” IEEE/ASME Journal of Microelectromechanical Systems, 2019, pp. 1380–1383.URL: https://ieeexplore.ieee.org.

[3] W.Zhao, H.Yang, F.Liu, Y.Su1, Ch.Li. High sensitivity rate-integrating hemispherical resonator gyroscope with dead area compensation for damping asymmetry. // Scientific Reports. www.nature.com/scientificreports (2021)

Excitation Mode with Continuous Restoration of the Velocities of Movement of Elastic Deformations of a Single-Axis Solid-State Gyroscope

V.M. Kotlov
FAI State Research Institute of Aviation Systems
Moscow, Russia
e-mail: vadimkot 366@ yandex.ru

S. E. Perelyaev
Ishlinsky Institute for Problems in Mechanics of RAS
Moscow, Russia
e-mail: pers 2030@ yandex.ru

Annotation. **To obtain inertial information, the new physical principle introduced [1] – the inertial properties of elastic waves – is replaced by a more general principle of inertial motion of elastic deformations of the resonator. The development of the theory of a ring wave solid-state gyroscope shown in [2] is considered for the problem of excitation and continuous continuation of resonator oscillations, including the conditions of non-perturbation by the exciting force of the absolute angular velocity of the resonator base and the motion of its elastic deformations.**

Key words: solid-state wave gyroscope (SSWG), angular velocity sensor (AVS), integrating SSWG, elastic deformations of a ring resonator.

I. INTRODUCTION

Previously [1] it was shown that it is possible to construct a wave solid-state gyroscope without its main part - an elastic standing wave. From the equations of its motion it follows that the gyroscopic effect carries elastic deformations of the resonator. In the presented study, the problem of excitation and continuous continuation of oscillations of a single-axis solid-state gyroscope on a ring resonator without an elastic wave is posed and solved. Elastic deformations of the resonator behave like free particles: with a fixed base of the resonator, they move along it with a constant speed relative to the fixed axes, and when the resonator rotates, they are carried along by it. Moving elastic deformations under the action of overloads behave like point solids, which can serve as sensors of inertial information, like ordinary inert solids.

These provisions, on the one hand, allow us to move from the principle of inertial properties of elastic waves to the principle of inertial motion of resonator deformations, on the other hand, make the term "wave" in the name of a solid-state gyroscope optional. The gyroscopic effect can be used in such a formulation both as a meter of the absolute angular velocity of rotation of the resonator body, and as a meter of the integral of the angular velocity.

The excitation mode and continuous continuation of oscillations of the ring resonator of a single-axis solid-state gyroscope should not disturb either the angular velocity of the resonator base or the speed of movement of its elastic deformations.

To accomplish the task, the classical method of V.F. Zhuravlev is used [2], developed by the authors in relation to a SSWG with a thin elastic ring resonator and presented by us in a joint report at the XXX St. Petersburg International Conference on Integrated Navigation Systems.

From the solution of the equations of motion obtained for the kinetic energy of a ring resonator, it follows that elastic deformations move along the resonator at a constant speed relative to fixed axes as a rigid body, without changing their configuration. This phenomenon in continuous media is called a "soliton", or solitary wave [2]. Moving elastic deformations under the action of overloads behave like point solid bodies.

II. STATEMENT OF THE PROBLEM

The aim of this work is to formulate and solve the problem of excitation of oscillations of a single-axis solid-state gyroscope based on a thin ring.

In this case, the condition of non-perturbation of both the angular velocity of the ring resonator base and the velocities of its elastic deformations must be met. The system of dynamic equations of motion of a thin ring resonator uses a new type of elastic waves (solitons). Thus, the term "wave" in the name of a solid-state gyroscope means a mechanical standing elastic wave of solitons.

In the classical monographs [2,3] by D.M. Klimov and V.F. Zhuravlev, the well-known equations of the dynamics of motion of a ring-shaped SSWG were obtained in the form of generally accepted equations of mechanics in partial derivatives for a thin elastic ring.

III. FUNDAMENTALS OF CREATING AN EXCITATION SYSTEM WITH CONTINUOUS RESTORATION OF MOVEMENT SPEEDS.

The system of equations of the dynamics of a ring resonator consists of two independent systems of equations of motion. The first of them, under the condition of constancy of the parameter, φ is equal to

$$\begin{cases} \ddot{r} - \left(2\dot{v}\Omega + v\dot{\Omega} + (r+R)\Omega^2\right) = 0, \\ \ddot{v} - \left(-2\dot{r}\Omega - (r+R)\dot{\Omega} + v\Omega^2\right) = 0. \end{cases} \quad (3.1)$$

where:

$$\ddot{r} = \frac{\partial}{\partial\tau}\dot{r},\ \ \ddot{v} = \frac{\partial}{\partial\tau}\dot{v},\ \ \dot{\Omega} = \frac{\partial}{\partial\tau}\Omega$$

Ω - angular velocity of rotation of the resonator base,

r, v - elastic deformations of the flexible ring [2].

The second is under the condition of constancy of the variable τ:

$$R\frac{\partial^2}{\partial\varphi^2}\,\mathrm{Sin}(Q) = 0 \quad (3.2)$$

The solution of systems of equations of the form (3.1), (3.2) represents the movement of resonator deformations:

$$\{r(\tau, \Delta\varphi) + R, v(\tau, \Delta\varphi)\} =$$

$$= \{\Delta Xno + Vxn\,\tau, \Delta Yno + Vyn\,\tau\}\cdot \cdot \begin{bmatrix} \mathrm{Cos}(\Delta\varphi+\psi) & -\mathrm{Sin}(\Delta\varphi+\psi) \\ \mathrm{Sin}(\Delta\varphi+\psi) & \mathrm{Cos}(\Delta\varphi+\psi) \end{bmatrix} + (\Delta\varphi) \quad (3.3)$$

Where:

$$f(\Delta\varphi) = \{a, b\}.\begin{bmatrix} \mathrm{Cos}(\Delta\varphi) & -\mathrm{Sin}(\Delta\varphi) \\ \mathrm{Sin}(\Delta\varphi) & \mathrm{Cos}(\Delta\varphi) \end{bmatrix} + +\{\mathrm{c}(\Delta\varphi), \mathrm{d}(\Delta\varphi)\} \quad (3.4)$$

$$a = R\Big(-\mathrm{Cos}\big(Qo\big) + \epsilon\,\mathrm{Sin}\big(Qo\big)\Big)$$
$$b = R(\mathrm{Cos}(Qo) - \epsilon\,\mathrm{Sin}(Qo)])\mathrm{tg}(Qo),$$
$$\mathrm{c}(\Delta\varphi) = \tfrac{1}{2}R\,\mathrm{Sec}(2Qo)(1 + \epsilon\,\Delta\varphi(1 - -\mathrm{Cos}(2Qo)) - \epsilon\,\mathrm{Sin}(2Qo)),$$
$$d(\Delta\varphi) = R\,\big((-1 + \epsilon\,\Delta\varphi)\mathrm{Cos}(Qo) + \epsilon\,\mathrm{Sin}(Qo)\big)\mathrm{tg}(Qo).$$
$$\Delta\varphi = (\varphi - \varphi o),$$
$$R\,\mathrm{Sin}(Qo) = (R\,\mathrm{Sin}(Q)_{|_{\varphi\to\varphi_o}},$$
$$\big(R\mathrm{Sin}(Qo)\big)_{s_0} = \left(\frac{d}{d\varphi}R\mathrm{Sin}(Q)\right)_{|_{\varphi\to\varphi_o}},$$

Above, the following symbols have been introduced:

$$\epsilon = -\frac{\big(RSin(Qo)\big)s_0}{RSin(Qo)},$$

where: $Qo, \varphi o$ – initial parameter values,

$\{\Delta Xno, \Delta Yno, Vxn, Vyn\}$ are constant values.

Elastic deformations of a ring resonator behave like free particles: when the base of the resonator is stationary, they move along it at a constant speed relative to the stationary axes, and when the resonator rotates, they are carried along by it.

From equations (3.1) for each $\Delta\varphi_i = Const$ we obtain an expression ψ as an integral of the projection of the absolute angular velocity onto the sensitivity axis:

$$\{\{r(\tau,\Delta\varphi) + R, v(\tau,\Delta\varphi)\}\} = = \{\Delta\mathrm{Xno} + \mathrm{Vxn}\cdot\tau, \Delta\mathrm{Yno} + \mathrm{Vyn}\cdot\tau\}\cdot \cdot \begin{bmatrix} \mathrm{Cos}(\Delta\varphi+\psi) & -\mathrm{Sin}(\Delta\varphi+\psi) \\ \mathrm{Sin}(\Delta\varphi+\psi) & \mathrm{Cos}(\Delta\varphi+\psi) \end{bmatrix} \quad (3.5)$$

It is convenient to consider the vector relation (3.5) as the following expression for complex numbers:

$$(r + R) + iv = (\Delta Xn + i\Delta Yn)Exp(-i(\varphi + \psi)) \quad (3.6)$$

where i – imaginary unit.

For $\varphi = const$ and with a known value $\psi(\tau)$

$$\psi(\tau) = \psi_0 + \int_0^\tau \Omega_0(\tau_1)d\,\tau_1 \quad (3.7)$$

We get that the expression:

$$Z(\tau) = ((r + R) + iv)Exp(-i(\varphi + \psi)) = =(\Delta Xn + i\Delta Yn) \quad (3.8)$$

The right side of which is a complex quantity that changes linearly over time $Z(\tau)$:

$$\Delta Xn + i\Delta Yn = \Delta Xno + i\Delta Yno + (Vxn + iVyn)\,\tau \quad (3.9)$$

where, its initial value $\Delta Xno + i\Delta Yno$ and the coefficient $(Vxn + iVyn)$ can be obtained from numerical processing $Z(\tau)$by the least squares method. Knowing the value of the quantity

$$\Delta Xn + i\Delta Yn$$

and assuming constancy of quantity

$$\Delta Xno + i\Delta Yno$$

and the coefficient

$$Vxn + iVyn$$

throughout the entire time period of the excitation system, we find from the following relationship:

$$(r + R) + iv = (\Delta Xn + i\Delta Yn)Exp(-i(\varphi + \psi))$$

Size $Exp(-i(\varphi + \psi))$ equal to the expression

$$Exp\big(-i(\varphi + \psi)\big) = ((r + R) + iv)/(\Delta Xn + i\Delta Yn) \quad (3.10)$$

When replacing $\{r + R, v\}$ in system (3.1) with variables $\{x, y\}$:

$$\begin{cases} x = Exp(-\eta\,\omega_0\,\tau)\big((r + R)Cos[\omega_0\,\tau] - v\,Sin(\,\omega_0\,\tau)\big) \\ y = Exp(-\eta\,\omega_0\,\tau)(v\,Cos(\,\omega_0\,\tau) + (r + R)Sin(\,\omega_0\,\tau)) \end{cases}$$

We obtain that the equations of the dynamics of the VTG motion are similar to the equations of D. Lynch:

$$\begin{cases} \ddot{x} - \big(-2\eta\omega_0\dot{x} + \big((1-\eta^2)\omega_0{}^2 - \Omega(2\omega_0 - \Omega)\big)x\big) = \\ \quad = -\,2(\omega_0 - \Omega)\dot{y} + \big(\dot{\Omega} - 2\,\eta(\omega_0 - \Omega)\,\omega_0\big)y \\ \ddot{y} - \big(-2\eta\omega_0\dot{y} + \big((1-\eta^2)\omega_0{}^2 - \Omega(2\omega_0 - \Omega)\big)y\big) = \\ \quad = +2(\omega_0 - \Omega)\dot{x} - \big(\dot{\Omega} - 2\,\eta(\omega_0 - \Omega)\,\omega_0\big)x \end{cases} \quad (3.11)$$

IV. Ring resonator excitation systems with continuous restoration of elastic deformation velocities.

The system of alignment and calibration of a single-axis solid-state gyroscope on a fixed base includes two modes. In the first of them (the alignment mode), we find the velocities of the resonator's elastic deformations using the absolute angular velocity of the base's rotation taken from the external system (3.8), (3.9). Then, after exciting the resonator's oscillations, we switch to the second mode (the flight mode). In this case, the excitation mode of the resonator should preserve both the angular velocity of the resonator's base's rotation and the velocities of its elastic deformations. In the second mode, knowledge of the velocity of the resonator's elastic deformations makes it possible to determine the absolute velocity of the resonator's base's rotation, even if it has changed during the transition from the first mode to the second.

The ring resonator oscillation excitation system changes the following variables:

$$\{\dot{r}, \dot{v}, r, v, \varphi, \Omega\} \quad (4.1)$$

Let us compose a system for exciting the oscillations of a ring resonator so that it does not disturb either the angular velocity of rotation of the resonator base or the velocity of its elastic deformations using the approach of V.F. Zhuravlev [4]:

$$\{Vxn, Vyn, \Delta Xn = \Delta Xno + Vxn\cdot\tau, \Delta Yn = \Delta Yno + Vyn\cdot\tau, \Omega\} \quad (4.2)$$

It should use the following operator in external differential form:

$$\mathrm{M[d[]]} = \frac{\mathrm{d[Vxn]}\wedge\mathrm{d[Vyn]}\wedge\mathrm{d[\Delta Xn]}\wedge\mathrm{d[\Delta Yn]}\wedge\mathrm{d[\psi]}\wedge\mathrm{d[]}\wedge\mathrm{d[\Omega]}}{\mathrm{d[r1]}\wedge\mathrm{d[v1]}\wedge\mathrm{d[r]}\wedge\mathrm{d[v]}\wedge\mathrm{d[\psi]}\wedge\mathrm{d[\varphi]}\wedge\mathrm{d[\Omega]}} \quad (4.3)$$

where $\wedge$ is the outer multiplication of differentials [5].

The action of this operator does not change the following parameters of the dynamics of the VTG movement:

$$M\left[d[Vxn]\right] = 0,\ \ M\left[d[Vyn]\right] = 0$$

$$M\left[d[\Delta Xn]\right] = 0,\ M\left[d[\Delta Yn]\right] = 0,\ M\left[d[\Omega]\right] = 0 \quad (4.4)$$

Moreover, the excitation of oscillations of a single-axis solid-state gyroscope is carried out in accordance with the action of the operator (4.3):

$$M\left[d[\varphi]\right] = 1 \quad (4.5)$$

In the exhibition mode, relying on relations (4.3), (4.4) and knowing the value of Ψ by the measurement time $\{(r+R), v\}$, we obtain the values $\{\Delta Xn, \Delta Yn, Vxn, Vyn\}$ from (3.9).

When switching to flight mode, knowing the values

$$\{\Delta Xn, \Delta Yn, Vxn, Vyn\}$$

it is possible to determine the values Ψ,Ω according to (3.10). In this case, the change in variables (4.1) is carried out in accordance with the expressions:

$$\frac{\partial}{\partial\varphi}\{\dot{r}, \dot{v}, r, v, \varphi, \Omega\} = \{\dot{v}, -\dot{r}, \dot{v}, r, v, -(r+R), 1, 0\} \quad (4.6)$$

When moving to the equations of the dynamics of the VTG motion similar to the equations of D. Lynch (3.11), the oscillation excitation system must use the following operator in the form of an external differential form:

$$\mathrm{M1[d[]]} = \frac{\mathrm{d[Vx\ \ ]\wedge\ d[Vy\ \ ]\wedge d[}\Delta Xn\mathrm{]\wedge\ d[}\Delta Yn\mathrm{]\wedge d[]\wedge d[}\Omega\mathrm{]}}{\mathrm{d[x1\wedge\ d[}y1\mathrm{]\wedge\ d[}x\mathrm{]\wedge d[}y\mathrm{]\wedge d[}\varphi\mathrm{]\wedge d[}\Omega\mathrm{]}} \quad (4.7)$$

a, the change of variables $\{x1, y1, x, y, \varphi, \Omega\}$, is carried out in accordance with the expressions:

$$\frac{\partial}{\partial\varphi}\{x1, y1, x, y, \varphi, \Omega\} = \\ = Exp(4\eta\omega_0 t)\{y1, -x1, y, -x, 1, 0\} \quad (4.8)$$

CONCLUSIONS

The new results of the theory of the uniaxial ring resonator of the SSWG include:

a) proof of the integrability of the kinematic condition of non-stretchability of the center line of a ring resonator and simplification of the approach to deriving the equations of the dynamics of motion of a uniaxial inertial sensor;

b) the holonomy of the condition of non-stretchability of the middle line of the ring resonator, which makes it possible to obtain the equations of the dynamics of its motion using the symbolic mathematics program package for engineering calculations " Wolfram" Research Mathematica » 5.2.0.

c) obtaining the equations of the dynamics of the motion of a uniaxial annular SSWG as three ordinary differential equations of the second order, independent of each other in time and angle, which allows for the practical use of ordinary differential equations of the Lynch-Zhuravlev type;

d) the inertial information that determines the gyroscopic effect depends on the inertial motion of the ring resonator deformations, which provides a justification for the inertial properties of standing waves;

e) the basics are formulated excitation systems with continuous restoration of the velocities of movement of elastic deformations of a ring resonator for the development of a system for the alignment and calibration of a single-axis solid-state gyroscope on a fixed base.

From the differential equations of the dynamics of the motion of a ring resonator obtained in this study it follows:

- the possibility of developing a new inertial angle sensor (integrating gyroscope) and angular velocity sensor (AVS) without exciting mechanical standing waves of elastic oscillations of the resonator edge;

- obtaining a procedure for excitation and continuous extension of solitary mechanical waves (solitons) as bending oscillations of a ring resonator, for developing a system for setting and calibrating a single-axis solid-state gyroscope on a fixed base.

REFERENCES

[1] Kotlov V.M., Perelyaev S.E. Derivation of equations of the dynamics of motion of a ring wave solid- state and pendulum gyroscopes by the method of holonomic mechanics. // XXX Anniversary International Conference on Integrated Navigation Systems. collection of materials conferences. Saint Petersburg, 2023. pp. 395-398.

[2] Zhuravlev V.F., Klimov D.M. Wave solid-state gyroscope. - M.: Nauka, 1985.

[3] Klimov D.M., Zhuravlev V.F., Zhbanov Yu.K. Quartz hemispherical resonator (solid-state wave gyroscope). - M.: Kim L.A., 2017.

[4] Zhuravlev V.F., On control of the oscillation form in resonant systems. // PPM, Vol. 56, 5, 1992, pp. 827–836.

[5] Dobronravov V.V. Fundamentals of Analytical Mechanics. - M: Higher School, 1976.

Increasing the Accuracy of a Micromechanical Gyroscope with a Ring Resonator in a Multimode Operating Mode

A.A. Maslov

dept. of Higher Mathematics
National Research University "MPEI"
Moscow, Russian Federation
maslov954@yandex.ru

I.G. Ninalalov

dept. of Robotics
National Research University "MPEI" Moscow, Russian Federation

ninalalov98@mail.ru

D.A. Maslov

dept. of Higher Mathematics
National Research University "MPEI"
Moscow, Russian Federation
dm_93@live.ru

I.V. Merkuryev

dept. of Robotics
National Research University "MPEI"
Moscow, Russian Federation
merkuryeviv@ya.ru

***Abstract*—We consider the dynamics of a micromechanical gyroscope with ring resonator in a multimode oscillation mode under external vibration and impact effects. To construct a mathematical model of the gyroscope resonator motion, the Hamilton-Ostrogradsky variational principle and the Bubnov-Galerkin method are used. The resulting differential equations are used to calculate the gyroscope errors caused by external effects and nonlinear properties of the resonator oscillations magnetoelectric control circuit.**

***Keywords*—micromechanical gyroscope, ring resonator, multimode oscillation mode**

I. Introduction

The relevance of the research topic is related to the need to improve the accuracy and stability of the metrological characteristics of micromechanical gyroscopes used in navigation and motion control of highly maneuverable objects [1,2]. One of the promising areas of development of micromechanical gyroscopes is the creation of precision inertial sensors with improved and stable metrological characteristics and with an automatically adjustable measurement range [2-5]. It is known that to build a high-precision micromechanical gyroscope, it is necessary to ensure not only the perfection of the gyroscope resonator manufacturing technology, but also to build a precision electronic control circuit and adjust the oscillation modes [6-8]. A feature of gyroscopes functioning as part of the orientation and navigation moving objects system is a significant level of accidental and shock external actions. At the same time, the operating conditions of the system change significantly, for example, with changing the ambient temperature. This work aims to develop a new mathematical model of a micromechanical gyroscope for a multimode oscillation mode of a ring single-crystal resonator on a moving base under intense random and impact external influences.

II. Equations of Small Oscillations of a Resonator

The micromechanical gyroscope belongs to the class of the generalized Foucault pendulum gyroscopes [2]. The principle of gyroscope operation is based on the elastic waves inertia effect of thin elastic resonator oscillations. Resonator is supported by a system of thin elastic torsion bars. By using plasma chemical etching technology, a thin elastic ring resonator 2 and an elastic suspension 3 are formed from a single crystal silicon plate 1 (Fig.1). Control and measurement of small vibrations of the ring resonator is carried out using a system of magnetoelectric sensors. Control and measurement circuits are formed by two nearly elastic supporting torsion bars and a corresponding section of the ring resonator (Fig.2) [9].

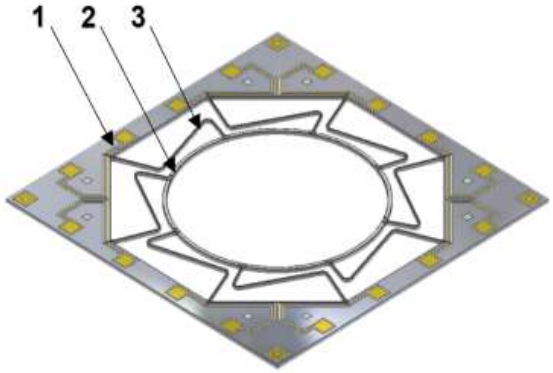


Fig. 1. Micromechanical vibration gyroscope

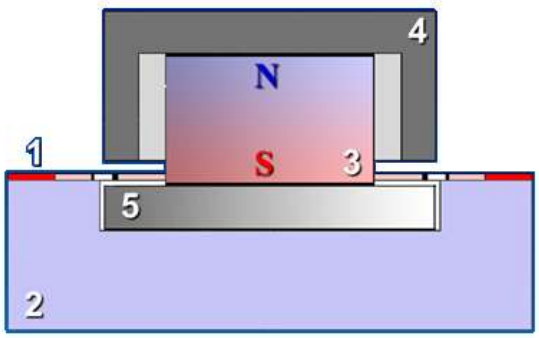


Fig. 2. Magnetoelectric control and measurement system

The permanent magnet 3 and the magnetic cores 4, 5 have to provide a constant magnetic field in the area of the ring

The research was supported by the Russian Science Foundation grant No. 25-21-00436.

resonator. When the conductive circuit moves in the constant magnetic field, the magnitude of the magnetic induction flux changes through the surface bounded by the considered electrically conductive circuit.

On the one hand, the phenomenon of electromagnetic induction is used to measure resonator vibrations, and on the other hand, it is used to excite and control vibrations by applying a control voltage to the contact pads of an electrically conductive circuit. The possibility of excitation and measurement of resonator oscillations in a multimode oscillation mode is discussed. A special feature of the control and measurement system is a small change in the magnetic field in the gap between the magnetic core and the ring resonator caused by the edge effect of the permanent magnet. A violation of the magnetic field symmetry can affect the accuracy and stability of the measurement characteristics of the gyroscope.

By using the Hamilton-Ostrogradsky variational method partial differential equations describing small oscillations of a gyroscope resonator on moving base are obtained:

$$\begin{aligned}
&\ddot{v}+2\Omega\dot{w}-\left(\dot{V}_1-\Omega V_2\right)\sin\varphi+\left(\dot{V}_2+\Omega V_1\right)\cos\varphi-\\
&-\delta^2\left(v''-w'\right)+\kappa^2\left(w^{III}-v''+\zeta v\right)+e_*\kappa^2\left(\dot{w}^{III}-\dot{v}''\right)=0,\\
&\ddot{w}-2\Omega\dot{v}+\left(\dot{V}_2+\Omega V_1\right)\sin\varphi+\left(\dot{V}_1-\Omega V_2\right)\cos\varphi-\\
&-\delta^2\left(v'-w\right)+\kappa^2\left(w^{IV}-v^{III}+\zeta w\right)+\\
&e_*\kappa^2\left(\dot{w}^{IV}-\dot{v}^{III}\right)=-\frac{q(t,\varphi)}{\rho S(1+\mu_0)},
\end{aligned}\qquad (1)$$

where $v(t,\varphi), w(t,\varphi)$ are the elastic displacements of the resonator in the circumferential and radial directions in the axes y_1y_2, connected with the resonator (Fig. 3). The dot denotes differentiation with respect to time t, the prime denotes the partial derivative with respect to angle φ.

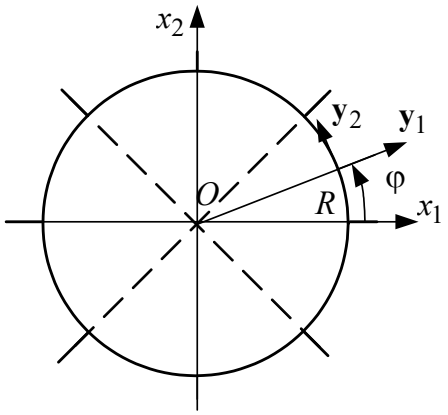


Fig. 3. Resonator calculation scheme

The boundary conditions in problem (1) are determined by the periodicity conditions. When obtaining differential equations (1), it is assumed that the base moves with a slowly changing angular velocity $\Omega(t)$ around the axis x_3 of the coordinate system $Ox_1x_2x_3$ connected with the gyroscope base. The time-varying velocity of the point O – the pole of the resonator in projections $V_1(t), V_2(t)$ onto the corresponding axes x_1, x_2 is also taken into account. In (1), the parameters are introduced

$$\kappa^2=\frac{EI}{\rho SR^4(1+\mu_0)},\quad \zeta\kappa^2=\frac{c_0}{\rho SR(1+\mu_0)},\quad \delta^2=\frac{12R^2}{h^2}\kappa^2,$$

characterizing the elastic properties of the system and corresponding the dimension of the square of the frequency; E is the Young's modulus in the calculation plane (111) of the silicon single crystal; $I=bh^3/12$ is the moment of inertia of the transverse rectangular section of the resonator; b is the height of the cross section; h is the thickness of the resonator; ρ, R are, respectively, the density of the material and the radius of the resonator; $S=bh$ is the cross-sectional area of the resonator; μ_0 and ζ are the relative added mass and the stiffness of the elastic suspension; e_* is the damping coefficient of the resonator oscillations; $q(t,\varphi)$ is the specific external force of the magnetoelectric system for controlling the resonator oscillations.

To solve the system of partial differential equations (1) describing the oscillations of a ring resonator in an elastic suspension on a movable base, the Bubnov-Galerkin method was used for the multimode approximation of elastic oscillations:

$$v=\sum_{k=1}^{N}\left(r_k(t)\cos k\varphi+s_k(t)\sin k\varphi\right),$$
$$w=\sum_{k=1}^{N}\left(f_k(t)\cos k\varphi+g_k(t)\sin k\varphi\right),$$

here $\mathbf{x}_k=\left(f_k,g_k,r_k,s_k\right)^T$ are the generalized coordinates describing the oscillations of the ring resonator in oscillation modes $(k=1,2,3,...)$, N is the number of oscillation modes.

The first oscillation resonator mode characterizes the displacement of the ring resonator as a whole in the direction of the axes x_1 and x_2. Figure 4 shows two forms of ring resonator oscillations in the second basic mode. In addition to these oscillation forms, there are oscillations in the circumferential direction, which can also be used to estimate the angular motion of the gyroscope base.

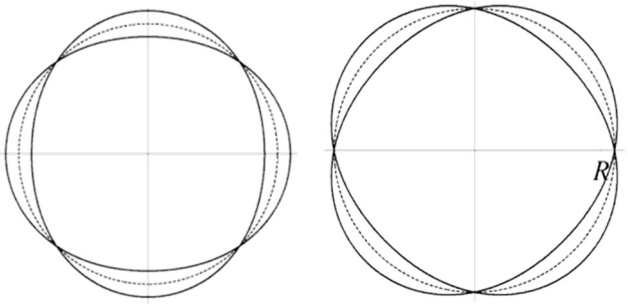


Fig. 4. Basic normal modes of vibration

In the calculated case of an axisymmetric MMG resonator, a system of ordinary differential equations describing the multimode oscillations of the resonator is obtained:

$$\ddot{\mathbf{x}}_1 + C_1\,\mathbf{x}_1 = -\left(\Omega J_1 + D_1\right)\dot{\mathbf{x}}_1 + B_1\mathbf{u} + \mathbf{e}_1(\mathbf{x}_1,\mathbf{x}_2) + \mathbf{w},$$
$$\ddot{\mathbf{x}}_2 + C_2\,\mathbf{x}_2 = -\left(\Omega J_2 + D_2\right)\dot{\mathbf{x}}_2 + B_2\mathbf{u} + \mathbf{e}_2(\mathbf{x}_1,\mathbf{x}_2),\ldots \quad (2)$$

where J_1, J_2 are block skew-symmetric matrices of gyroscopic forces; C_k and D_k $(k=1,2,\ldots)$ are matrices of positional and dissipative forces; B_1, B_2 are matrices of the influence of control signals $\mathbf{u}(t)$, used to excite and control oscillations according to a given oscillation shape of the resonator $\mathbf{w} = \left(-\dot{V}_1 + \Omega V_2, -\dot{V}_2 - \Omega V_1, -\dot{V}_2 - \Omega V_1, \dot{V}_1 - \Omega V_2\right)^T$ is the vector of disturbing external effects caused by vibrations and impacts of the base; $\mathbf{e}_1(\mathbf{x}_1,\mathbf{x}_2)$, $\mathbf{e}_2(\mathbf{x}_1,\mathbf{x}_2)$ are small nonlinear terms determined by edge effects in the magnetic field of the permanent magnet.

Let's pay attention that shocks and random vibrations of the base affect the dynamics of a balanced resonator only in the first form of vibrations. Instrumental errors in the manufacture of the gyroscope lead to a violation of the symmetry of the resonator and to the connections between the oscillation modes [10, 11].

We note, if the elastic system moves along the first mode, the ring resonator moves as a whole shift. Impacts and random vibrations of the base affect to the dynamics of the balanced resonator only along the first mode of oscillations. Instrumental errors in the manufacture of the gyroscope lead to a violation of the resonator symmetry and to connections between the oscillation modes.

The exploitation an elastic suspension for the ring resonator allows to use the first mode of resonator oscillations for determining the angular and linear motion of the resonator, as well as the second and third modes for duplicating the inertial information about the angular motion of the gyroscope base.

In addition to the analytical model of the resonator motion, a digital model of the resonator is built in the Fidesis engineering design system (Fig. 5). The comparison of analytical and numerical solutions of two mathematical models of a micromechanical gyroscope on a moving base is carried out. The constructed finite element model is used to estimate the parameters of the mathematical model of the resonator (1).

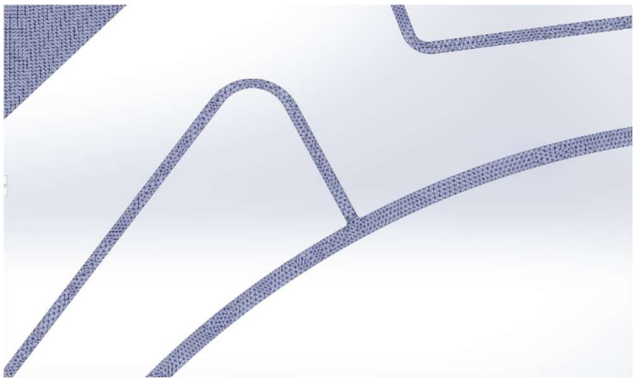


Fig. 5. Calculated finite element scheme

The deduced finite element model is used to evaluate the adequacy of the developed analytical model of a ring resonator and to estimate the parameters of the resonator mathematical model (1) including the reduced mass and stiffness of the elastic suspension.

III. Investigation of Nonlinear Dynamics of a Ring Resonator

To study the nonlinear dynamics of the MMG, the hypothesis of the inextensibility of the resonator midline is applied. The normalized differential equations that describe the bending resonator oscillations according to the second fundamental form are obtained:

$$\ddot{f} + f = \varepsilon\left(-2\gamma\dot{f} + 2\nu\dot{g} + \frac{3}{4}\alpha_2 u_1(\tau) f^2 + \right.$$
$$\left. + \frac{1}{2}\alpha_2 u_2(\tau) fg + \frac{1}{4}\alpha_2 u_1(\tau) g^2 - u_1(\tau)\right),$$
$$\ddot{g} + g = \varepsilon\left(-2\gamma\dot{g} - 2\nu\dot{f} + \frac{3}{4}\alpha_2 u_2(\tau) g^2 + \right. \quad (3)$$
$$\left. + \frac{1}{2}\alpha_2 u_1(\tau) fg + \frac{1}{4}\alpha_2 u_2(\tau) f^2 - u_2(\tau)\right),$$

here γ is the damping coefficient; ν is the normalized angular velocity of the gyroscope base; α_2 is the parameter characterizing the small change in the modulus of the magnetic induction vector in the gap between the magnetic circuit and the ring resonator; $u_1(\tau)$, $u_2(\tau)$ is the control action for exciting two normal modes of resonator oscillations characterized by normalized generalized coordinates $f(\tau)$ and $g(\tau)$; the dot denotes differentiation with respect to the dimensionless time $\tau = \omega_2 t$; ω_2 is the characteristic frequency of resonator oscillations according to the fundamental mode of oscillations; ε is a small parameter characterizing the smallness of damping, angular motion, and control actions. The nature of the nonlinearity is caused by small change in the modulus of the magnetic induction vector in the gap between the ring resonator and the magnetic circuit (Fig. 6).

The nature of the nonlinearity is caused by small changing in the modulus of the magnetic induction vector in the gap between the ring resonator and the magnetic core in the direction from the external normal to the resonator (Fig.6). The dots in Fig. 6 show the results of measurements of the modulus of the magnetic induction vector in the direction of the external normal to the ring resonator. The solid line in Figure 6 shows a polynomial approximation that sufficiently accurate shows the nature of magnetic field changing near the resonator. The main contribution is made by the small quadratic dependence of the magnetic field on the normalized displacement of the resonator in the radial direction.

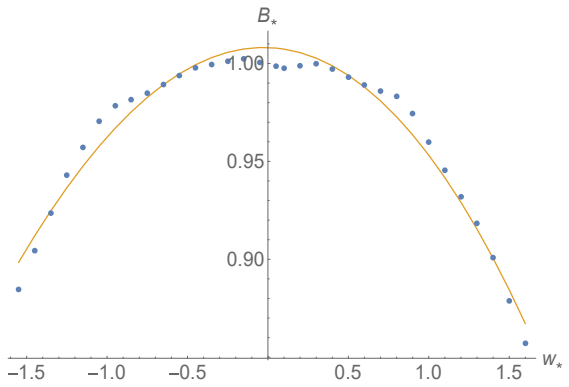


Fig. 6. The changing in the modulus of the magnetic induction vector in the direction of the external normal to the ring resonator

We note, that nonlinearities are fundamental errors of all without exception gyroscopes of the generalized Foucault pendulum class. Equations (3) are nonlinear nonstationary Mathieu–Hill equations [2].

To excite the resonator oscillations according to the second fundamental oscillation mode we select the control voltages in the form, where is the frequency detuning, is the normalized frequency detuning taking a value of the unity order. System (3) is investigated by the Krylov–Bogolyubov averaging method [2]. The differential equations for the measured slow variables p_1, q_1, p_2, q_2 have the form

$$\begin{aligned}
\dot{q}_1 &= \varepsilon\left(-\lambda p_1 - \gamma q_1 + \nu q_2 + \frac{1}{16}\alpha_2\left(3p_1q_1 + p_2q_2\right)\right),\\
\dot{p}_1 &= \varepsilon\left(\frac{1}{2} - \gamma p_1 + \nu p_2 + \lambda q_1 + \right.\\
&\left. + \frac{1}{32}\alpha_2\left(-3p_1^2 - p_2^2 - 3\left(3q_1^2 + q_2^2\right)\right)\right), \qquad (4)\\
\dot{q}_2 &= \varepsilon\left(-\lambda p_2 - \gamma q_2 - \nu q_1 + \frac{1}{16}\alpha_2\left(p_2q_1 + p_1q_2\right)\right),\\
\dot{p}_2 &= \varepsilon\left(-\gamma p_2 - \nu p_1 + \lambda q_2 + \frac{1}{16}\alpha_2\left(-p_1p_2 - 3q_1q_2\right)\right).
\end{aligned}$$

The obtained nonlinear differential equations are studied by using numerical methods. Stationary solutions of system (4) are found, the influence of the mathematical model parameters on the Lyapunov stability of stationary solutions and on the accuracy of gyroscope measurements are investigated.

It is shown that a small change in the magnetic field changing near the ring resonator causes a small deviation from the calculated value of the stationary singular point of the linear system, but under certain system parameters it can cause the phenomenon of oscillation stall and transition to a new stationary mode significantly different from the linear case.

Methods for numerically solving nonlinear nonstationary oscillatory systems in the Fidesis engineering design system based on a calculation calculator and additionally developed external software modules are considered.

IV. Conclusion

The deduced mathematical model of the MMG describes in a single interconnected form the oscillations of an elastic system on a moving vibrating foundation. Numerical and analytical methods that allow dividing the movements of the elastic system into "fast" and "slow" movements and assessing the influence of the mathematical model parameters on the dynamics and accuracy of the MMG are proposed.

Vibrations of the Ring resonator oscillations in the circumferential direction can be used to estimate the angular motion of the gyroscope base in addition to the known multimode methods for determining inertial information [6]. A small change in the magnetic field near the ring resonator causes the phenomenon of oscillation stall and transition to a new stationary mode significantly different from the linear case.

By the analysis of the Lyapunov stability of the nonlinear system near singular points their types are identified and the stability regions in the parameter space of the mathematical model are deduced.

References

[1] Acar C., Shkel A., MEMS Vibratory Gyroscopes Structural Approaches to Improve Robustness. 2009. Springer. 256p.

[2] Zhuravlev V.F. Inertial Navigation System of a New Type // Mechanics of Solids. 2023. Vol. 58. no. 7. pp. 2445-2449.

[3] Martynenko Y. G. , Merkuriev I.V., Podalkov V.V. Dynamics of a ring micromechanical gyroscope in the forced-oscillation mode // Gyroscopy and Navigation 2010, no. 1, pp. 43–51.

[4] Yoon S.W., Lee S., Najafi K. Vibration sensitivity analysis of MEMS vibratory ring gyroscopes // Sensors and Actuators A, 2011, no. 171, pp. 163– 177

[5] Maslov A.A., Maslov D.A., Ninalalov I.G., Merkuryev I.V. Hemispherical Resonator Gyros (An Overview of Publications) // Gyroscopy and Navigation, 2023, vol. 14, no. 1, pp. 1–13. DOI: 10.1134/S2075108723010054.

[6] Perelyaev S.E. Fundamental Questions of the Theory of Combined Free Two-Mode Gyroscopes // Mechanics of Solids, 2021, vol. 56, no. 4, pp. 485-494.

[7] Perelyayev S.E. Theory of New Gyroscopic Sensors of the “Generalized” Foucault Pendulum Series Operating in Two Modes // Mechanics of Solids, 2021, vol. 56, no. 8, pp. 1611-1621.

[8] Younis M.I. MEMS Linear and Nonlinear Statics and Dynamics. Springer. New York. 2011. 453 P.

[9] Распопов В.Я. Raspopov V.Ya. Mikromekhanicheskie pribory. Moskva. Mashinostroenie, 2007. 400 p.

[10] Zhuravlev V.P., Klimov D.M. Spatial Effect of Inertness of Elastic Waves on a Sphere // Mechanics of Solids. 2021. Vol. 56. No 3. PP. 293-295.

[11] Raspopov V.Y., Matveev V.V., Likhosherst V.V., Volchikhin A.I., Shepilov S.I. Design of a Coriolis Vibratory Gyroscope and Orientation and Stabilization System Based on it.// 27th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2020 - Proceedings. 27. 2020. P. 9134029.

Ultra-Compact Magnetic Heading Sensor for UAV

Pazychev D. B.
chief designer
Integral Ltd
Moscow, Russia
d.pazychev@integral-group.pro

Fadeev A. S.
Deputy General Director
JSC "TsENKI"
Moscow, Russia

Gerdi V. N.
Dean of the “PS” faculty
Bauman Moscow State Technical University
Moscow, Russia

***Abstract*-The paper considers a small-sized digital magnetic heading sensor "DMHS-micro" of the company "Integral". The article will describe the design of the sensor, describe the algorithm of its operation, present the procedure for its calibration, and also present the results of testing the sensor on a light UAV, giving an idea of its accuracy.**

Keywords-magnetic heading sensor, magnetometer, navigation system, navigation algorithm, UAV

I. INTRODUCTION

Modern navigation systems used on unmanned aerial vehicles (UAVs) have compact dimensions, low weight and, in most cases, are built on the basis of inexpensive and compact micromechanical sensors (MEMS), which causes their small dimensions and low accuracy. The calculation of the heading channel in such systems using the gyroscopic method is very difficult due to the high instability of the zero offset of the gyroscopes, which directly affects the accuracy of calculating the initial value of the heading angle (1):

$$\delta H = \frac{\omega_E^{dr}}{U \cdot \cos(\varphi)} \qquad (1)$$

where:

δH - error in the initial value of the heading angle

ω_E^{dr} - the given value of the zero gyroscope offset in the projection on the eastern axis

$U \cdot \cos(\varphi)$ - horizontal projection of the angular velocity of the Earth's rotation at latitude φ.

Using formula (1), it is not difficult to calculate that determining the initial value of the heading angle with an error of the order of 1 ° at the latitude of Moscow will require an angular velocity sensor with a zero offset of the order of 0.15 °/h, which makes it impossible to use micromechanical gyroscope with a random zero offset of the order of 15 ... 50 °/h to solve such a problem.

The second possible way to calculate the true heading angle is to use multi-antenna GNSS receivers (Figure 1 and Figure 2), which calculate the relative position of the antennas among themselves.

Such systems have a certain set of disadvantages that significantly limit the range of their application:

- By their principle of operation, such systems work only in conditions of good and stable GNSS signal reception.
- The GNSS requires a certain convergence time during which the true exchange rate signal is not available (Figure 3)
- If the GNSS signal is lost, the transition process takes time to restore an accurate course angle decision



Fig. 1. Small-sized attitude and heading reference system (AHRS) "NV-1-K" of the company "Integral" with two GNSS antennas for calculating the true course



Fig. 2. Radio Complex's Fairway RK-2106 receiver

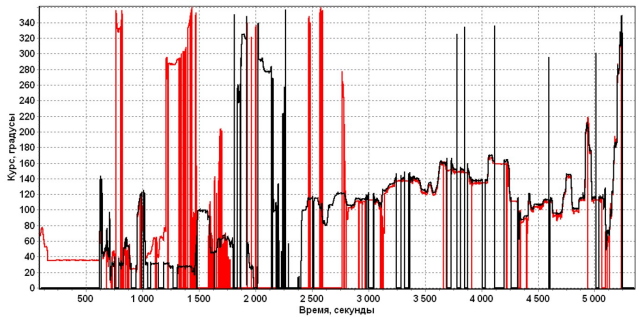


Fig. 3. Behavior of the true heading signal according to the GNSS system (red color) and the GNSS tracking angle (black color) of the NV-1-K AHRS during vehicle tests in Moscow

However, there is another way to calculate the true heading angle – using the magnetic heading signal and magnetic declination (2) (Fig. 4), which, in turn, can be taken from tabular values for these location coordinates.

$$H = H_M + \Delta H_M \qquad (2)$$

where:

H - true course

H_M - magnetic course

ΔH_M - magnetic declination

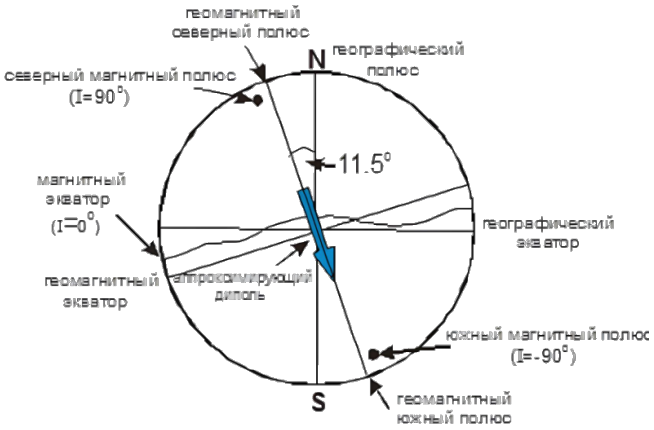


Fig. 4. True and magnetic geographic poles

Thus, the use of an inexpensive and compact digital magnetic heading sensor (DMHS) makes it fundamentally possible to determine the direction to true north, even with the condition of determining and using indirectly measured values. However, this method also has a number of features that you should pay attention to:

The device is sensitive to interference in the form of additional magnetic fields

Due to these features, it is difficult to choose a place to install a DMHS. It is necessary to install the device where the magnetic fields of the object itself have a minimal effect – in the tail or wing of the UAV aircraft scheme.

The magnetic field may change, and the magnetic declination may also be unstable-this depends on the current coordinates of the object.

To calculate the true course of an object, you need to get the exact value of the magnetic declination.

II. Calculation of the Magnetic Exchange Rate

To calculate the magnetic course of an object, it is necessary first of all to know the direction of the horizontal component of the Earth's magnetic field, located in the horizon plane and directed towards magnetic North (Fig.5).

The calculation of this component is carried out using the signals of magnetic field sensors that are part of the DMHS, the axes of which coincide with the construction axes of the object. To convert their readings to the horizon plane, you will need a matrix of guiding cosines that connects the coordinate system associated with the object and the geographical trihedron coordinate system.

To recalculate the readings of magnetic field sensors in the horizon plane, roll and pitch angle signals are used, which are also calculated in the DMHS based on the signals of the gyroscope and accelerometers included in the device. A similar matrix is calculated in strapdown ahrs and strapdown inertial navigation systems (SINS), so the DMHS also includes three small-sized angular velocity sensors (gyroscopes) and three accelerometers. The DMHS operation algorithm, in its essence, is an algorithm for strapdown AHRS based on the SINS algorithm using radial and course correction. The radial correction is intended for correcting the calculation of roll and pitch angles, and the course correction is intended for correcting the course vertically from the magnetic course readings, while the error in calculating the course angle during horizontal flight does not exceed the error of the magnetic course itself.

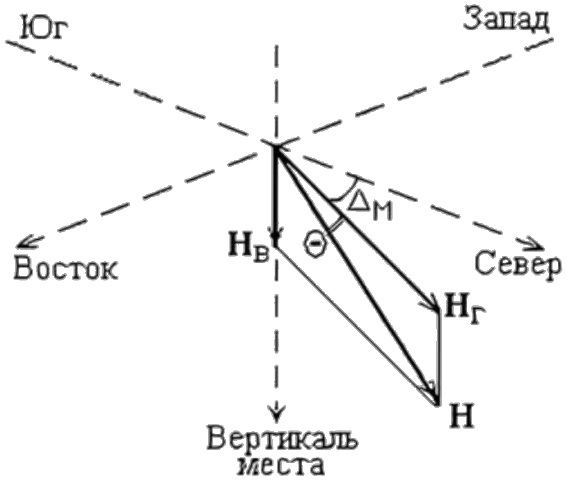


Fig. 5. Horizontal component of the Earth's magnetic field

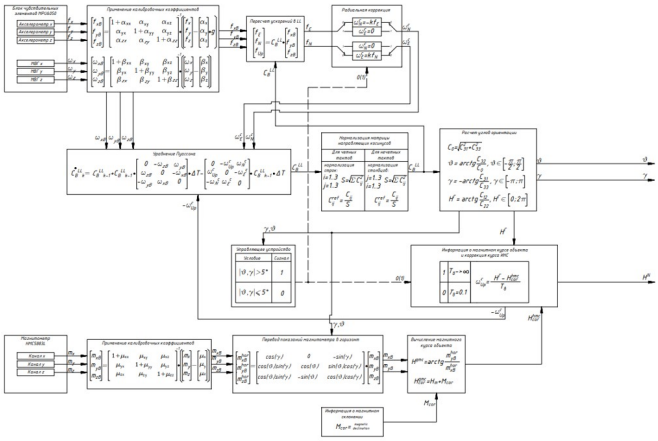


Fig. 6. DMHS operation algorithm

Thus, the calculation of the magnetic course is reduced to the following formula (3).

$$H_M = \arctan(\frac{M_Y}{M_X}) \quad (3)$$

where

M_X, M_Y - horizontal projections of the magnetic field based on the readings of magnetic field sensors.

III. Magnetic Heading Errors

Among the main errors of the magnetic course, at least 4 components should be distinguished (4):

$$\delta H_M = \delta_1 + \delta_2 + \delta_3 + \delta_4 \quad (4)$$

where

δ_1 - a constant error that has the same value for all courses of the aircraft. This constant error is mainly

determined by the inaccurate installation of the DMHS on board the aircraft.

δ_2 - semicircular deviation. It changes periodically when changing course, twice taking the maximum value (in absolute value) when changing course in the range from 0 to 360°. The reason for its appearance is in the magnetic fields created by hard magnetic objects on board the aircraft, as well as in the permanent fields created by electrical equipment.

δ_3 - fourth deviation. It is a periodic function of the course, but it takes the maximum value four times when the course changes in the range from 0 to 360°. It appears as a consequence of fields created by soft magnetic objects. These objects can be magnetized under the influence of the Earth's magnetic field, but since the orientation of such objects relative to the Earth's field changes with a change in course, the nature of the dependence of δ_3 on the course differs from the similar dependence for δ_2. In addition, the quarter deviation cannot be eliminated with the help of a permanent magnet system.

δ_4 - the residual error of the DMHS after the calibration procedure. It is based on the residual calibration errors of the DMHS sensors.

To eliminate permanent, semicircular and quarter deviations, in most cases it is possible to carry out the procedure for writing off magnetic deviations (Fig. 7). This procedure is based on a sequential turn of the aircraft around the vertical axis with points along the course angle of 45 ° (0, 45°, 90°, 135°, 180°, 225°, 270 °, 315°). At each of the points, the magnetic course issued from the device is compared with the true value of the angle. Based on the results of this procedure, a table of correction coefficients of deviation errors of the DMHS is formed, stored in its flash memory.

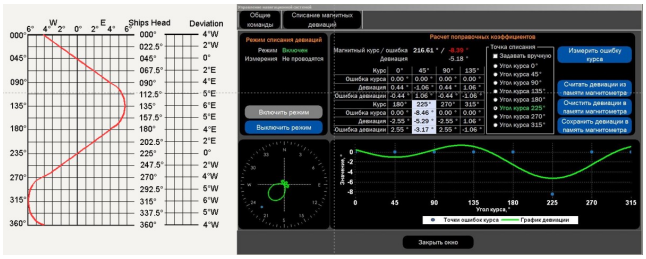


Fig. 7. Free software for writing off magnetic deviation of DMHS-1 and DHMS-micro magnetometers from Integral LLC

Thus after carrying out the procedure for writing off magnetic deviations the error in calculating the DMHS magnetic course is determined by the residual deviation error and the design error of the device δ_4, which is reduced by a special calibration procedure.

IV. DMHS CALIBRATION

Since the DMHS consists of three types of sensors at once, the general error model of the device can be presented in the following form (5).

$$\vec{\omega}_B^{Err} = \begin{bmatrix} \beta_{XX} & \beta_{XY} & \beta_{XZ} \\ \beta_{YX} & \beta_{YY} & \beta_{YZ} \\ \beta_{ZX} & \beta_{ZY} & \beta_{ZZ} \end{bmatrix} \cdot \vec{\omega}_B^{Tr} + \begin{bmatrix} \beta_X(\mathrm{t}) \\ \beta_Y(\mathrm{t}) \\ \beta_Z(\mathrm{t}) \end{bmatrix}$$
$$\vec{a}_B^{Err} = \begin{bmatrix} \alpha_{XX} & \alpha_{XY} & \alpha_{XZ} \\ \alpha_{YX} & \alpha_{YY} & \alpha_{YZ} \\ \alpha_{ZX} & \alpha_{ZY} & \alpha_{ZZ} \end{bmatrix} \cdot \vec{a}_B^{Tr} + \begin{bmatrix} \alpha_X(t) \\ \alpha_Y(t) \\ \alpha_Z(t) \end{bmatrix} \quad (5)$$
$$\vec{M}_B^{Err} = \begin{bmatrix} m_{XX} & m_{XY} & m_{XZ} \\ m_{YX} & m_{YY} & m_{YZ} \\ m_{ZX} & m_{ZY} & m_{ZZ} \end{bmatrix} \cdot \vec{M}_B^{Tr} + \begin{bmatrix} m_X(t) \\ m_Y(t) \\ m_Z(t) \end{bmatrix}$$

where

$\beta_{ii}, \alpha_{ii}, m_{ii}$ - errors in the scale coefficients of gyroscopes, accelerometers, and magnetic field sensors, respectively.

$\beta_{ij}, \alpha_{ij}, m_{ij}$ — errors of non-orthogonality of the gyroscopes axes, accelerometers, and magnetic field sensors, respectively.

$\beta_i(\mathrm{t}), \alpha_i(\mathrm{t}), m_i(\mathrm{t})$ — temperature-dependent zero offset errors of gyroscopes, accelerometers, and magnetic field sensors, respectively.

To compensate for errors in sensor zero offsets, in most cases, temperature compensation is used, which consists in determining the coefficients of the polynomial of the dependence of the sensor readings on the external temperature (6).

$$f_j(T_i) = K_{j,N} \cdot T_i^N + K_{j,N-1} \cdot T_i^{N-1} + \ldots + K_{j,0} \quad (6)$$

In this case, to determine the error coefficients of the error model (6), when the sensor itself is stationary, it is necessary to change only its internal temperature, which can be performed using an inexpensive and compact temperature chamber. To determine the above coefficients, long field tests will be required, during which the temperature of the sensor must change between the lowest and the highest value for operation. The mathematical model of sensor errors (6) considered above is certainly not complete: the presented model contains only coefficients of dependence of sensor zero offsets on their current temperatures T_i. At the same time, it can be said that the readings of compact and inexpensive MEMS sensors are also greatly influenced by the rate of change in the external temperature $\dot{T}_i$ this should undoubtedly be reflected in the mathematical model of sensor errors. However, this circumstance will further complicate the procedure for identifying coefficients in the calibration process.

To identify the residual errors of accelerometers, it is possible to use a rotary table, which will allow placing the system in 6 specified positions (Fig. 8).

To calibrate the residual errors of the gyroscopes, it is necessary to organize rotations around three connected axes of the turntable, while simultaneously comparing the readings of the gyroscope and the current value of the angular velocity of rotation.

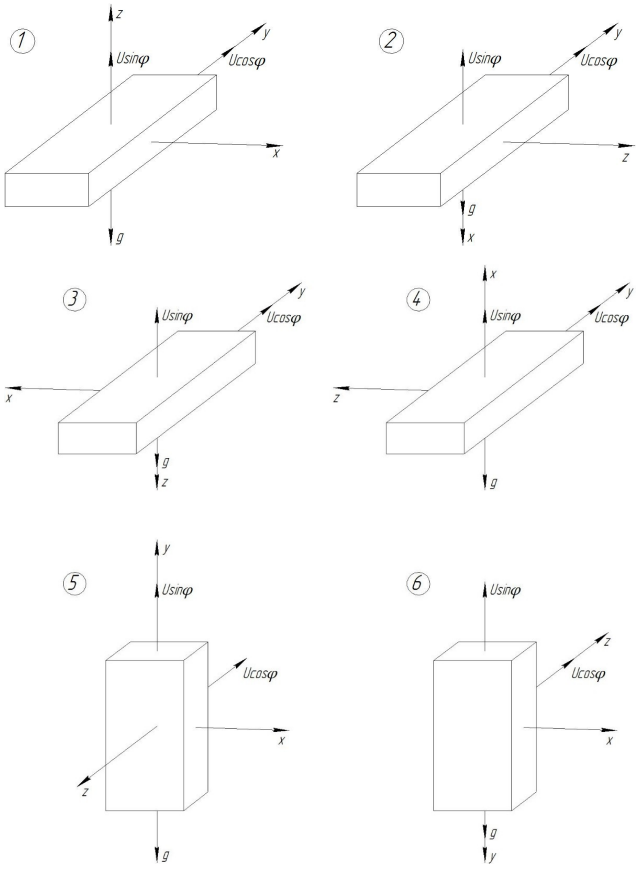


Fig. 8. Positions of the DMHS on the rotary table for accelerometer calibration

For calibration of magnetic field sensors, in most cases, the simplest procedure is used to rotate the DMHS around all three axes in order to equalize the magnetic field along three axes (Fig. 9)

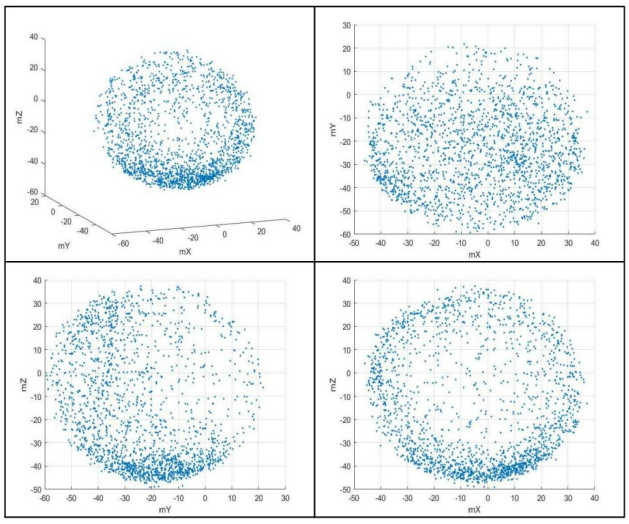


Fig. 9. Readings of DMHS magnetic field sensors during rotation around three axes.

The coefficients estimated using this method are characterized by a high level of error and do not allow their estimation depending on the changing temperature. Based on this, the company "Integral" has developed a technological stand for calibration of magnetic field sensors, consisting of a magnetically shielded housing and a system of 3 mutually orthogonal inductors that allow you to create a magnetic field of different strength and direction. Using this stand, it becomes possible to identify the error model (5) with high accuracy to ensure high-quality measurements of the magnetic field after the calibration procedure of the sensors.

V. DMHS-MICRO MAGNETOMETER

For use on compact and ultra-compact UAVs, Integral LLC developed the DMHS-micro magnetometer (Fig. 10).

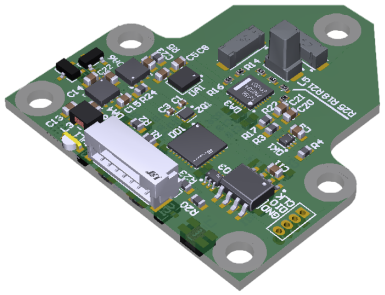


Fig. 10. External view of the DHMS-micro magnetometer

The DMHS-micro magnetometer is acompact device based on MEMS sensors for determining the roll angles, pitch and magnetic course of a moving object. A distinctive feature of the DMHS -micro magnetometer is its small dimensions, weight, and the ability to connect to existing UAV control controllers via UART and CAN interfaces (using UAVCanthe UAVCan / droneCAN protocol).

The use of the DMHS-micro magnetometer is possible under the operating conditions shown in Table 1.

TABLE I. TECHNICAL CHARACTERISTICS OF THE DMHS-MICRO MAGNETOMETER

Parameter	Value
Dimensions	45 x 40 x 15 mm
Weight	20 g
Voltage	= 5 V (NOT protected)
Power consumption	100 MW
Output interface	1 x UART 1 x CAN
Calculation frequency	20 Hz
Ambient temperature	from minus 40° C to plus 70° C
Relative humidity	from 5% to 98 %
Magnetic field measurement range	± 800 uT
Диапазон измерений μrtesla Angular velocity measurement range	± 250 ° / sec (range can be changed to ± 500 º / sec, ± 1000 º / sec or ± 2000 º / sec
Linear acceleration measurement range	± 2g (range can be changed to ± 4g, ± 8g or ± 16g)

The accuracy characteristics of the DMHS-micro magnetometer in the confidence interval of 1 sigma are presented in Table 2.

TABLE II. ACCURACY CHARACTERISTICS OF THE DMHS-MICRO MAGNETOMETER

Parameter	Value
Roll, Pitch (static)	2.0 ^{0}
Roll, pitch (dynamics)	5.0 ^{0}
Magnetic course (static) After the procedure for writing off magnetic deviations	1.0 ^{0}
Magnetic exchange rate (dynamics) After the procedure for writing off magnetic deviations	2.0 ^{0}

VI. Field tests of the DMHS-micro magnetometer

Full-scale tests of the DMHS-micro magnetometer were performed as part of the Micro UAV navigation kit of Integral LLC (Fig. 11)

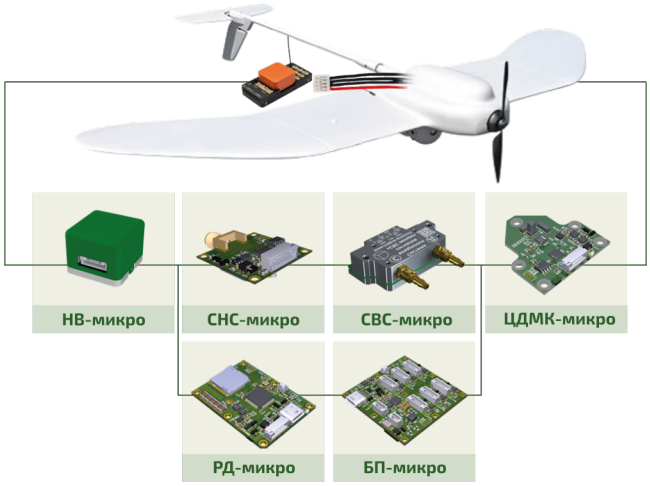


Fig. 11. Micro navigation kit for UAVs of Integral LLC

This complex consists of several devices connected to each other by one common information network via the CAN interface, which allows all devices of the complex to exchange information among themselves. The complex consists of the following devices:

- AHRS
- air data system (ADS)
- magnetic heading sensor.

The basis of the complex is the NV-micro AHRS in the computer of which the general high-frequency navigation solution of the entire complex is formed, including angular velocities, accelerations, angles, linear velocities and geographical coordinates of the UAV. A distinctive feature of the AHRS is temperature stabilization system for, which makes it possible to increase the accuracy and stability of sensor readings over the entire temperature range of using the AHRS. The temperature stabilization system keeps the temperature of the MEMS sensor at +50 ° C and is a heating system with a single heating transistor located under the MEMS sensor (Fig. 12).

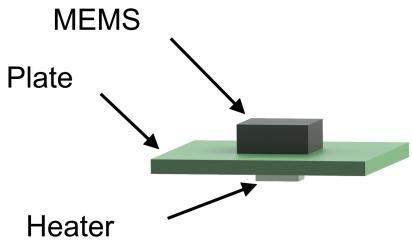


Fig. 12. Thermal stabilization system of the NV-micro AHRS

To eliminate the accumulation of linear velocity errors of the complex, it includes the ADS-micro air signal system, which provides the entire complex with information about static and total pressures, instrument and true airspeeds, and barometric altitude of the UAV. A distinctive feature of the system is the presence of factory calibration of the system sensors in accordance with the following mathematical model (3):

$$P_{St}^{Err} = \rho_{SS}(t) \cdot P_{St} + \rho_S(t)$$
$$P_{Dn}^{Err} = \rho_{DD}(t) \cdot P_{Dn} + \rho_D(t) \quad (7)$$

where $\rho_{SS}(t)$ $\rho_{DD}(t)$ - errors in the scale coefficients of static and total pressure sensors, respectively;

$\rho_S(t)$ и $\rho_D(t)$ - errors of zero offsets of static and total pressure sensors, respectively.

P_{St} и P_{Dn} - true values of static and total pressures, respectively.

P_{St}^{Err} и P_{Dn}^{Err} - errors in the readings of static and total pressure sensors, respectively.

By analogy with gyroscopic sensors, each of the coefficients of the model (7) can be represented as a power function (6), the coefficients of which, depending on the temperature, are subject to identification.

Unlike the NV-micro AHRS, which uses ultra-compact micromechanical angular velocity and acceleration sensors, the ADS-micro air signal system uses precise pressure sensors with very large dimensions. This circumstance does not allow us to build a high-quality and compact system for temperature stabilization of sensor data, which leads to the need for classical temperature compensation of pressure sensor errors. Within the framework of the presented mathematical model of sensor errors (7), Integral LLC developed a technological stand for calibration of pressure sensors, which is a system of pressure setters, tubes and heat-cold chambers combined into one common automated system with software control. Using this stand, it becomes possible to identify the error model (7) even up to high-order polynomials, which ensure high accuracy of pressure measurements after the calibration procedure of the sensors.

To eliminate the accumulated errors of the azimuth channel of the NV-micro AHRS, the complex includes a digital magnetic course sensor "DMHS-micro", which allows you to measure the components of the magnetic field, the magnetic course and the magnetic declination of the UAV.

Additionally, the complex can use satellite navigation systems "GNSS-micro", power supply "PS-micro" and data logger "DR-micro" to save all data from all devices to the SD card for subsequent analysis.

The above-mentioned complex has a number of distinctive features:

- Flexibility in placing its components on the UAV: the navigation system is placed in a place with the lowest possible level of vibrations (it is also possible to use vibration-isolating supports), the ADS is located in a place as close as possible to the incoming air flow, and the magnetic course sensor is located in a place with the level of minimal magnetic disturbances (wing or tail for an aircraft UAV scheme).
- The possibility of performing individual or all components of the complex in a bodyless version, which allows not only to significantly reduce the dimensions of the complex, but also its weight (up to 150 g in a bodyless version).
- Convenient switching between individual components of the complex. Data exchange within the components of the complex and signal transmission to the UAV

controller is performed via just one CAN interface using the standard UAVCAN protocol (DroneCAN).

This complex was placed on an airplane-type UAV with vertical take-off and landing (Fig 13)

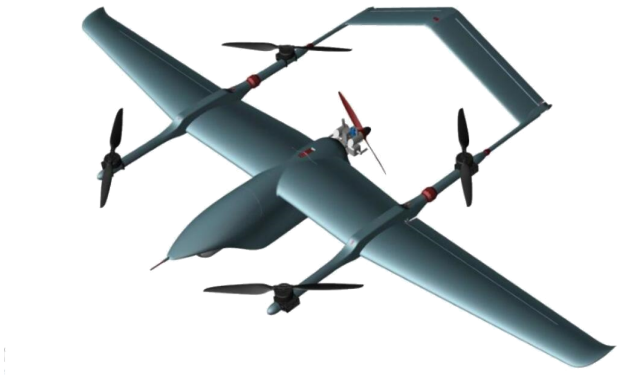


Fig. 13. Appearance of an airplane-type UAV with vertical take-off and landing

Several UAV flights were performed along a similar route-moving along the same closed oval trajectory in order to check the accuracy of navigation of the micro kit (Figs 14, 15).

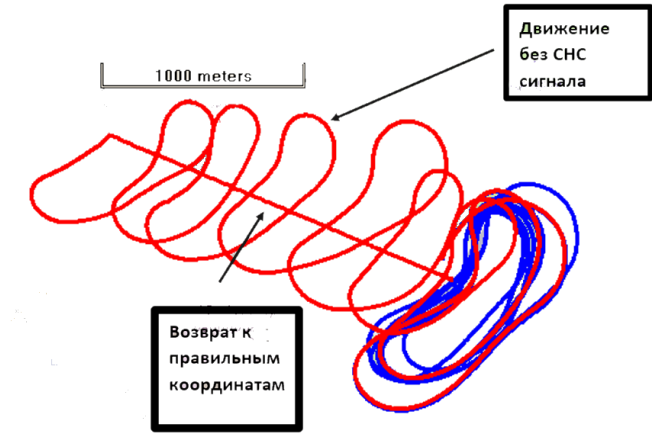


Fig. 14. Trajectory of the UAV movement in one of the flights. Red color – the trajectory of the navigation system in the process of disabling correction from the GNSS, blue color – the trajectory of the GNSS

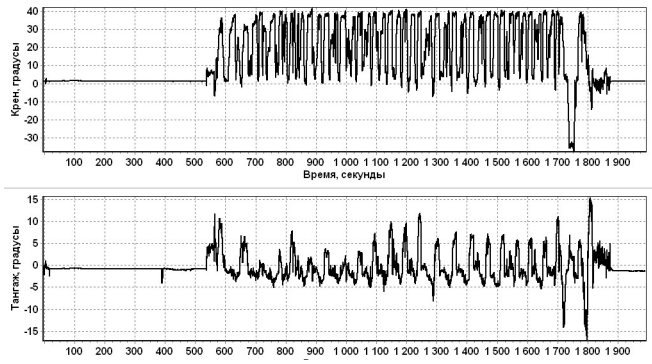


Fig. 15. Beh avior of roll and pitch angles of the NV-micro AHRS in one of the UAV flights

The speed of the UAV movement on this route was about 120 km/h, the height of movement was not more than 300 meters (Figs. 16, 17).



Fig. 16. Comparison of the airspeed of the ADS-micro air signal system (red) with the ground speed readings of the satellite navigation system (black) in one of the UAV flights



Fig. 17. Coparison of the barometric altitude of the ADS-micro air signal system (red) with the absolute altitude of the satellite navigation system (black) in one of the UAV flights

During the tests, comparisons were also made between the DMHS-micro and the GNSS tracking angle, as well as with the terrain map during the flight (Fig. 18)

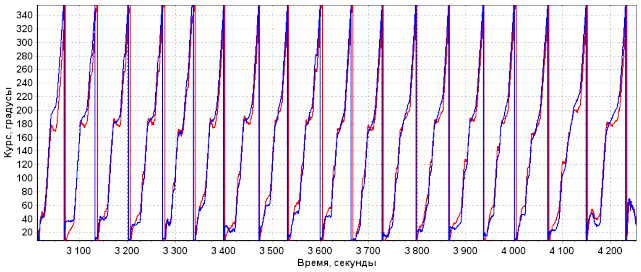


Fig. 18. Comparison of the DMHS-micro magnetic course (red) and the SNA track angle (blue) in one of the UAV flights

Comparison of the trajectory of the aircraft, the readings of the magnetic course and the track angle of the GNSS allowed us to conclude that the accuracy of the DMHS-micro readings is at the level of 2 ° error during movement.

VII. Conclusion

The DMHS-micro magnetometer presented in this article is distinguished by its ultra-compact dimensions and weight. Despite these features, the device allows you to determine the angle of the magnetic course in a wide temperature range with high accuracy.

The proposed magnetometer algorithm uses the classical SINS algorithm with control signals. Changing the signal correction coefficients during the operation of the algorithm allows you to achieve not only optimal results from the point of view of accuracy, but also to calculate the algorithm parameters with minimal delays to obtain the most highly dynamic indicators.

Field tests conducted on a small-sized UAV demonstrated acceptable accuracy for piloting and high dynamic characteristics of the DMHS-micro magnetometer.

References

[1] Golovan A. A., Parusnikov N. A. Matematicheskie osnovy navatsionnykh sistem [Mathematical foundations of navigation systems]. Part 1. Moscow: Maks Press, 2011, 136 p.

[2] Dmitriev S. P. Inertial methods in engineering geodesy. Saint Petersburg: SSC RF "Central Research Institute "Electropribor",

[3] Salychev, O.S., MEMS-based inertial navigation: Expectations and reality, Moscow: Bauman MSTU Press, 2012, 207 p.

[4] Kuzovkov N. T., Karabanov S. V., Salychev O. S. Continuous and discrete control systems and identification methods. Moscow: Mashinostroenie, 1978, 222 p.

[5] D. B. Pazychev and R. N. Sadekov, "MEMS Temperature Stabilization," 2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, Russia, 2020, pp. 1-4, doi: 10.23919/ICINS43215.2020.9133899

[6] D. B. Pazychev and R. N. Sadekov, "Simulation of INS Errors of Various Accuracy Classes," 2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, Russia, 2020, pp. 1-3, doi: 10.23919/ICINS43215.2020.9133869.

[7] D. B. Pazychev, K. S. Bakulev and R. N. Sadekov, "Low-Cost Navigation System for UAV," 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2023, pp. 1-6, doi: 10.23919/ICINS51816.2023.10168469

[8] https://www.integral-group.pro/

Magnetic Compass with a Correction System Built into the Autopilot Circuit

D. G. Gryazin
Concern CSRI Elektropribor, JSC,
ITMO University,
St. Petersburg, Russia
volnagdg@yandex.ru

T. V. Paderina
Concern CSRI Elektropribor, JSC,
ITMO University,
St. Petersburg, Russia
paderinata@gmail.com

***Abstract*—To improve high-quality operation of a vessel's autopilot with a magnetic compass circuit, it is necessary to reduce the dynamic errors of the compass. This is made possible due to the introduction of a correction system into the measuring circuit of the compass. The correction system is based on a complementary filter using a MEMS gyroscope with a vertical sensitivity axis, installed on the compass bowl. Some fragments of the recorded realizations of yaw angles generated by the magnetic compass Azimuth KM-05D with the correction system operating and disabled under field testing conditions are presented. The test results have shown that the proposed magnetic compass allows "smoothing" of yaw angles, which brings such a compass closer to a gyrocompass (in terms of its characteristics) in the autopilot circuit, thus ensuring accurate keeping of the vessel on course.**

Keywords—magnetic compass, correction system, MEMS gyroscope, autopilot

I. Introduction

All vessels are currently equipped with automatic course-keeping systems, the main part of which is an autopilot. In the majority of cases, a gyrocompass is used as the main course detector which implements feedback in the autopilot.

At the same time, many autopilots, e.g. Navitron NT 888G, Navitron NT 921G (Great Britain), Miran AR-1-1 (Russia), and Arkas (Denmark), provide for the operation mode from a magnetic compass (MC). Such a situation may arise if the ship's gyrocompass fails, but in the modern practice of ship navigation, it is rather often that the MC is used as an only source of information about the vessel's heading because in accordance with international requirements, only MCs can be used on vessels with a displacement of up to 500 tons, so that they have no gyrocompasses.

However, conventional MCs have significant dynamic errors that manifest themselves under pitch and roll conditions, which is explained by two factors. Firstly, the influence of redistributed ship magnetism forces on the magnetic system of the compass card. This effect causes heeling deviation [1,4] that can be eliminated only for a specific latitude of fix. Secondly, the influence of tangential and centripetal (translational) accelerations on the card, which is a short-period pendulum, caused by the MC installation at a certain distance from the ship oscillation center [2]. The errors mentioned above result in unstable operation of the MC and errors in yaw angle measurement.

These errors are most pronounced during the MC operation in high latitudes, which is an additional negative factor in solving the heading indication problem in such latitudes, since the guiding magnetic moment of the card, arising from the horizontal component of the Earth's magnetic field, is already small. Given the current trend towards an increase in the number of vessels, including small-tonnage vessels in the northern seas, which is due to the increasingly intensive use of the Northern Sea Route, suppression of the MC dynamic errors becomes especially relevant [3].

It should also be noted that MC dynamic errors adversely affect the quality of the autopilot performance. The MC operates on the principle of compensation for the vessel's deviation from the specified course, i.e. its control algorithm continuously compares the real and specified heading values. Deterioration in the quality of its performance during rolling results in frequent rudder displacement and the inability to compensate for yaw angles.

Thus, when MCs are used in autopilots, primarily on small-tonnage fishing, coastal and other vessels, there is a need to increase the MC accuracy to a level as close to the accuracy of conventional gyrocompasses as possible.

MC dynamic errors can be reduced by using a correction system included in the MC, based on the integration of output signals of the MC magnetic heading sensor (MHS) and a MEMS gyroscope with a vertical sensitivity axis, installed on the MC bowl.

Such a correction system was implemented in the MC Azimut KM-05D.

In this paper, the possibility of using MCs with this kind of correction system in autopilots of small vessels is theoretically substantiated and practically confirmed using the results of the field tests of the MC Azimuth KM-05D.

II. Magnetic Compass Dynamic Error Correction System

The system of MC dynamic error correction at pitch and roll is based on the application of the problem of optimal integration of the output signals from the compass MHS and MEMS gyroscope (from here on called "gyro") with a vertical sensitivity axis that generate similar information—the sum of the vessel's heading and yaw angles. However, these two measuring devices have different dynamic characteristics since, unlike the MC sensitive element, the gyro is insensitive not only to the action of translational accelerations, but also to the redistributed forces of ship magnetism under the the vessel's heeling. For design considerations, the gyro can be installed directly on the compass bowl.

It is known that a complementary filter can be successfully applied for this kind of measuring devices [5]. The integration method implemented by this filter is aimed to compensate for the dynamic errors caused by the action of

translational accelerations and ship magnetism at pitch and roll by means of the gyro, as well as for the errors caused by the low-frequency components of the gyro drift using the MHS readings. In this case, for the gyro, the complementary filter is a high-pass filter (HPF), whereas for the MHS, a low-pass filter (LPF) [6]. Therefore, the dynamic errors arising at pitch and roll in the angle representing a sum of the vessel's heading and yaw angles generated by the MC will be reduced.

The equation of the complementary filter has the form [7]:

$$\hat{\alpha} = W_{HPF}(p) \cdot \alpha_g + W_{LPF}(p) \cdot \alpha_{dhs}, \quad (1)$$

where $W_{HPF}(p)$ is the HPF translational function,

$W_{LPF}(p)$ is the LPF translational function,

α_g, α_{mhs} are the sums of the heading and yaw angles, respectively, derived from the gyro and MHS information; $\hat{\alpha}$ is the estimate of the sum of the heading and yaw angles produced by the filter.

In a simplest case, the corresponding links with translational functions can be used as the HPF and LPF:

$$W_{HPF}(p) = \frac{Tp}{Tp+1}, \quad W_{LPF}(p) = \frac{1}{Tp+1}, \quad (2)$$

where T is the time constant of the HPF and LPF.

In order for the LPF and FHF filters not to affect the useful signal, the following invariance condition must be fulfilled [8]:

$$W_{LPF}(p) + W_{HPF}(p) = 1 \ . \quad (3)$$

Taking into account the invariance condition, the error in estimating the angle $\Delta\hat{\alpha} = \alpha - \hat{\alpha}$ will depend on the MHS errors $\Delta\alpha_{mhs}$, gyro $\Delta\alpha_g$, and the time constant T:

$$\Delta\hat{\alpha} = \frac{1}{Tp+1}\Delta\alpha_{mhs} + \frac{Tp}{Tp+1}\Delta\alpha_g \ . \quad (4)$$

Let us write a difference equation to find the current estimate of the heading and yaw angles $\hat{\alpha}$ [7]:

$$\hat{\alpha}_t = \frac{T}{T+\Delta t}\left[\hat{\alpha}_{t-1} + \omega_g \Delta t\right] + \frac{\Delta t}{T+\Delta t}\alpha_{mhs}, \quad (5)$$

where $\hat{\alpha}_{t-1}$ is the estimate of the sum of the heading and yaw angles at the previous point in time;

ω_g is the angular rate from the gyro;

Δt is the sampling interval.

In the MC correction system, we use a low-cost MEMS gyro with the following characteristics:

- the angular rate measurement in the range from 75 to 500 deg/s;
- bias instability – max 24 deg/h,
- ARW – max 0.28 deg/√h,
- bandwidth – up to 160 Hz.

Taking into account the capabilities of the gyro used: the sampling interval Δt = 0.01 s and, also, based on the simulation results, we chose the filter time constant T = 10 s. Figure 1 shows a block diagram of the MC correction system based on a complementary filter [9].

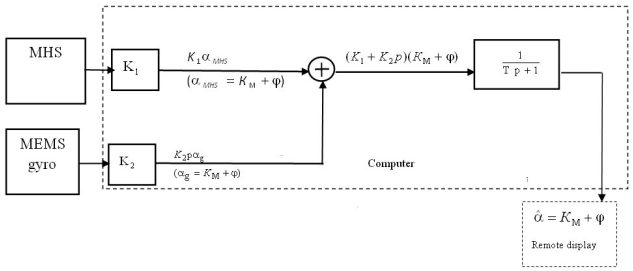


Fig. 1. A block diagram of the MC correction system based on a complementary filter [9].

To ensure the invariance condition (2), factors K_1 and K_2 (Fig. 1) should be selected based on the condition that $T = \frac{K_2}{K_1}$. If K_1 is selected equal to 1, then K_2 should be equal to the time constant: $K_2 = T$. Thus, the sum of the heading and yaw angles $K_M + \varphi$ (K_M is the magnetic heading, φ is the yaw angle) in Fig. 1 should be generated according to the following algorithm:

1. The MHS output signal is fed to the link with the gain factor K_1, and the gyro output signal, to the link with the gain factor K_2.

2. The MHS signal $K_1 \cdot \alpha_{mhs}$ (where $\alpha_{mhs} = K_M + \phi$) is summed with the gyro signal $K_2 p\alpha_g$ (where $\alpha_g = K_M + \phi$);

3. The resulting sum of the signals is fed to input of the aperiodic link $\frac{1}{T\,p+1}$. The signal from the output of the aperiodic link $\hat{\alpha} = K_M + \varphi$, representing the estimate of the sum of the heading and yaw angles, without MC dynamic error caused by pitch and roll, is delivered to the remote display.

III. THE RESULTS OF THE BENCH AND FIELD TESTS OF THE MC WITH A CORRECTION SYSTEM

The efficiency of the MC correction system based on a complementary filter with the time constant T = 10 s aimed to reduce the MC dynamic errors under the pitch and roll was experimentally estimated using a yawing test bench [10]. The MC functioning with the correction system was tested at a latitude of 60°N (the horizontal component of the magnetic field strength was (15±1) μT) while the yawing test bench was set to a mode of an angular motion in yaw in the range of ±10º angles with different periods from 4 to 30 s. In this case, permanent magnets placed near the bowl in the extreme (10º and minus 10º) positions during the angular motion generated an effect that simulated the appearance of a heeling deviation, leading to an additional turn of the card by approximately two times, in the range of angles of ±20º. The MC correction system must return the MC readings to the

yaw angles set by the test bench, i.e. to the angles in the range of ±10º. The efficiency of the MC correction system can be estimated with the use of the efficiency factor, which is calculated as follows:

$$K_{ef} = \left|DK^{M}_{no\ cor}\right| / \left|DK^{M}_{cor}\right|, \qquad (6)$$

where $DK^{M}_{no\ cor}$, DK^{M}_{cor} are the MC errors representing the limiting differences between the MC noncorrected and corrected outputs and true yawing angles.

Figure 2 shows the dependence K_{ef} on the periods of disturbing effects (during the bench tests, these are the periods of the specified harmonic oscillations of the yaw angles). The K_{ef} values lie in the range from 3.7 to 9.8, while the yawing periods vary from 4 to 30 s; note that the maximum K_{ef} value was observed for the period of 18 s.

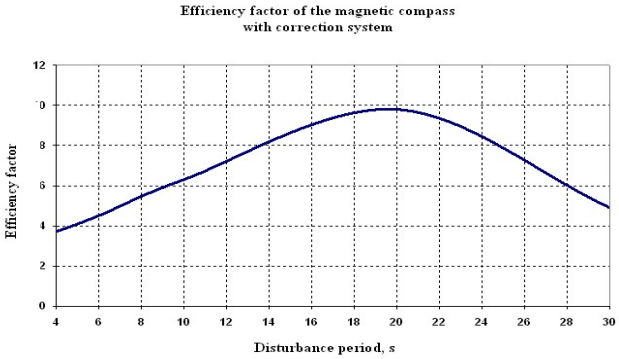


Fig. 2. Dependence of the MC correction system efficiency factor on the periods of disturbances.

The field tests of the MC Azimut KM-05D were carried out on board the Azimut ship with displacement of 100 tons. In the tests, the magnetic compass was installed on the bridge deck. The additional equipment included a strapdown attitude determination system with a GNSS receiver to determine the ship oscillation angles and a ship gyrocompass used as a source of the reference heading. The additional vertical magnets, which are normally used to compensate for heeling deviation, were removed from the binnacle during the measurements in order to test the effect of noncompensated heeling deviation on the correction system [11].

The MC field tests have shown that under roll with an amplitude of up to 17º and pitch with an amplitude of up to 4.5º and periods of the both close to 4 s, the correction system reduces the MC error; in so doing, its efficiency factor K_{ef}, depending on pitch and roll periods, remains in the range from 2.4 to 3.5. This is consistent with the test results obtained on the yaw test bench with additional magnets simulating the effect of heeling deviation, where K_{ef} for yawing periods of about 4 s was 3.7.

When the MC is included in the autopilot circuit, it is important to ensure its high-quality operation, which mainly depends on two factors: the number of rudder shifts and the values of the shift angles themselves (it is desirable that these angles have minimum values). With a large number of rudder shifts *òa!*, (more than 400 per hour), the accuracy of keeping the vessel on course does not increase, this only results in accelerated wear of the rudder gear and loss of useful power of the vessel's power supply unit [12, 13].

Figures 3–5 show typical fragments of recorded yaw-angle realizations measured by the MC with the correction system operating and disabled.

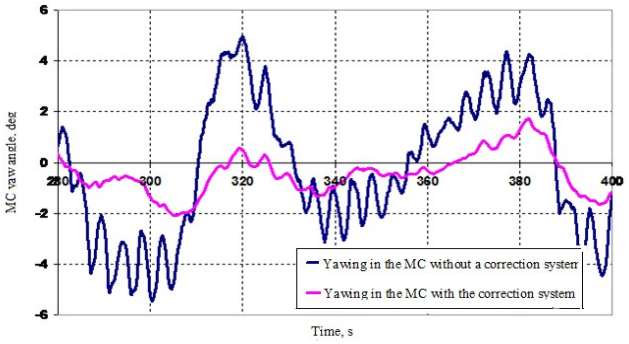


Fig. 3. Yaw angle in the MC with the correction system operating and disabled (vessel heading K = 40º).

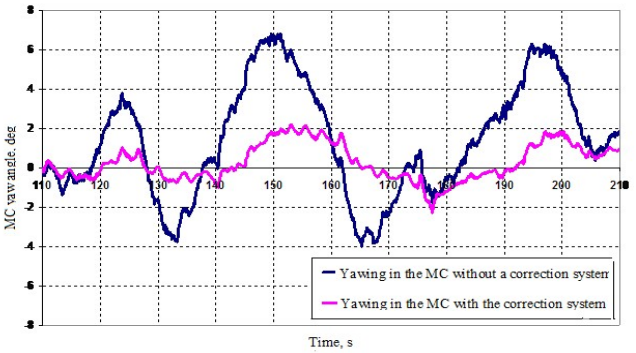


Fig. 4. Yaw angle in the MC with the correction system operating and disabled (vessel heading K = 20º).

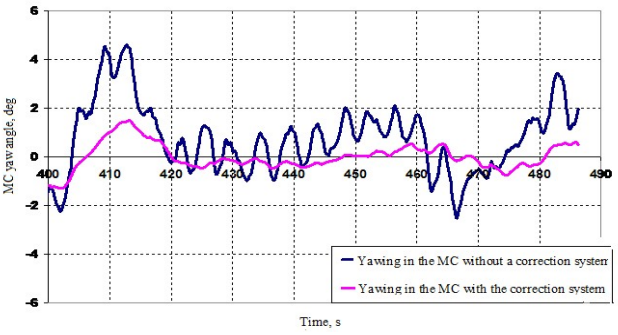


Fig. 5. Yaw angle in the MC with the correction system operating and disabled (vessel heading K = 190º).

Based on the plots shown in Figs. 3–5, we can conclude that the correction system included in the MC makes it possible to "smooth" the MC yaw angles, bringing such an MC closer to a gyrocompass, in terms of its characteristics, included in the autopilot circuit to ensure accurate steering and keeping the vessel on course.

IV. Conclusions

The results of the tests allow us to draw the following conclusions:

The results of the tests allow us to draw the following conclusions:

• the bench tests have shown that the correction system of the MC Azimut KM-05D based on a complementary filter makes it possible to reduce dynamic errors by a factor of 3.5–9.8, depending on the prevailing frequencies (periods) of the ship's pitch and roll;

• the field tests of the MC Azimut KM-05D with the correction system in the Lake Ladoga water area have confirmed the "smoothing" effect of yaw angles, thereby proving the possibility of using such an MC in the autopilot circuit.

The use of the MC with the correction system in the autopilot circuit will reduce the number of rudder shifts and the values of these shift angles, bringing such an MC closer to the characteristics of a gyrocompass included in the autopilot circuit to ensure accurate keeping of the vessel on course.

REFERENCES

[1] Nechaev, P.A. and Grigoriev, V.V., *Magnitno-kompasnoe delo* (Magnetic Compass Technology), Moscow: Transport, 4th edition,

[2] Rybaltovskii, N.Yu., *Magnitno-kompasnoe delo* (Magnetic Compass Technology), Leningrad: Vodtransizdat, 1954.

[3] Kardashinskii-Braude, L.A., *Sovremennye sudovye magnitnye kompasy* (Modern Ship Magnetic Compasses), State Scientific Center of the Russian Federation – Central Research Institute Elektropribor, St. Petersburg. 1999.

[4] Kardashinskii-Braude, L.A., *Deviatsiya magnitnykh kompasov* (Deviation of Magnetic Compasses), Leningrad: TsNII Rumb, 1986.

[5] Matveev V.V. Analysis of complementary filters in constructing a gimballess vertical gyro, Izvestiya Tul'skogo gos. universiteta. Tekhnicheskie nauki, 2019, no. 8, pp. 153–164.

[6] Ivoilov, A.Yu., MEMS sensors application in the development of two-wheeled robot automatic stabilization system, *Sbornik nauchnykh trudov NGTU*, 2017, no. 3 (89), pp. 32–51.

[7] Stepanov, O.A. and Mansur, M., Integrated processing algorithms for correction of navigation system using nonlinear measurements, *Izvestiya TulGU. Tekhnicheskie nauki*, 2016, no. 6, pp. 89–102.

[8] Zo M'u Nain, Shchagin, A.V., Le Vin' Tkhang, and Khtin Linn, U., Complementary filter for angle estimation using a MEMS gyro and accelerometer, *Inzhenernyi vestnik Dona*, no. 3, 2020.

[9] Matveev, Yu.V., Gryazin, D.G., and Paderina, T.V., Method for measuring the ship's magnetic heading using a correction system, Patent RU 2804444, registered 29.09.2023.

[10] Gryazin, D.G., Method for determining the dynamic error of a magnetic compass with a roll/pitch correction system and a device for its implementation, Patent RU 2783479, registered 14.11.2022.

[11] Gryazin, D.G. and Sergachev, I.V., Influence of external factors on transient processes of magnetic compass in high latitudes, *Journal of Instrument Engineering*, 2020, vol. 63, no. 10, pp. 921–929.

[12] Preobrazhenskii, A.V. Optimization of the autopilot parameters on a mathematical model of a ship. *Nauchnyye problemy vodnogo transporta*. Vol. 45, 2015.

[13] Jung, V.L., Vessel heading control system, *Aktual'nye problemy aviatsii i kosmonavtiki.* (Current Problems of Aviation and Cosmonautics) *Sbornik materialov VI Mezhdunarodnoi nauchno-prakticheskoi konferentsii,* Vol. 1, Krasnoyarsk, April 2020.

An Adaptive Initial Alignment Filtering Algorithm for Strapdown Inertial Navigation under Large Misalignment Angles

Xiaoge Ning
Department of Automation, School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
Beijing Aerospace Times Optical-electronic Co. Ltd
Beijing, China
nxg01@126.com

Jixun Huang
Beijing Institute of Aerospace Control Devices
Beijing, China
huangjixun008@163.com

Jianxun Li
Department of Automation, School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
lijx@sjtu.edu.cn

***Abstract*—Initial alignment is the basis for obtaining accurate misalignment angles of Strapdown Inertial Navigation Systems (SINS), and its accuracy and efficiency will directly affect the SINS performance.In this paper, aiming at the problem of measurement noise deviating from the prior in the practical application of SINS, an adaptive filtering method based on chi-square detection is proposed. Through the simulation experiment of SINS initial alignment, it is verified that the proposed method has higher accuracy and stronger robustness compared with the traditional method. The method in this paper can avoid the maximum 0.6° jump of the attitude alignment result that occurs in the traditional method when there is a measurement jump.**

Keywords—adaptive filtering method; SINS inertial alignment; chi-square detection.

I. INTRODUCTION

Initial alignment is the basis for obtaining accurate misalignment angles of Strapdown Inertial Navigation Systems (SINS), and its accuracy and efficiency will directly affect the SINS performance[1-3].

The traditional linear error models for initial alignment based on small misalignment angles have become mature and have been used in a variety of ways, covering static, dynamic and shaky disturbance bases, and are usually based on linear differential error equation models derived from perturbations of the Eulerian Φ angle or ψ angle[4-5].

However, with the development of SINS technology and the emergence of different application scenarios, the classical small misalignment angle error model is subject to large errors and inaccuracies in complex applications such as high dynamics and large misalignment angles. As a result, initial alignment techniques and non-linear filtering techniques for large misalignment angles have been researched and developed, and the SINS error model has changed from the traditional linear model to a non-linear model.

In practical initial alignment applications, there often exist errors in the process noise covariance matrix Qk or the measurement noise covariance matrix Rk, or unreasonable initial settings of filter parameters, which can lead to a decrease in the precision of Kalman filtering and even cause filter divergence, rendering the filtering results unusable[6].

To address this issue, this paper proposes a chi-square detection based self-adaptive initial alignment filtering algorithm for large misalignment angles in strapdown inertial navigation systems, which can improve the precision of large misalignment angle initial alignment under abnormal measurement conditions. Simulation experiments show that, the proposed method has better robustness and adaptability compared with traditional methods under abnormal measurement conditions.

The remainder of the paper is organized as follows. In Section II, an adaptive filtering method based on chi-square detection is proposed. Section III presents the SINS alignment simulation process and results. In Section IV conclusions are drawn.

II. ADAPTIVE FILTERING METHOD BASED ON CHI-SQUARE DETECTION

For Kalman filtering in SINS alignment, achieving the optimal estimation of the state of a stochastic system requires a precise knowledge of both the structural characteristics of the system and the characteristics of the noise. However, in practical applications, there are often situations where there are errors in the process noise covariance matrix or the measurement noise covariance matrix, or where the initial values of the filtering parameters are set unreasonably. These situations will reduce the accuracy of the Kalman filter, and may even cause the filter to diverge, making the filtering results unusable.

In 1969, A. P. Sage and G. W. Husa proposed the Sage-Husa adaptive filtering algorithm, which can estimate the parameters of the system noise and measurement noise of a stochastic system. But in engineering applications, it is difficult to simultaneously estimate all the noise parameters, including the mean and variance of the system noise and the mean and variance of the measurement noise. Therefore, usually only the variance of the measurement noise is adaptively estimated.

However, the traditional Sage-Husa measurement noise adaptive filtering algorithm assumes that the variance of the measurement noise changes slowly, which is not suitable for situations where the variance of the measurement noise changes rapidly. In response to this, an adaptive algorithm suitable for rapid changes in the variance of the measurement noise is proposed based on the chi-square detection method.

The filtering measurement update formula is:

This work was supported by Chinese National Programs for High Technology Research and Development （863-706）（2013AA706503）

$$\begin{cases} \mathbf{K}_k = \mathbf{P}_{k/k-1}^{xz}(\mathbf{P}_{k/k-1}^{zz})^{-1} \\ \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k/k-1} + \mathbf{K}_k\tilde{\mathbf{z}}_{k/k-1} \\ \mathbf{P}_k = \mathbf{P}_{k/k-1} - \mathbf{P}_{k|k-1}^{xz}(\mathbf{P}_{k|k-1}^{zz})^{-1}(\mathbf{P}_{k|k-1}^{xz})^T \end{cases} \tag{1}$$

The statistical quantity λ_k can be constructed from the measurement innovation $\tilde{\mathbf{z}}_{k/k-1}$ and its mean square error matrix $\mathbf{P}_{k/k-1}^{zz}$.

$$\lambda_k = (\tilde{\mathbf{z}}_{k/k-1})^T(\mathbf{P}_{k/k-1}^{zz})^{-1}\tilde{\mathbf{z}}_{k/k-1} \tag{2}$$

And the statistical quantity λ_k follows a chi-square distribution with freedom degree of m: $\lambda_k \sim \chi^2(m)$. Under normal measurement, the value of λ_k is relatively small, while in the case of an anomaly, the value of λ_k is relatively large. For this reason, a threshold value T_{Dm} can be selected as the judgment threshold for measurement anomalies.

Therefore, the filtering measurement update calculation will be carried out when $\lambda_k \le T_{Dm}$, and the measurement update calculation will not be carried out when $\lambda_k > T_{Dm}$, thus avoiding the influence of measurement outliers on the filtering performance.

In practical applications, considering the multi-dimensional measurement situation, a chi-square test can be performed on each measurement component one by one to avoid the problem that the abnormality of any measurement component will lead to a decrease in the utilization rate of other normal measurement components at the same time.

The result of the traditional chi-square detection method is either 0 or 1, that is, the measurement information is either used or discarded. However, in practical application environments, the measurement information is often not limited to the two extreme states of normal and abnormal, but there are rich intermediate states. If only a simple binary judgment is made, the utilization rate of the measurement information will decrease. Therefore, based on the chi-square test statistic, the binary value of the weight can be improved to a continuous value. When the measurement information is within the normal range, the weight should be the maximum value to make full use of the measurement information. When the measurement innovation is at a medium level (in a suspicious state), the weight should be appropriately reduced, and when the measurement information is obviously abnormal, the measurement weight should be greatly reduced (when the weight is close to 0, it is equivalent to discarding the measurement).

At this time, the filtering measurement update formula is improved as follows:

$$\begin{cases} \mathbf{K}_k = \mathbf{P}_{k/k-1}^{xz}\sqrt{\boldsymbol{\gamma}_k}(\mathbf{P}_{k/k-1}^{zz})^{-1} \\ \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k/k-1} + \mathbf{K}_k\tilde{\mathbf{z}}_{k/k-1} \\ \mathbf{P}_k = \mathbf{P}_{k/k-1} - \mathbf{P}_{k/k-1}^{xz}\sqrt{\boldsymbol{\gamma}_k}(\mathbf{P}_{k/k-1}^{zz})^{-1}(\mathbf{P}_{k/k-1}^{xz}\sqrt{\boldsymbol{\gamma}_k})^T \end{cases} \tag{3}$$

$$\begin{cases} \sqrt{\boldsymbol{\gamma}_k} = \mathrm{diag}\left(\sqrt{\gamma_k^{(1)}} \;\; \sqrt{\gamma_k^{(2)}} \;\cdots\; \sqrt{\gamma_k^{(m)}}\right) \\ \gamma_k^{(i)} = \begin{cases} 1 & \lambda_k^{(i)} \le T_{D1} \\ T_{D1}/\lambda_k^{(i)} & \lambda_k^{(i)} > T_{D1} \end{cases} \\ \lambda_k^{(i)} = \tilde{z}_{k/k-1}^{\mathrm{T}}\left(P_{ZZ,k/k-1}^{(ii)}\right)^{-1}\tilde{z}_{k/k-1}^{(i)} = \left(\tilde{z}_{k/k-1}^{(i)}\right)^2 / P_{ZZ,k/k-1}^{(ii)} \end{cases} \tag{4}$$

Among them, $\lambda_k^{(i)}$ is the chi-square test statistic of each measurement component, and it satisfies $\lambda_{k(i)} \sim \chi^2(1)$. When the significance level is set to 0.05, $T_{D1} \approx 3.8$.

By combining the adaptive filtering method proposed above with the the cubature Kalman filter (CKF), the adaptive CKF filtering method can be obtained.

III. SINS ALIGNMENT SIMULATION TEST

The SINS updating algorithm used in the initial alignment of the static base condition includes attitude-updating and velocity-updating, the corresponding differential equations are expressed with (5) to (6):

$$\dot{\mathbf{C}}_b^n = \mathbf{C}_b^n[(\boldsymbol{\omega}_{ib}^b - \boldsymbol{\omega}_{ie}^b)\times] \tag{5}$$

$$\dot{\mathbf{v}}^n = \mathbf{C}_b^n\mathbf{f}^b + \mathbf{g}^n \tag{6}$$

Where *b* frame is the measurement reference frame, *n* frame is the navigation reference frame (local level coordinate frame), *i* frame is the inertial reference frame, $\mathbf{C}_b^n$ is the direction cosine matrix of the attitude converting from *b* frame to *n* frame, $\boldsymbol{\omega}_{ib}^b$ is the rotational angular rate of *b* frame relative to *i* frame (coordinatized in the *b* frame), $\boldsymbol{\omega}_{ie}^b$ is the rotational angular rate of the earth (coordinatized in the *i*-frame), $(\bullet)\times$ denotes the antisymmetric matrix, the vector $\mathbf{v}^n$ expresses the eastern velocity, northern velocity and vertical velocity (coordinatized in the n frame), respectively, $\mathbf{f}^b$ is the output of the accelerometer (coordinatized in the *b* frame), $\mathbf{g}^n$ is the gravity vector (coordinatized in the *n* frame).

The large misalignment angle error model of the SINS are expressed with (7) in the case of ignoring the instrument bias:

$$\begin{cases} \dot{\boldsymbol{\alpha}} = \mathbf{C}_{\omega}^{-1}(\mathbf{I}_3 - \mathbf{C}_n^{n'})\boldsymbol{\omega}_{ie}^n \\ \delta\dot{\mathbf{v}}_{en}^n = \left[\mathbf{I}_3 - (\mathbf{C}_n^{n'})^T\right]\mathbf{C}_b^{n'}\tilde{\mathbf{f}}^b \end{cases} \tag{7}$$

Where $\boldsymbol{\alpha}$ denotes the vector of large angles between the true n-frame and the calculation reference frame (abbreviated as *n*' frame), $\mathbf{I}_3$ is a third-order identity matrix, $\mathbf{C}_n^{n'}$ is the attitude matrix between the true n-frame and the calculation reference frame, $\mathbf{C}_b^{n'}$ is the attitude matrix between the measurement reference frame and the calculation reference frame , $\boldsymbol{\omega}_{ie}^n$ is the rotational angular rate of the earth (coordinatized in the n-frame), $\tilde{\mathbf{f}}^b$ is the real accelerometer outbut, $\mathbf{C}_{\omega}^{-1}$ is the matrix described as follows:

$$\mathbf{C}_{\omega}^{-1} = \frac{1}{\cos\alpha_x}\begin{bmatrix} \cos\alpha_y\cos\alpha_x & 0 & \sin\alpha_y\cos\alpha_x \\ \sin\alpha_y\sin\alpha_x & \cos\alpha_x & -\cos\alpha_y\sin\alpha_x \\ -\sin\alpha_y & 0 & \cos\alpha_y \end{bmatrix} \tag{8}$$

The measurement is the velocity error, and Equation (7) is the dynamic model, the state is described as follows:

$$\mathbf{X} = \left[(\boldsymbol{\alpha})^T \quad \left(\delta \mathbf{v}_{en}^n\right)^T \right]^T \tag{9}$$

The simulation conditions are set as follows: the gyro constant drift is ε = 0.01 °/h, random drift is 0.001 °/$\sqrt{h}$; the accelerometer zero bias is ∇ =50 μg, random drift is 10($\mu g/\sqrt{Hz}$); the local geographic latitude is 40°, longitude is 110°; initial attitude is set as [0º;0º;0º]; the initial misalignment angles are set as $\boldsymbol{\alpha}_0 = [1°;1°;5°]$; the sampling period of inertial device is ts=0.01s and the simulation period is 600s. The general parameters of the initial alignment algorithms are set with the same values. The initial value of state is set as $\mathbf{X}_0 = [\mathbf{0}_{6\times1}]$. The matrix of Q is determined according to the random drift of IMU, and Qk=diag{0.001(°/$\sqrt{h}$); 0.001(°/$\sqrt{h}$); 0.001(°/$\sqrt{h}$); 10($\mu g/\sqrt{Hz}$); 10($\mu g/\sqrt{Hz}$); 10($\mu g/\sqrt{Hz}$);}^{2}×ts. The matrix of Rk is set as diag{0.01(m/s); 0.01(m/s); 0.01(m/s)}^{2}/ts.

Considering that short-term gusts and other sudden linear interference often occur during the SINS initial alignment on quasi-static bases, carrier interference linear velocities of 2 m/s are added during six time intervals: 230s~232s, 280s~282s, 330s~332s, 460s~462s, 498s~500s, and 540s~542s.

In this simulation, the cubature Kalman filter (CKF3) and the adaptive cubature Kalman filter proposed in this paper (ACKF3) are respectively used, and the results are shown in Fig.1-Fig.3. The mean attitude error values during the last 50 seconds are calculated, as shown in Table I.

TABLE I. THE ATTITUDE ERROR BY DIFFERENT FILTERS

Table Head	Alignment results at 600s	
	CKF3	***ACKF3***
East misalignment	-0.0028°	-0.0028°
North misalignment	0.0028°	0.0028°
Upward misalignment	-0.0693°	-0.0578°

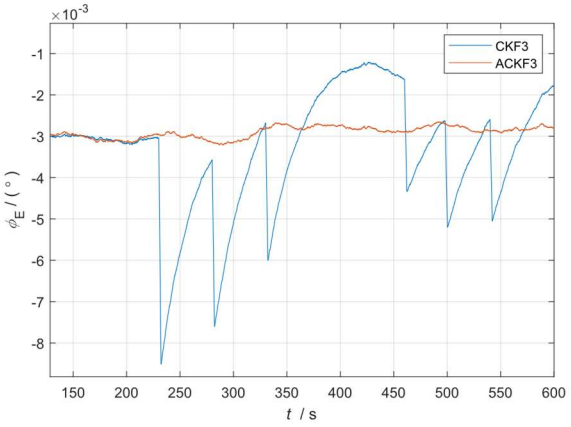


Fig. 1. East misalignment angle estimated by different filters.

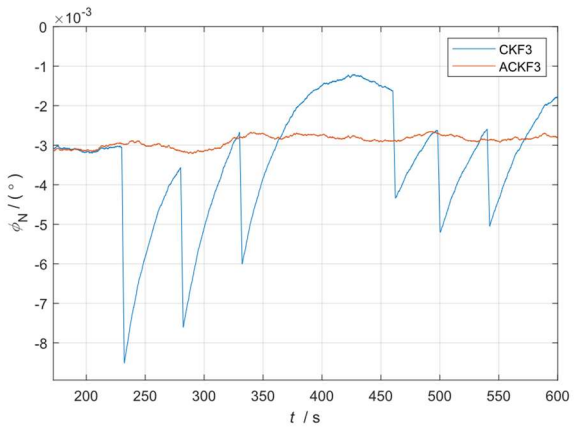


Fig. 2. North misalignment angle estimated by different filters.

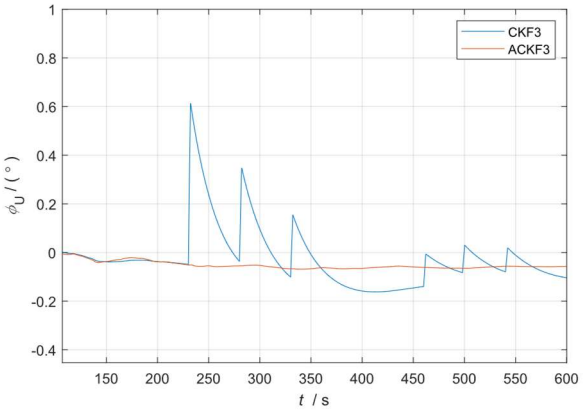


Fig. 3. Upward misalignment angle estimated by different filters.

As can be seen from Table I, the filtering accuracy of the horizontal misalignment angle of the ACKF3 method proposed in this paper is close to that of the traditional method, but the accuracy of the upward misalignment angle is better than that of the traditional method.

From the convergence process of the attitude errors in Fig.1 to Fig.3, it can be seen that the traditional filtering method highly depends on the accuracy of the system model noise modeling. When there is a sudden interference in the measurement, the estimated results of the filtering show a jump phenomenon (the jagged part of the blue curve in Fig.1 to Fig.3), ultimately leading to a decrease in the filtering accuracy. In Fig.3, the upward misalignment result exhibit a maximum sudden change of 0.6°.

However, the adaptive cubature Kalman filter proposed in this paper has good adaptability. When there are measurement anomalies, it can still perform filtering more accurately, being almost unaffected by measurement anomalies. The filter has higher accuracy and stronger robustness.

IV. CONCLUSION

In this paper, aiming at the problem of measurement noise deviating from the prior in the practical application of SINS, an adaptive filtering method based on chi-square detection is proposed.

Finally, through the simulation experiment of SINS initial alignment, it is verified that the proposed method has

higher accuracy and stronger robustness compared with the traditional method. The method in this paper can avoid the maximum 0.6° jump of the attitude alignment result that occurs in the traditional method when there is a measurement jump.

ACKNOWLEDGMENT

The authors acknowledge the reviewers who give many detailed suggestions.

REFERENCES

[1] D.H. Titterton and J.L. Weston, Strapdown Inertial Navigation Technology, 2nd ed. London, United Kingdom: The Institute of Electrical Engineers, 2007.

[2] P.G. Savage, Strapdown Analytics, 2000.

[3] P.D. Groves, Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems, Second Edition. 2013, pp. 1-776.

[4] C. Jekeli, Inertial Navigation Systems with Geodetic Applications, de Gruyter, 2012.

[5] D. Simon, Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches, John Wiley & Sons, 2006.

[6] G. Yan, F. Liu, Z. Li, Q. Zhou, "An adaptive Kalman filter based on soft Chi-square test," Navigation Positioning and Timing, vol. 10, no. 4, pp. 81-86, July 2023.

Calibration of a Strapdown Inertial Measurement Unit with Two Angular Rate Sensors

L. Vodicheva
JSC Academician N.A. Semikhatov Scientific and Production Association of Automatics
Yekaterinburg, Russia
avt@npoa.ru

N. Shmakov
JSC Scientific and Production Association of Electromechanics
Miass, Russia
101@npoe.ru

D. Kabanova
JSC Academician N.A. Semikhatov Scientific and Production Association of Automatics
Yekaterinburg, Russia
avt@npoa.ru

Ya. Savel'eva
JSC Academician N.A. Semikhatov Scientific and Production Association of Automatics
Yekaterinburg, Russia
avt@npoa.ru

Yu. Parysheva
JSC Academician N.A. Semikhatov Scientific and Production Association of Automatics
Yekaterinburg, Russia
avt@npoa.ru

***Abstract*—Bench calibration of a strapdown inertial measurement unit (SIMU) including a triad of accelerometers and two angular rate sensors is considered. The technique, the calibration bench and algorithms, the results of theoretical error analysis and the results of testing a SIMU prototype based on the micro-electromechanical inertial measurement unit MPU-9250 and two fibre-optic gyroscopes (FOG) FOGS170B are described.**

Key words—strapdown inertial measurement unit, micro-electromechanical accelerometer, fibre-optic gyroscope, calibration, error analysis

I. INTRODUCTION

The determination of the orientation of a longitudinal axes of a stationary measurement device—optical, in particular—relative to the horizontal plane (elevation angle) and to the North direction (azimuth angle) can be implemented with the help of an SIMU including a triad of accelerometers and two angular rate sensors (ARS). Using two ARS instead of three ones allows reducing notably the SIMU cost since accelerometers are relatively low-cost devices whereas the cost of each ARS can be much higher than the cost of the rest of an SIMU. In this case, there is a challenge in correct estimation of the azimuth angle.

A technique of gyrocompassing with the help of such an SIMU was presented at the previous Saint Petersburg International Conference on Integrated Navigation Systems [1]. An SIMU prototype was constructed at the Scientific and Production Association of Automatics on the base of the inertial measurement unit MPU-9250 and two fibre-optic gyroscopes (FOG) FOGS170B, which were installed in a common housing. Errors of estimation with the help of such a device are about 1-2° for an elevation angle, 3-5° for an azimuth angle. To improve accuracy, the SIMU needed calibration, which was implemented with the help of the test equipment of Scientific and Production Association of Electromechanics.

II. SIMU MODEL AND ESTIMATED ERRORS

The SIMU schematic with sensor axes orientation relative to the SIMU body axes is shown in Fig. 1. X_A, Y_A, Z_A are accelerometer sensitive axes (SA); $X_{FOG\text{-}1}, X_{FOG\text{-}2}$ are FOG SA; *OXYZ* is a right-handed orthogonal body frame (BF) fixed rigidly to the mechanical device for connection of the SIMU body to the measurement axis of the optical device, which is a tube.

In addition, introduce an orthogonal sensor frame (SF) $OX_IY_IZ_I$ referenced geometrically to the accelerometer axes X_A and Y_A. Nominally, SF axes coincide with the accelerometer sensitive axes. Misalignments between the BF and SF axes are parameterised by three small angles $\Delta\alpha_x, \Delta\alpha_y, \Delta\alpha_z$, which are estimated during calibration. With zero misalignments, the accelerometer axes coincide with the BF axes, the FOG-1 SA is directed opposite to the axis *OZ*, the FOG-2 SA is directed opposite to the axis *OY*.

The accelerometer error model includes biases $\tau_x^A, \tau_y^A, \tau_z^A$, scale factor errors $\rho_x^A, \rho_y^A, \rho_z^A$ and misalignments $\beta_{yx}^A, \beta_{zx}^A, \beta_{zy}^A$ (orthogonality errors).

Parameters for FOG-1 misalignments are defined as orthogonality errors $\Delta\xi_1, \Delta\eta_1$ between its SA and BF axes *OX*, *OY*; for FOG-2 as orthogonality errors $\Delta\xi_2, \Delta\zeta_2$ between its SA and BF axes *OX*, *OZ*. The FOG error model includes biases τ_1^G, τ_2^G and scale factor errors ρ_1^G, ρ_2^G.

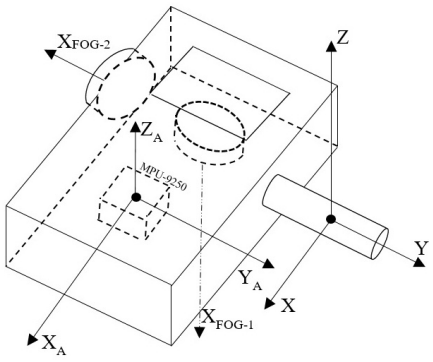


Fig. 1. The SIMU schematic

III. Test Equipment and Arrangement

The test equipment used for SIMU calibration consists of two parts. One of them is single-axis turntable PS-132T with rotation axis of its face plate aligned along the local vertical. The other one is two-axis tilt-and-turn bed SNP-522 with a plate for mounting an SIMU under test. To fix precisely an SIMU on the plate, design engineers of Science and Production of Automatics developed and produced a special fixing device to provide misalignments of a SIMU body relative to the plate about an arc min.

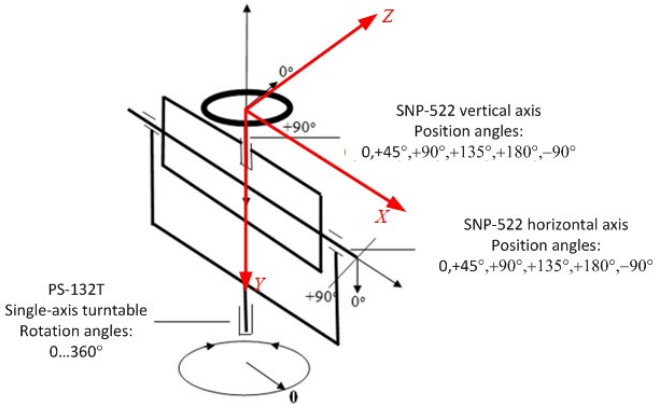


Fig. 2. The test equipment schematic

SNP-522 mounted on the PS-132T face plate provides for alignment of any axis of an SIMU body fixed in it parallel, perpendicular or at the angle 45° relative to the rotation axis of the face plate. The turntable angular rate range is specified from ±1° to ±200°. The relative error in setting the angular rate is about 1%. The test equipment schematic and orientation of an SIMU body axes at zero position angles and at zero rotation angle are shown in Fig. 2.

The tests include static and dynamic parts. In the static tests, the SNP-522 plate with an SIMU mounted on it is rotated and fixed in 11 various preset positions in turn and sensor data are collected during 3-5 minutes. Accelerometer parameters, misalignments between the SF and BF axes, and FOG biases are determined in the static tests.

In the dynamic tests, the BF axes are directed vertically in turn and SNP-522, along with an SIMU fixed in it, is rotated with the help of the PS-132T by an integer number of complete rotations first clockwise and then counterclockwise. The rotation angular rate is constant and equal to 20°/s. FOG misalignments and scale factor errors are estimated in the dynamic tests.

During testing, the sensor outputs are collected in arrays referenced to time points, test type (static or dynamic) and orientation number. The SIMU accuracy parameters are estimated in postprocessing.

IV. Calibration Technique

Estimation of accelerometer parameters

The accelerometer parameters are estimated with the help of a scalar method [2], which does not require precise alignment of their axes relative to a programme orientation. Measurements are carried out in nine BF programme orientations optimum according to the criterion of calibration errors minimum [3]. The SA of each accelerometer is directed along the local vertical upwards and downwards and the SA of each couple of accelerometers are oriented in a vertical plane symmetrically relative to the horizontal plane.

The accelerometer parameters are determined by solving the linearized system of nine equations with nine unknowns

$$Z = H \cdot X \text{, where} \tag{1}$$

$X = \left[\frac{\tau_x^A}{g}, \frac{\tau_y^A}{g}, \frac{\tau_z^A}{g}, \rho_x^A, \rho_y^A, \rho_z^A, \beta_{yx}^A, \beta_{zx}^A, \beta_{zy}^A\right]^T$ is the column matrix of estimated parameters; Z is the column matrix of measurements, its *n-th* element is calculated with the help of the formula $z^{(n)} = \frac{\hat{g}_n^2 - g^2}{2g^2}$, $\hat{g}_n^2 = \hat{g}_{n,x}^2 + \hat{g}_{n,y}^2 + \hat{g}_{n,z}^2$ is the square of the measured vector module in *n-th* orientation; $H_n =$

$$\left[\frac{\hat{g}_{n,x}}{g}, \frac{\hat{g}_{n,y}}{g}, \frac{\hat{g}_{n,z}}{g}, -\frac{\hat{g}_{n,x}^2}{g^2}, -\frac{\hat{g}_{n,y}^2}{g^2}, -\frac{\hat{g}_{n,z}^2}{g^2}, -\frac{\hat{g}_{n,x} \cdot \hat{g}_{n,y}}{g^2}, \frac{\hat{g}_{n,z} \cdot \hat{g}_{n,x}}{g^2}, -\frac{\hat{g}_{n,y} \cdot \hat{g}_{n,z}}{g^2}\right]$$

is *n*-th row of the state-to-measurement matrix *H*; $\hat{g}_{n,x}, \hat{g}_{n,y}, \hat{g}_{n,z}$ are averaged accelerometer outputs.

Estimation of misalignments between the SF and BF axes

To first-order approximation, misalignments between the SF and BF axes are defined by the following expression:

$$\begin{bmatrix} X_I \\ Y_I \\ Z_I \end{bmatrix} \approx \begin{bmatrix} 1 & -\Delta\alpha_z & \Delta\alpha_y \\ \Delta\alpha_z & 1 & -\Delta\alpha_x \\ -\Delta\alpha_y & \Delta\alpha_x & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2)$$

These parameters are estimated using outputs of accelerometers already calibrated. To estimate them, BF axes have to be aligned along the local vertical in turn and it is worth using all the measurements obtained during the calibration of accelerometers. Thus, the expressions for determination of misalignments between the SF and BF axes take the form:

$$\Delta\alpha_x \approx \frac{\hat{g}_{z_I}^{y\uparrow} - \hat{g}_{z_I}^{y\downarrow}}{2g} \approx -\frac{\hat{g}_{y_I}^{z\uparrow} - \hat{g}_{y_I}^{z\downarrow}}{2g};\ \Delta\alpha_y \approx \frac{\hat{g}_{x_I}^{z\uparrow} - \hat{g}_{x_I}^{z\downarrow}}{2g} \approx -\frac{\hat{g}_{z_I}^{x\uparrow} - \hat{g}_{z_I}^{x\downarrow}}{2g};$$

$$\Delta\alpha_z \approx \frac{\hat{g}_{y_I}^{x\uparrow} - \hat{g}_{y_I}^{x\downarrow}}{2g} \approx -\frac{\hat{g}_{x_I}^{y\uparrow} - \hat{g}_{x_I}^{y\uparrow}}{2g}, \text{ where} \quad (3)$$

$\hat{g}_{x_I}, \hat{g}_{y_I}, \hat{g}_{n_I}$ are projections of the measured specific force resolved from the accelerometer axes in the SF axes.

Estimation of FOG biases

According to FOGS170B Datasheet, the scale factor error (1σ) is $5\cdot10^{-5}$, bias is 0.02°/h. Contribution of the scale factor error to the total error of the Earth angular rate is by two orders of magnitude less than that of bias, thus, scale factor errors can be neglected when estimating biases.

The scalar method of calibration cannot be applied to two sensors. Without azimuth reference, the vertical component of the Earth angular rate can be used as a calibration standard. Thus, the FOG sensitive axis, or rather a corresponding BF axis, is aligned with the local vertical. In this orientation, the first measurement is performed, after which the SIMU is rotated about the vertical axis by 180° and then the second measurement is taken. This adds two more orientations to the static tests. Using the sum of measurements in two orientations allows eliminating the horizontal component of the earth angular rate caused by the vertical alignment error due to misalignments between the FOG and BF axes. The measurement error of the vertical component of the Earth angular rate itself is about 0.0019°/h or 0.4% at latitude 55° under FOG axis alignment error 1°. The expressions for FOG bias estimation have a simple form:

$$\tau_k^G = \frac{1}{2}\left(\hat{\Omega}_k^{(1)} + \hat{\Omega}_k^{(2)}\right) - \Omega_H,\ k = 1,2, \text{ where} \quad (4)$$

$\hat{\Omega}_k^{(1)}, \hat{\Omega}_k^{(2)}$ are averaged FOG outputs in the first and second orientations; Ω_H is the vertical component of the Earth angular rate; k is a FOG number.

Estimation of FOG scale factor errors and misalignments

The FOG errors depending on the measured angular rate are estimated using the rotation of the SIMU mounted on the SNP-522 plate about each BF axis, which is initially aligned vertically downwards.

While the SIMU is rotating about the axis X, misalignments of FOG-1 and FOG-2 axes relative to the BF axis X $\Delta\xi_1$ and $\Delta\xi_2$ are estimated; during the SIMU rotation about the axis Y, estimates of misalignment between FOG-1 axis and the BF axis Y $\Delta\eta_1$ and the FOG-2 scale factor error ρ_2^G are obtained; and when it is rotated about the axis Z, we obtain estimates of misalignment between the FOG-2 axis and the BF axis Z and the FOG-1 scale factor error ρ_1^G.

The standard for FOG calibration is a turn angle about the PS132-T rotation axis. The Earth angular rate measured by the FOG along with the rotation velocity is a disturbance. To minimize the influence of a transverse component of the Earth angular rate, it is required (1) to rotate an SIMU with a constant angular rate; (2) to turn it by an integer number of complete rotations.

The influence of the longitudinal component of the Earth angular rate can be eliminated by the SIMU rotation about the chosen axis, first, counterclockwise (+) and then clockwise (–). The expressions for estimation of FOG misalignments have a simple form as well:

$$\Delta\chi = \frac{\hat{\varphi}^+ - \hat{\varphi}^-}{n\cdot 4\cdot\pi};\ \hat{\rho}_k^G = \frac{n\cdot 4\cdot\pi}{\hat{\varphi}^+ - \hat{\varphi}^-} - 1, \text{ where} \quad (5)$$

$\hat{\varphi}^+, \hat{\varphi}^-$ is an integral of the absolute angular velocity measured by the FOG over the interval of the FOG rotation about the corresponding axis counterclockwise and clockwise; n is the number of complete rotations; $\Delta\chi = \Delta\xi_1, \Delta\eta_1, \Delta\xi_2, \Delta\zeta_2$; $k = 1,2$.

V. Calibration Error Analysis

Verification of the technique and algorithms and theoretical error analysis of calibration were implemented with the help of a simulation mathematical model. The test procedure taking into account the errors of the test equipment, averaged sensor outputs collected during the tests and the algorithms for SIMU accuracy parameters estimation in the post-testing data processing were simulated.

Calibration errors were estimated with the help of statistical simulation based on 10,000 runs. Estimated parameters and equipment errors were simulated as constant random variables with zero mean values uniformly distributed from run to run. Values of 3σ (σ is the standard deviation) were assigned for accelerometer biases $5\cdot10^{-3}$ m/s^2, scale factor errors $5\cdot10^{-2}$% and misalignments 1.5′; for FOG biases 0.06°/h and scale factor errors $5\cdot10^{-3}$%; for misalignments between the body axes and the accelerometer axes 0.5° and the body axes and the FOG axes 1°.

The test equipment errors (3σ) were simulated as misalignments between the turntable body and the geographic frame (30″), between the SIMU body and the SNP-522 plate (1′), and between the SNP-522 plate and the required programme orientation (6″).

The simulation results have shown that the testing equipment can ensure the accuracy of three correct significant decimal digits for estimation of accelerometer parameters, the accuracy of two correct significant decimal digits for FOG biases, scale factors and misalignments of sensor axes to a SIMU body.

VI. Test results

Two series of tests with an interval of one month were carried out on the test equipment of Scientific and Production Association of Electromechanics. The values of accuracy parameters of the accelerometers and FOG obtained as a result of calibration have not exceeded the manufacturer's specifications for MPU-9250 and FOGS170B devices in both series (runs).

Integral estimation of accelerometer calibration accuracy can be obtained by checking the module of the measured vector of the normal reaction to the gravity, since in this case we have complete set for vector measurements. Such estimation has shown that the relative errors of the measured vector modules averaged over eleven orientations prior to the calibration were 2.15% in the first run, 2.46% in the second run; and after the calibration 0.11%. Thus, the integral error of specific force measurement has decreased by 20 times.

Using only two FOG does not allow applying this approach to estimate FOG calibration accuracy in the general case as well as using the scalar method for FOG calibration. Yet fine azimuth alignment of the turntable was carried out during testing. The outer horizontal axis of the turntable was directed to the East with accuracy of a few arc min. As a result, it was possible to choose SNP-522 plate orientations in a way that the FOG sensitive axes were oriented in the meridian plane. That allowed determining the module of a measured vector by its two projections and applying the scalar method along with the vector one to estimation of FOG biases. The results of estimation of FOG biases are given in Table I.

Table I

Results of estimation of FOG biases				
Parameters	Vector calibration		Scalar calibration	
	First series of tests	Second series of tests	First series of tests	Second series of tests
τ_1^G, °/h	–0.0140	–0.0105	–0.0486	–0.0361
τ_2^G, °/h	–0.0810	–0.0195	–0.1504	–0.1178

The results of the scalar calibration turned out to be somewhat different from the results of the vector calibration. The relative error of the measured module of the Earth angular rate, which was 67% prior to the calibration, after calibration became 33% for the vector method and 15% for the scalar method. Thus, the FOG biases obtained in the scalar calibration were used as compensation coefficients in the algorithms of the basic task, which is determination of orientation angles.

Misalignments between the accelerometer axes and the SIMU body axes turned out to be more than expected 30′ and made up 80-160′; misalignments between the FOG axes and the body turned out to be less than expected 1° and made up 6-48′.

The SIMU tests in the basic mode were implemented with the elevation angle equal to 45° and several values of azimuth angle. The errors of alignment of the SNP-522 plate with the SIMU mounted on it with respect to the horizontal plane and the North direction were about several minutes. Due to sensor noise and high sensitivity of FOG to external disturbances and vibration, the errors of determination of orientation, especially for azimuth angle, depend on time of the sensor output averaging. When averaging over the interval of 280 s, the calibration allows reducing an elevation angle estimation error from 2° to 0.2°, and an azimuth angle estimation error from 5° to 0.25°. When averaging over the intervals of 30 s and 5 s, an elevation angle estimation error remains almost the same, whereas the azimuth angle estimation error increases notably: for the 30 s interval of up to 1°, for the 5 s interval of up to 2°.

VII. Conclusion

The technique developed for calibration of an SIMU including a triad of accelerometers and two FOG allows estimating accelerometer parameters with three correct significant decimal digits, FOG biases and misalignments between the sensor axes and the SIMU body with two correct significant decimal digits.

The results of testing the SIMU in the basic mode have shown that the calibration results in decreasing the elevation angle estimation error from 2 to 0.2°, the azimuth angle estimation error from 5 to 0.25°.

References

[1] Vodicheva, L, Yakimov, V., Parysheva, Yu., and Kabanova, D. Gyrocompassing Technique for a Strapdown Inertial Measurement Unit with Two Angular Rate Sensors, 31st St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2024, pp. 375–378.

[2] Izmailov, E.A., Lepe, S.N., Molchanov, A.V., and Polikovsky, E.F. Scalar Method for Calibrating and Balancing Strapdown Inertial Navigation Systems, 15th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2008, pp. 151–159.

[3] Vodicheva, L. and Parysheva, Yu. Estimating the Accuracy Parameters of Sensors in a Strapdown Inertial Measurement Unit with the Use of a Relatively Coarse Turntable, Gyroscopy and Navigation, 2019, Vol. 10, No. 4, pp. 303-312.

Adaptive Robust Filtering and Calibration of Inertial MEMS - Sensor Signals

A.V. Chernodarov
"NaukaSoft" Research & Production Association, Ltd., MAI University, Moscow, Russia
e-mail: chernod@mail.ru

P.S. Gorshkov
"NaukaSoft" Research & Production Association, Ltd., Moscow, Russia
e-mail: contacts@naukasoft.ru

I.V. Mitrofanov
"NaukaSoft" Research & Production Association, Ltd., Moscow, Russia,
e-mail: contacts@naukasoft.ru

A.P. Patrikeev
"NaukaSoft" Research & Production Association, Ltd., Moscow, Russia,
e-mail: apatrikeev@naukasoft.ru

Abstract **— This paper is devoted to improving the technology of bench calibration of micro-electro-mechanical sensors (MEMS) for inertial measurement units (IMU). A combined calibration procedure is considered, including processing of MEMS signals in the frequency and time domains. Filtering in the frequency domain makes it possible to eliminate the influence of noise on the accuracy of estimating systematic MEMS errors. Estimation of MEMS systematic errors in the time domain allows the use of mathematical models of signal drifts to improve their observability relative to bench standards. At the same time, this allows for parametric identification of MEMS error models, which are subsequently used during the operation of the IMU. It is proposed to estimate the systematic errors of MEMS by processing the residuals of their signals relative to the reference values for the three IMU positions on the bench using a sequential modification of the extended Kalman filter (EKF). In comparison with the traditional calibration procedure based on the least squares method, the use of such a modification of the EKF does not require matrix inversion. The results of semi-natural experiments based on the recorded MEMS data are analyzed.**

Keywords — inertial measurement unit, micro-electro-mechanical sensors, bench calibration, signal filtering, parametric identification, extended Kalman filter.

I. INTRODUCTION

At present, the problem of reducing the weight, dimensions and cost of navigation systems (NS) while maintaining their accuracy characteristics remains relevant. A possible solution to this problem is associated with the inclusion of inertial measurement units (IMU) based on micro-electro-mechanical systems (MEMS) and sensors [1] in the NS. However, MEMS have low accuracy due to large drifts in the signals of the sensors included in their composition: gyros and accelerometers located along three orthogonal axes. It seems that the contradiction between the required accuracy characteristics and the required weight and dimensions of MEMS can be resolved, including on the basis of improving the calibration technology of the sensors. Systematic drifts of sensor signals are estimated, as a rule, during bench calibration. The accuracy of such calibration is significantly affected by random noise and failures in sensor signals. Therefore, adaptive signal filtering and robust failure parrying must be performed during the calibration process. The adaptation is related to the adjustment of the filter coefficients based on the results of observations. Robust failure parrying is based on the use of confidence coefficients for observations. The differences between known and measured geophysical invariants are used as observations for sensor calibration. The following parameters can be invariants:

- gravity acceleration;
- projections of the angular velocity vector of the Earth's rotation onto the axes of the IMU placed on the stand.

During bench calibration of the MEMS, observations are formed so that each of the three IMU axes is sequentially in the vertical position. As a rule, 18 parameters are estimated, namely: systematic biases of accelerometer signals and gyros drifts, as well as angular deviations of the measuring axes of the sensors from their mutually orthogonal positions.

The traditional formulation of the calibration problem [2,3] is associated with the estimation of systematic biases of the sensor signals and the identification of the elements of angular skews matrix of the measuring axes of the IMU. Simultaneous solution of the problem of estimation and identification by precise computational methods is not always possible. In this paper, models of calibration procedures are formed in the form of state observers for each IMU position. In such models, observations of invariants are related to the IMU error vector through a matrix of coefficients dependent on sensor signals. The use of such models makes it possible to solve the calibration problem using the least squares method [2,3]. However, in this case it is necessary to invert the coefficient matrix in the generalized observation model of dimension 18×18, which is an inaccurate and unstable operation. To avoid such an operation, the calibration problem is transformed into a form that allows the use of a sequential modification of the extended Kalman filter (EKF) [4] to assess the systematic errors of the sensors. In this modification, the operation of inverting the 18×18 covariance matrix of observation errors

is reduced to a sequence of 18 division operations. In addition, the following additional calibration features can be implemented with EKF:

- sensors noise filtering in the frequency-time domain, estimation and consideration of their statistical characteristics;
- an adaptive robust protection of sensor calibration procedures from abnormal observations [5];
- use of zero velocity observations and known stand coordinates to improve the accuracy of sensor error estimation [6];
- identification of drift and diffusion coefficients for dynamic sensor error models [5].

The purpose of the paper is to study the procedures for calibration the IMU using bench standards and the EKF for estimation the errors of MEMS sensors.

II. Bench Calibration of an Inertial Measuring Unit Based on MEMS Sensors

During bench calibration of an IMU, the errors vector, as a rule, includes systematic biases of signals from inertial sensors, namely gyros and accelerometers, as well as angular skews of the axes of the measuring trihedrons relative to their ideal orthogonal positions. For example, for accelerometers included in the IMU, the errors vector will have the form

$$x_{a\delta} = [x_a^{\mathrm{T}} x_{\delta a}^{\mathrm{T}}]^{\mathrm{T}}, \quad (1)$$

where $x_a = [\Delta a_x \Delta a_y \Delta a_z]^{\mathrm{T}}$ (2)

is a vector of systematic biases of accelerometers signals relative to their reference a_x, a_y, a_z values; $ox_r y_r z_r$ is a coordinate system associated with IMU. For the specified coordinate system, we accept a sequence of rotations that ensures the alternate arrangement of accelerometers in a vertical position, namely: in the first position, the oz_r axis of the IMU is vertical, in the second position, the oy_r axis is vertical, in the third position, the ox_r axis is vertical;

$$x_{\delta a} = [\delta_{a(xy)} \delta_{a(xz)} \delta_{a(yx)} \delta_{a(yz)} \delta_{a(zx)} \delta_{a(zy)}]^{\mathrm{T}} \quad (3)$$

is the vector of angular skews of the measuring axes of accelerometers relative to their mutually orthogonal positions, as shown in Fig. 1 for MEMS MPU-6050 [7].

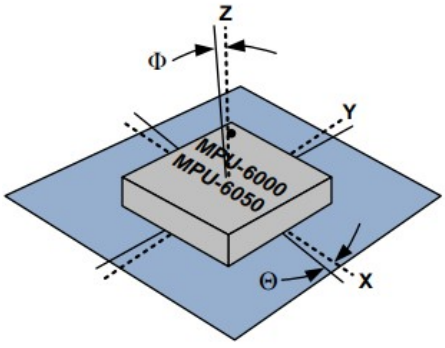


Fig. 1. . Angular skews of the IMU axes

The technology for calibrating MEMS sensors for three IMU positions on a single-axis stand is considered. It is assumed that the ideal measuring trihedron *oxyz*, connected to the turn table of the stand, has orthogonal axes. In addition, the *oz* axis is directed vertically, and the *ox* and *oy* axes are located in the horizontal plane. The mutual orientation of the $ox_r y_r z_r$ and *oxyz* trihedrons is determined by the angles ψ, ϑ, γ. In this instance, we assume that the original directions of axes of the above frames coincide, and turns are made counterclockwise as follows: the first turn is done by an angle of ψ radn. about the third axis *oz*; the second turn is performed by an angle of ϑ radn. about a new position of the first axis ox'; the third turn is done through an angle of γ radn. about a new position of the second axis oy'.

When the oz_r axis of the IMU is installed in a vertical position on the horizontal plane of the turn table of the stand, the accelerometers signals model will have the form

$$\begin{bmatrix} a_{x(r)} \\ a_{y(r)} \\ a_{z(r)} \end{bmatrix} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} + \begin{bmatrix} \Delta a_x \\ \Delta a_y \\ \Delta a_z \end{bmatrix} + \begin{bmatrix} 0 & \delta_{xy} & \delta_{xz} \\ \delta_{yx} & 0 & \delta_{yz} \\ \delta_{zx} & \delta_{zy} & 0 \end{bmatrix} \begin{bmatrix} a_{x(r)} \\ a_{y(r)} \\ a_{z(r)} \end{bmatrix}, \quad (4)$$

where $[a_x a_y a_z]^{\mathrm{T}} = [0\ 0\ g]^{\mathrm{T}}$ is the vector of accelerometers signals for the reference IMU; *g* is the acceleration of gravity at the location of the stand; $[a_{x(r)} a_{y(r)} a_{z(r)}]^{\mathrm{T}}$ is a vector of signals from real accelerometers.

Taking into account equation (4), we can form observations Z_a for the accelerometers errors vector with the such orientation of the IMU:

$$Z_{a(z)} = \begin{bmatrix} a_{x(r)} \\ a_{y(r)} \\ a_{z(r)} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = C_{a\psi} x_{a\delta} + \vartheta_{a\psi}, \quad (5)$$

where

$$C_{a\psi} = \begin{bmatrix} 1 & a_{y(r)} & a_{z(r)} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{x(r)} & 1 & a_{z(r)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{x(r)} & a_{y(r)} & 1 \end{bmatrix}$$

$x_{a\delta} = [\Delta a_x \delta_{xy} \delta_{xz} \delta_{yx} \Delta a_y \delta_{yz} \delta_{zx} \delta_{zy} \Delta a_z]^{\mathrm{T}}$ is an expanded view of the error vector of accelerometers as part of the IMU; $\vartheta_{a\psi} = [\vartheta_{a_{\psi x}} \vartheta_{a_{\psi y}} \vartheta_{a_{\psi z}}]^{\mathrm{T}}$ is a vector of observation errors associated, for example, with deviations of the stand from the horizontal position.

From equation (5), an approximate estimate of the accelerometers errors vector can be found using the least squares method [3]:

$$\hat{x}_{a\delta} = C^*_{a\psi} Z_{a(z)}, \quad (6)$$

where $C^*_{a\psi} = (C^{\mathrm{T}}_{a\psi} C_{a\psi})^{-1} C^{\mathrm{T}}_{a\psi}$ is a pseudo inverse matrix.

For a correct solution to the estimation problem (6), it is necessary that the $C_{a\psi}$ matrix, which takes into account the relationship between the observation vector $Z_{a(z)}$ and the accelerometers errors $x_{a\delta}$, be square. This can be realized by additionally performing two more measurement procedures, sequentially placing the IMU vertically along the oy and ox axes.

When setting the axis oy_r to a vertical position by rotating the IMU around the *ox* axis by an angle $\vartheta = \pi / 2$, the accelerometers signals model will have the form

$$P_1 [a_{x(r)} a_{y(r)} a_{z(r)}]^{\mathrm{T}} = [0\ 0\ \mathrm{g}]^{\mathrm{T}} + C_{a\vartheta} x_{a\delta}\ , \qquad (7)$$

where $C_{a\vartheta} = P_1 C_{a\psi}$; $P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ is a matrix that takes into account the permutation of the measuring axes of the IMU. Equation (7) is written taking into account the fact that the physical orientation of the reference trihedron associated with the horizontal position of the turntable of the stand does not change.

Taking into account equation (7), it is possible to form the observation vector $Z_{a(y)}$ for the errors of accelerometers with the orientation of the IMU $oxy'z'$:

$$Z_{a(y)} = \begin{bmatrix} a_{x(r)} \\ a_{z(r)} \\ a_{y(r)} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = C_{a\vartheta} x_{a\delta} + \vartheta_{a\vartheta}\ . \qquad (8)$$

When setting the axis ox_r to a vertical position by rotating the IMU around the *oy* axis by an angle $\gamma = -\pi / 2$, the accelerometers signals model will have the form

$$P_2 [a_{x(r)} a_{y(r)} a_{z(r)}]^{\mathrm{T}} = [0\ 0\ \mathrm{g}]^{\mathrm{T}} + C_{a\gamma} x_{a\delta}\ , \qquad (9)$$

where $C_{a\gamma} = P_2 C_{a\psi}$; $P_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

Taking into account equation (9), it is possible to form the observation vector $Z_{a(y)}$ for the errors of accelerometers with the orientation of the IMU $ox'yz'$:

$$Z_{a(x)} = \begin{bmatrix} a_{z(r)} \\ a_{y(r)} \\ a_{x(r)} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = C_{a\gamma} x_{a\delta} + \vartheta_{a\gamma}\ . \qquad (10)$$

Taking into account equation (5), (8), and (10), an extended observation vector Z_a for accelerometers errors can be formed:

$$Z_a = H_a x_{a\delta} + \vartheta_a\ , \qquad (11)$$

where $Z_a = [Z_{a(z)}^{\mathrm{T}} Z_{a(y)}^{\mathrm{T}} Z_{a(x)}^{\mathrm{T}}]^{\mathrm{T}}$; $H_a = [C_{a(\psi)}^{\mathrm{T}} C_{a(\vartheta)}^{\mathrm{T}} C_{a(\gamma)}^{\mathrm{T}}]^{\mathrm{T}}$; $\vartheta_a = [\vartheta_{a\psi}^{\mathrm{T}} \vartheta_{a\vartheta}^{\mathrm{T}} \vartheta_{a\gamma}^{\mathrm{T}}]^{\mathrm{T}}$.

The H_a matrix is square and has a dimension of 9×9. Therefore, the solution of equation (11) with respect to the error vector $x_{a\delta}$ will be correct.

To calibrate the MEMS angular velocity sensors (AVS), it is necessary to perform reference rotations of the IMU using a rotary table sequentially relative to each of the three measuring axes. To estimate the errors of the AVS in equations (4) - (11), instead of the acceleration of the force gravity *g*, the reference angular velocity ω of the turntable rotation is set. In addition, the vertical component of the angular velocity of the Earth's rotation $\Omega \sin\varphi$ is taken into account. Instead of accelerometers signals, the corresponding AVS readings are recorded.

By analogy with equations (4) – (11), when calibrating the IMU, an $H_\omega = [C_{\omega(\psi)}^{\mathrm{T}} C_{\omega(\vartheta)}^{\mathrm{T}} C_{\omega(\gamma)}^{\mathrm{T}}]^{\mathrm{T}}$ matrix can be formed to link the $Z_\omega = [Z_{\omega(z)}^{\mathrm{T}} Z_{\omega(y)}^{\mathrm{T}} Z_{\omega(x)}^{\mathrm{T}}]^{\mathrm{T}}$ observations with the error vector of the AVS.

It should be noted that the operation of inverting a extended $H = \begin{bmatrix} H_a \\ H_\omega \end{bmatrix}$ matrix of 18×18 dimension may be unstable when processing $Z = \begin{bmatrix} Z_a \\ Z_\omega \end{bmatrix}$ observations. In addition, when solving the $Z = Hx + \vartheta$ equation with respect to the $x = [x_{a\delta}^{\mathrm{T}} x_{\omega\delta}^{\mathrm{T}}]^{\mathrm{T}}$ vector, the traditional least squares method does not take into account the noises $\vartheta = [\vartheta_a^{\mathrm{T}} \vartheta_\omega^{\mathrm{T}}]^{\mathrm{T}}$ for observations.

To exclude the above limitations, it is proposed to process observations Z using the adaptive robust (AR) modification [5] of the EKF.

Prediction: $m_0 = x_{i/i-1} = \Phi_i \hat{x}_{i-1/i-1}$;

$$M_0 = P_{i/i-1} = \Phi_i P_{i-1/i-1} \Phi_i^T + \Gamma_i Q_{i-1} \Gamma_i^T\ .$$

Updating: $\nu_j = z_j - H_j m_{j-1}$; $\beta_j = \nu_j / \alpha_j$;

$$\psi_j = \psi(\beta_j)\ ;\ \psi'_j = \psi'(\beta_j)\ ;$$

$$\tilde{\alpha}_j^2 = H_j M_{j-1} H_j^T \psi'_j + \alpha_j^2\ ;$$

$$K_j = M_{j-1} H_j^T / \tilde{\alpha}_j^2\ ;$$

$$m_j = m_{j-1} + K_j \psi_j \alpha_j\ ;$$

$$M_j = (K_j H_j \psi'_j - I) M_{j-1} (K_j H_j \psi'_j - I)^T + K_j \alpha_j^2 \psi'_j K_j^T\ ;$$

$$j = \overline{1, l}\ ;\ \hat{x}_{i/i} = m_l ; P_{i/i} = M_l\ ,$$

where $\psi(\beta_j) = \left. \frac{\partial \rho(\beta)}{\partial \beta} \right|_{\beta=\beta_j}$; $\psi'(\beta_j) = \left. \frac{\partial^2 \rho(\beta)}{\partial \beta^2} \right|_{\beta=\beta_j}$;

$\rho(\beta) = -\ln f(\beta)$; $f(\beta)$ is a probability density function. The functions $\psi(\beta_j)$ and $\psi'(\beta_j)$ can be formed with due regard for «a priori» assumptions as to the distribution laws of the valid signal and noise as shown in [5]; m_j, $\hat{x}_{i/i}$ are the estimates of the state vector *x* at the *i*-th step after the *j*-th

component and the whole vector Z_i, of observations are processed, respectively; M_j, $P_{i/i}$ are the covariance matrices of errors of the estimates mentioned; in the problem under consideration, $l = 18$; I is an identity matrix; Φ_i, Γ_i are transition matrices for the state vector x_i and the disturbance vector ξ_i; Q_i is a covariance matrix for the disturbance vector; α_j is a scaling parameter, which can be, for example, in the EKF, the root mean square value of measurement errors $\sqrt{R_j}$.

III. Analysis of the Results of Studies

The IMU based on MEMS MPU-6050 [7] is considered as an object of studies (see Fig. 2). IMU signals are recorded on flash memory via a STM32 microcontroller [8] with a frequency of 500 Hz. After calibration, full-scale experiments were carried out on a car trip, taking into account the following conditions:

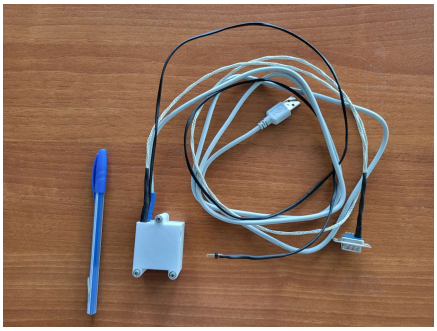


Fig. 2. IMU based on MEMS MPU-6050

• based on the recorded data, an algorithm of the strapdown inertial navigation system (SINS) was implemented in post-processing;

• when reckoning the coordinates, the estimates of the systematic errors of the MEMS obtained during calibration were written off from the sensor readings.

It was assumed that the global navigation satellite system (GNSS) was unavailable for SINS correction at the final section of the trajectory for 400 seconds. Inertial reckoning of parameters was performed at this section.

Figs. 3 and 4 show the circular positional errors ΔS of the inertial reckoning of geodetic coordinates: in Fig. 3, without taking into account the estimates of the calibration coefficients; in Fig. 4, taking into account and predicting such estimates, where

$$\Delta S = \sqrt{\delta_\varphi^2 + \delta_\lambda^2}\,;\ \delta_\varphi = (\varphi_{SINS} - \varphi_{GNSS})R_\varphi\,;$$

$\delta_\lambda = (\lambda_{SINS} - \lambda_{GNSS})R_\lambda \cos\varphi_{GNSS}$; φ and λ are respectively geodetic latitude and longitude; R is the radius of curvature of the earth's ellipsoid at the SINS location.

It can be seen that compensation of the predicted SINS-MEMS error estimates in the section of autonomous inertial

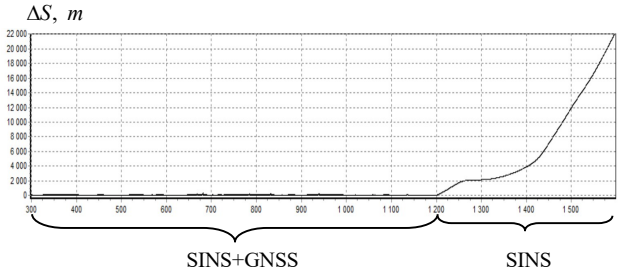


Fig. 3. Circular position error of SINS without taking into account the estimates of calibration coefficients

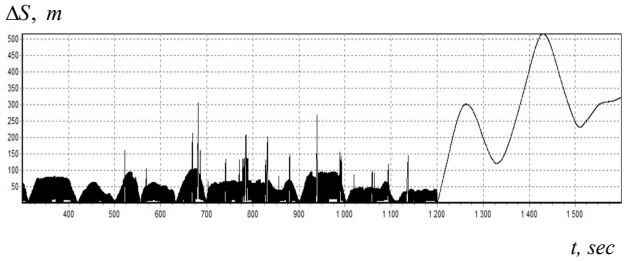


Fig. 4. Circular position error of SINS taking into account estimates of calibration coefficients

dead reckoning of flight parameters made it possible to reduce the SINS circular position error by almost two orders of magnitude on a time interval of 400 seconds.

IV. Conclusions

The conducted studies have shown the possibility of performing combined calibration of MEMS-based IMUs. Such calibration may include the following procedures:

• filtering and parrying abnormal sensor signals;

• estimation of systematic errors and angular skews of the measuring axes of sensors using AR EKF.

References

[1] Wang, D., Dong, Y., Li, Z., Li, Q., and J. Wu, J., Constrained MEMS-Based GNSS/INS Tightly Coupled System with Robust Kalman Filter for Accurate Land Vehicular Navigation, IEEE Trans. Instrum. Meas., vol. 69, no. 7, 2020, pp. 5138–5148.

[2] Titterton, D., and Weston, J., Strapdown Inertial Navigation Technology. Second Edition. Progress in Astronautics and Aeronautics Series, vol. 207. AIAA, Reston, 2004.

[3] Akimov, P., Derevyankin A., and Matasov, A., Garantiruyushchij podhod i ll - approksimaciya v zadachah ocenivaniya parametrov BINS pri stendovyh ispytaniyah (Guaranted Approach and ll Approhimation in the Problems of SINS Parameter Estimation Under Bench Testing), Moscow: Moscow University Press, 2012.

[4] Maybeck, P.S., Stochastic Models, Estimation and Control, New York., Academic Press, 1982, vol. 2.

[5] Chernodarov, A.V., Kontrol, diagnostika i identifikatsiya aviatsionnykh priborov i izmeritelno-vychislitelnykh kompleksov (Monitoring, Diagnostics, and Identification of Aviation Instruments and Measuring-and-Computing Complexes), Moscow, Nauchtekhlitizdat, 2017.

[6] Chernodarov, A., Patrikeev, A.P., and Perelyaev, S.E., "Inertial Navigation and Geophysical Invariants," 28th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2021, IEEE, 9470815.

[7] MPU-6050 Data Sheet 3.4. chipdip.ru›product/mpu-6050

[8] STM32. ru.wikipedia.org›STM32

Accuracy Improvement of SINS on Laser Gyroscopes Under Vibration

S.V. Topilskaya
Inertial System Development and Maintenance Department "Mars" Moscow Experimental Design Bureau
Moscow, Russian Federation
office@mokb-mars.ru

A.O. Kolesnikov
Inertial System Development and Maintenance Department "Mars" Moscow Experimental Design Bureau
Moscow, Russian Federation
office@mokb-mars.ru

Yu.V. Bakin
Inertial System Development and Maintenance Department "Mars" Moscow Experimental Design Bureau
Moscow, Russian Federation
office@mokb-mars.ru

***Abstract* – Laser gyroscopes with mechanical vibrating support are widely applied in aircraft-grade strapdown inertial navigation systems (SINS). Modern SINSs are required to ensure highly accurate autonomous determination of piloting and navigation parameters under high operational impact. The accuracy performance study for the SINS exposed to vibratory impacts has revealed that the tolerance errors of the SINS are attributable to the vibrational errors of laser gyroscopes in the range of frequencies between 700 and 900 Hz. This paper represents design options of a laser gyroscope vibro-suspension that decrease the vibrational error of the gyroscope.**

Keywords: laser gyroscope, mechanical impacts, inertial instruments, vibro-suspension.

I. INTRODUCTION

Modern SINS are exposed to rather high mechanical impacts (shocks, vibrational impacts) during use. Under the mechanical impacts, the random errors of inertial sensors are supplemented by vibrational errors, which leads to additional errors in determination of piloting and navigation parameters [1]. As a result of this study, it was found that the growth of the orientation angle error is due to the occurrence of vibration drift of the laser gyroscope, arising out of the available critical vibration amplitudes in the vibro-suspension structure in the high frequency region (Fig.1).

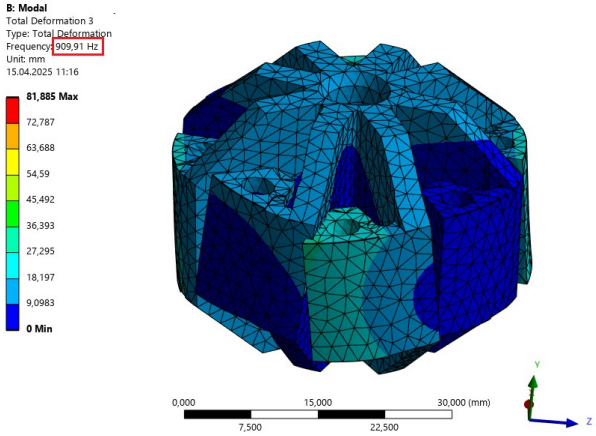


Fig. 1. Form of critical amplitudes of vibro-suspension oscillations

This work is aimed at developing a vibro-support design that will minimize or eliminate the gyroscope oscillation amplitudes in the high frequency range of vibrational impact.

II. STUDY OF THE RESULTS OF THE CONDUCTED RESEARCH

Mars MEDB, Branch of the Federal State Unitary Enterprise Dukhov Automatics Research Institute (VNIIA), developed and manufactured a SINS on laser gyroscopes. Upon testing of the SINS stability under vibration exposure, the accuracy parameters of the SINS were found to misalign with the terms of reference. The core element of the SINS structure is a sensor package (SP) containing gyroscopes and accelerometers, service electronics, and shock absorption elements [2]. The SP design is shown in Fig. 2.

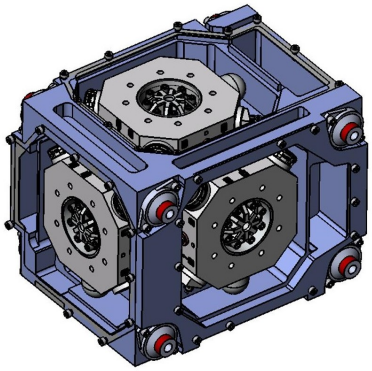


Fig. 2. SINS sensor package

The theoretical research yielded the probable causes of the growing error of the SINS orientation under the influence of standard vibration σ=3g with exposure time t=1h:

1 Poor dampening of the SP oscillations by the shock absorption system in the high frequency range of vibrational impact;

2 An offset of the SP's center of masses relative to the shear center of the shock absorption system;

3 Insufficient accuracy of computation of the SP orientation;

4 Conical movement of the laser gyroscope's vibro-suspension in the high frequency range of the vibrational impact.

The experimental research of the identified causes has defined that a higher level of dampening of the SP oscillations in high frequency range, matching the center of masses of the SP with the shear center of the shock absorption system and a higher data output rate of the gyroscope had all failed to noticeably improve the accuracy performance of the SINS under the standard vibration.

However, the rework of the SP software in terms of raising the gyroscope data output rate allowed for the more detailed exploration of the signals from the gyroscopes (Fig.3) and the determination of the attribution of the growing orientation angle error to a vibrational drift of the laser gyroscope, which occurs due to the conical movement of the gyroscope and is rather difficult to consider algorithmically. This movement is implemented by an angular oscillation of the "blades" of the vibro-suspension (Fig.1).

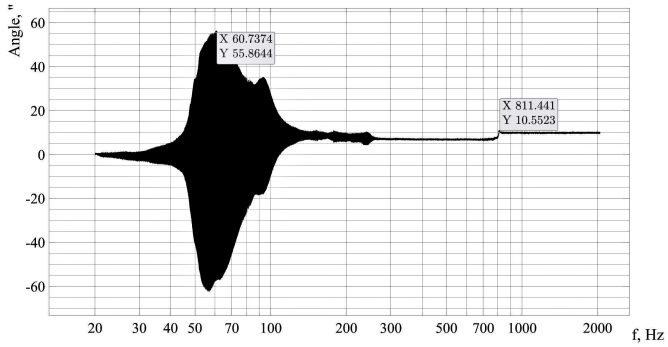


Fig. 3. Vibrational drift of a gyroscope

III. LASER GYROSCOPE SIMULATION MODEL

The sensitive element of the SINS in question is a ring laser [3] with 280 mm perimeter. The ring laser is designed in a single monoblock. The outputs of the ring laser are two sine signals, whose frequencies are identical and contain information on the projection of the angular speed of rotation onto the sensitive axis of the gyroscope, and the phase difference sign is defined by the sign of projection of the angular speed.

To excite oscillations of the ring laser, a vibro-suspension with piezoelements is attached to the ring laser through a compensating ring and used. This ensures that the gyroscope moves in a single plane. The resonance frequency of the "vibro-suspension – ring laser" system is 380±20 Hz [3].

The research included numerical simulation of the movement of a laser gyroscope with a vibro-support under the influence of wide-band random vibration [4]. The resultant simulations have shown that the vibro-suspension, in addition to the oscillations on the basic frequency of 380±20 Hz, flexurally oscillated on the doubled intrinsic frequency of the vibro-suspension (Fig.1), which lead to conical movement of the sensitivity axis of the laser gyroscope relative to the vibro-suspension's axis.

Since it is difficult to measure the doubled frequency movement with the desired accuracy and to take it into account algorithmically, a new design of the laser gyroscope vibro-suspension was suggested for the development. The new vibro-suspension design should ensure the absence of oscillation critical amplitudes and allow for the installation of gyroscope in the SP structure without any modifications.

Based on the formulated requirements, two options of the vibro-suspension design were developed (Fig.4). The first vibro-suspension option allows to decrease the oscillation amplitudes of the "ring laser – vibro-suspension" mechanical system at the doubled frequency of the vibro-suspension. The second vibro-suspension design option eliminated them completely. The second option differs from the first one in the absence of a compensating ring, which improves the rigidity of suspension of the vibro-suspension doubled frequency.

Each computation model (Fig.5) of the vibro-suspension was a weight and dimension equivalent of the resonator, as installed onto the vibro-suspension with and without the compensating ring (for the second design option of the vibro-suspension).

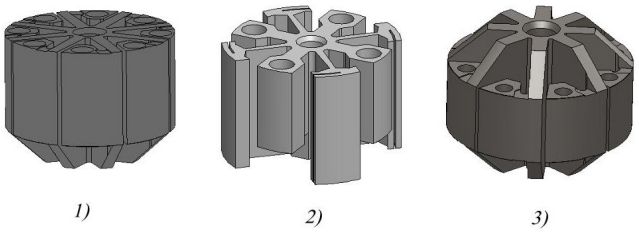


Fig. 4. Vibro-suspension design options, where 1 is the first design option, 2 – the second one; 3 – the initial (standard) version

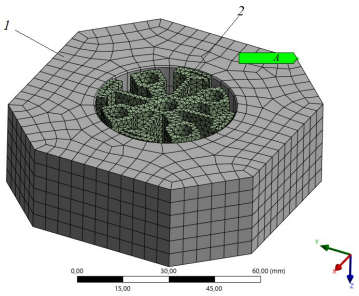


Fig. 5. Vibro-suspension computation model, where 1 is simulated resonator, 2 – the vibro-suspension; 3 – A is a representative point

IV. COMPUTATION RESULTS

The research included the estimations of the vibro-suspension's intrinsic frequencies and movement in representative points of the structure under the σ=3g standard vibration exposure.

The second oscillation mode for the first design option is 1396 Hz; for the second one – 1650 Hz (Fig.6). The higher frequency of the oscillation second mode helps materially reduce the oscillation amplitudes of the gyroscope in the high frequency range under the standard vibration (Fig.7).

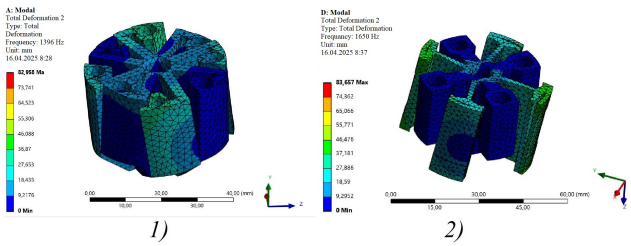


Fig. 6. Computation results of the oscillation second modes, where 1 is the first design option, 2 – the second one;

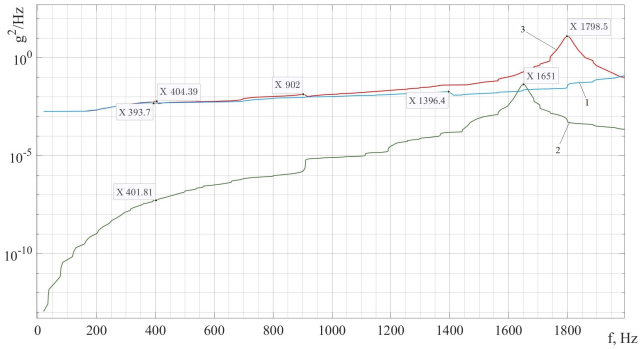


Fig. 7. Gyroscope oscillation amplitudes in representative points, where 1 is for the first option, 2 – for the second one; 3 – for the initial version.

It is suggested to experimentally determine the actual effect of higher rigidity of the vibro-suspension on the accuracy performance of the SINS by way of a separate process specimen using the developed program of research.

V. CONCLUSIONS

The main results of the work are:

- determination of the causes of growing errors of SINS built on laser gyroscopes under the vibration, based on the results of the previous researches;
- selection of the most suitable method to reduce the vibrational errors of the SINS, based on the results of the previous researches.

The suggested option to minimize the SINS errors is to create an upgraded gyroscope design version that will preclude the critical amplitudes of oscillations for the SINS design under the vibration.

REFERENCES

[1] Analytical review of state-of-the-art research and development in the area of navigation abroad. Ed.2. St. Petersburg, Concern CSRI Elektropribor JSC, 2018.

[2] Rivkin S.S., Berman Z.M., Okon I.M. Determination of parameters of orientation of an object by a strapdown inertial system. St. Petersburg, CSRI Elektropribor, 1996. 226 p.

[3] Kuznetsov A.G., Molchanov A.V., Chirkin M.V., Izmaylov E.A. Precision laser gyroscope for self-contained inertial navigation // Kvantovaya Elektronika. 2015 Vol. 45. No. 1. P. 78–88

[4] Timoshenko S.P. Oscillations in engineering / Timoshenko S.P.; reworked by D.K. Yangomper; translated from English by Y.G. Panovko from 3rd US edition. – Ed. 2, reprinted. Moscow: URSS, 2006. 439 p.

Construction and Verification Analysis of Digital Model for Dynamic Characteristics of Inertial Measurement Unit System

W. F Luo, J. Xu, W. M Geng, J. M Tang, T. J Wang
Xi'an Aerospace Precision Electromechanical Institute, Xi'an 710000,China
641654582@qq.com
Tel: +86(0)029-85618943

During the development and manufacturing process of inertial measurement unit (IMU) products, dynamic frequency characteristics serve as critical factors in evaluating product performance. The traditional approach involves initially conducting angular vibration and sweep frequency vibration tests on physical prototypes prior to filter design, the filter is then designed based on the test results and the system's dynamic characteristics requirements. Subsequently, additional angular vibration and sweep frequency vibration tests are performed on the physical prototypes to validate the filter's performance, based on these validation results, the filter is optimized and refined. Finally, a comprehensive final verification test is conducted to ensure that the filter fully meets the system requirements.

To address the issues of multiple iterations, extended test periods, and low efficiency in this process, this paper proposes a method for constructing and validating a digital model of the inertial measurement channel. This approach aims to conduct theoretical dynamic characteristic analysis through early-stage modeling and simulation of the inertial measurement channel, followed by physical verification analysis. Consequently, this method optimizes the dynamic characteristic testing process, enhances test efficiency, and reduces the overall cycle time.

The specific approach is as follows: First, identify the component units that influence the dynamic characteristics of the angular velocity and acceleration measurement channels in the IMU product. Second, perform detailed modeling for each component unit of both channels to derive their transfer functions and determine the frequency characteristics of each channel. Third, design filters for the angular velocity and acceleration measurement channels based on system requirements, and conduct system-level filtering simulations to ensure that the simulation results meet the test specifications. Finally, perform angular vibration and sweep frequency tests using physical prototypes, and compare and analyze the discrepancies between the theoretical modeling simulation results and the actual test results.

The test results demonstrate that the theoretical analysis is highly consistent with the actual test results. Compared to traditional methods, this approach significantly shortens the test process, reduces test costs, and enhances test efficiency while ensuring the accuracy of the test results. This method provides important guidance for both the theoretical analysis and engineering practice of the dynamic characteristics of IMU systems.

Key Words: Inertial Measurement Unit(IMU),Dynamic Frequency Characteristics, Digital Model Construction, Verification and Analysis

I. INTRODUCTION

The inertial measurement unit (IMU) is a critical unit in aerospace weapon systems, tasked with the essential mission of real-time sensing and outputting the state information of angular and linear motions of the carrier in three directions[1]. The dynamic frequency characteristics of the IMU refer to the variation in output response with respect to input frequency under external linear or angular motion excitation. As weapon systems increasingly demand high-dynamic and large-maneuverability performance, the frequency characteristics of IMU products have become a key indicators for evaluating their performance[2].

Currently, the evaluation methods for the frequency characteristics of the IMU primarily rely on repetitive experimental methods, including angular vibration test[3] and linear vibration test[4]. Specifically, the process involves conducting multiple unfiltered physical tests, and subsequently designing digital filters based on analyses. Following this, filtered tests are conducted to optimize and verify the filters through repeated iterations between theoretical design and physical test. This process is time-consuming, costly, and challenging to meet the current demands for low-cost, tight-schedule, and short-cycle.

In light of this, the paper proposes a method for constructing and validating a digital model of the inertial measurement channel. By analyzing the dynamic characteristics of channels through modeling and simulation, extensive physical testing can be reduced while maintaining accuracy. This significantly decreases repetitive tests, optimizes the evaluation process, reduces costs, and improves efficiency. This method provides robust theoretical support and valuable engineering references for future research into the frequency characteristics of IMU products.

II. IDENTIFICATION AND MODELING OF CRITICAL COMPONENTS OF THE FREQUENCY CHARACTERISTICS

A. . Identification and Modeling of Frequency Characteristics for Angular Velocity Channel

Taking a specific fiber optic IMU as an example, this study identifies the key links that influence the frequency characteristics of the angular velocity channel. The schematic diagram illustrating the signal transmission is presented in Fig. 1.

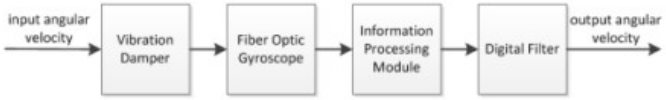


Fig. 1. Schematic diagram of signal transmission in angular velocity measurement channel.

From Fig. 1, the key components affecting the frequency characteristics of the angular velocity channel are the damper, fiber optic gyroscope, information processing module, and digital filters. These components are modeled and simulated in the following sections.

a) Development of the Damper Model in Angular Velocity Channel

Based on classical vibration analysis theory, angular vibration systems[5] can be modeled in Fig. 2.

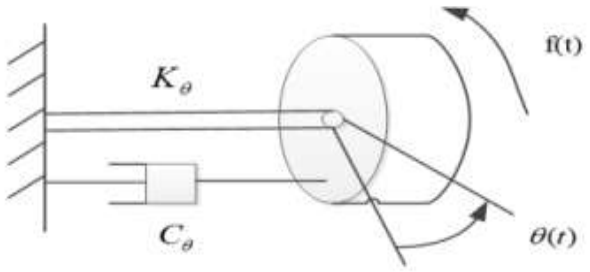


Fig. 2. Angular vibration damping system model.

The equation of motion is given by：

$$I\ddot{\theta}+C_\theta(\dot{\theta}-\dot{f})+k_\theta(\theta-f)=0 \quad (1)$$

Applying Laplace transformation to the above equation of motion yields the damper's transfer function as：

$$H_{j1}(s)=\frac{\theta(s)}{F(s)}=\frac{c_\theta s+k_\theta}{Is^2+c_\theta s+k_\theta} \quad (2)$$

Upon normalization，

$$H_{j1}(s)=\frac{\theta(s)}{F(s)}=\frac{2\cdot\xi_r\cdot\omega_r\cdot s+\omega_r^{\,2}}{s^2+2\cdot\xi_r\cdot\omega_r\cdot s+\omega_r^{\,2}} \quad (3)$$

where $\xi_r=\frac{c_\theta}{2\sqrt{Ik_\theta}}$ is the damping ratio； $\omega_r=\sqrt{k_\theta/I}$ is the natural frequency.

The natural frequency under angular vibration is 190 Hz, with a damping ratio of 0.15. Substituting these characteristic parameters into the transfer function results in:

$$H_{j1}(s)=\frac{3.5814e2\cdot s+1.4252e6}{s^2+3.5814e2\cdot s+1.4252e6} \quad (4)$$

The S-domain transfer function of the damper is converted to the Z-domain using the bilinear transformation[6] with a sampling frequency of 2000Hz, resulting in $H_{j1}(z)$.

b) Development of the Fiber Optic Gyroscope(FOG) Model

The FOG, an inertial unit for sensing and measuring carrier angular motion, uses the Sagnac effect[7] to detect the phase between two light waves proportional to the carrier's angular rate. Based on the optical wave interference principle, the signal is transmitted via the optical path, modulated by the circuit, and outputs angular velocity information. The FOG consists of optical and electrical components, with its functional schematic diagram shown in Fig. 3.

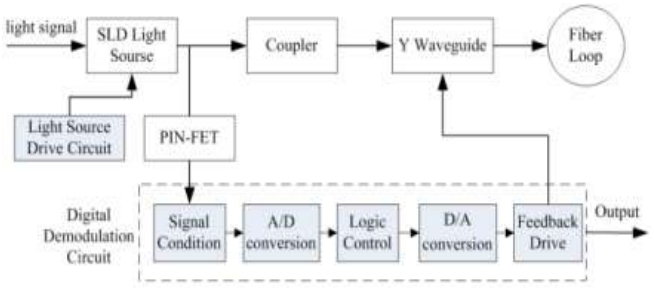


Fig. 3. Schematic diagram of functional principles in FOG.

The functional schematic diagram of the FOG can be modeled as a closed-loop control system, expressed as follows:

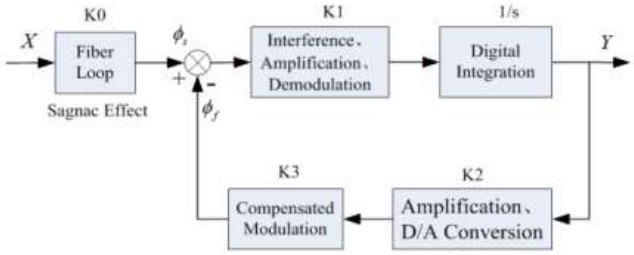


Fig. 4. Closed-loop control system model for FOG.

In this figure, $\phi_s=K_0\cdot x=\frac{2\cdot\pi\cdot L\cdot D}{\lambda\cdot c}\cdot x$ phase difference of the Sagnac effect is proportional to the input angular rate. After simplification, the interference, amplification, filtering, and demodulation stages are equivalent to a proportional link K1. Similarly, the D/A conversion and post-amplification stages are equivalent to a proportional link K2. The feedback compensation's amplitude adjustment stage simplifies to K3.

By simplifying and analyzing the closed-loop model, the system's transfer function is derived as follows:

$$H_{j2}(s)=\frac{Y(s)}{X(s)}=\frac{1}{M}\cdot\frac{1}{\tau s+1} \quad (5)$$

where， $M=\frac{K_2\cdot K_3}{K_0}$ ， $\tau=\frac{1}{K_1\cdot K_2\cdot K_3}$.

This FOG has a ring length of 2000m, a diameter of 110mm, a wavelength of 1550nm, a light source power of 3mW, a photoelectric conversion efficiency of 0.02V/uW, and a signal conditioning amplification factor of 8. Substituting these parameters into the transfer function gives:

$$H_{j2}(s)=\frac{1}{7.0082e\text{-}4\cdot s+1} \quad (6)$$

The transfer function of the FOG is transformed from the S-domain to the Z-domain, resulting in $H_{j2}(z)$.

c) Development of the Information Processing Module Model in Angular Velocity Channel

The information processing module collects and reprocesses the angular velocity data from Section b), including temperature error compensation, data accumulation, and smoothing. Its model acts as a phase delay link with the transfer function:

$$H_{j3}(s)=e^{-s\tau} \quad (7)$$

The first-order Pade approximation[8] is applied to the transfer function, effectively simulating phase delay in medium and low-frequency regions. The approximated transfer function follows:

$$H_{j3}(s)=\frac{-\tau s+2}{\tau s+2} \qquad (8)$$

The actual measurement of the information processing delay time is 0.3ms. Substituting the characteristic parameters into the transfer function gives:

$$H_{j3}(s)=\frac{-3e-4s+2}{3e-4s+2} \qquad (9)$$

The transfer function is transformed from the S-domain to the Z-domain, resulting in $H_{j3}(z)$.

d) Development of the Digital Filter Model in Angular Velocity Channel

Based on the frequency characteristics in Sections a) to c) and aligned with the channel dynamic requirements, the digital filter was designed. The requirements of IMU are as follows: amplitude-frequency (within 90Hz) ensure that the frequency is no less than 50 Hz at -1 dB attenuation and no less than 70 Hz at -3 dB attenuation, with a resonant peak not exceeding 2 dB. Phase-frequency (within 90Hz)ensure phase delay of no more than 4.5ms.

The digital filter, designed in the Z-domain, includes low-pass and notch sections. Its transfer function in the Z-domain is $H_{j4}(z)$,the frequency characteristics shown in Fig. 5.

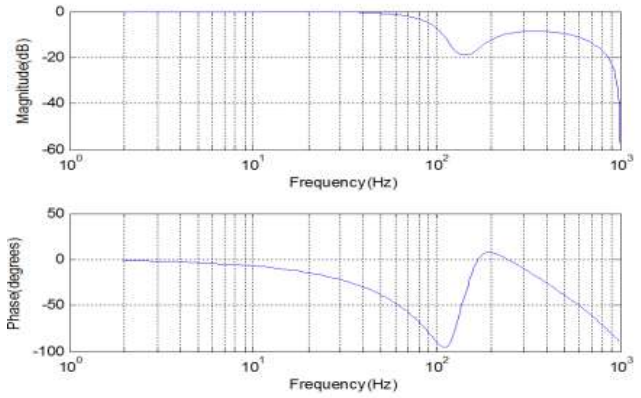


Fig. 5. Frequency characteristics of the digital filter in angular velocity measurement channel.

Comprehensive Model in Angular Velocity Channel

The transfer functions of each link in the angular velocity channel from Section a) to d) are cascaded to obtain the overall transfer function.

$$H_j(z)=H_{j1}(z)\cdot H_{j2}(z)\cdot H_{j3}(z)\cdot H_{j4}(z) \qquad (10)$$

The amplitude-frequency and phase-frequency characteristic shown in Fig. 6.

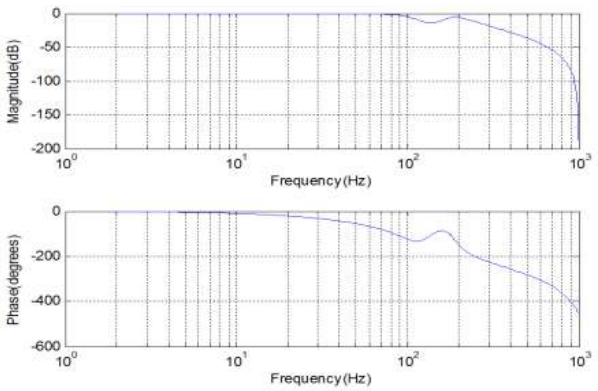


Fig. 6. F Frequency characteristics in angular velocity measurement channel.

The theoretical amplitude-frequency and phase-frequency characteristics are derived from Fig. 6, with specific data presented in Table I.

TABLE I. THEORETICAL RESULTS OF ANGULAR VELOCITY CHANNEL FREQUENCY CHARACTERISTICS

Demand of frequency characteristics		Theoretical results	Y/N
freq. at -1dB (Hz)	≥50	64.5803	Y
freq. at -3dB (Hz)	≥70	88.6135	Y
resonant peak (dB)	≤2	-0.0003	Y
system delay (ms)	≤4.5	≤3.5634	Y

As shown in Table I, the theoretical modeling and simulation of the angular velocity channel confirm that the -1dB and -3dB frequencies, as well as the resonant peak, meet system requirements. The phase-frequency delay time is also within the specified range. The dynamic frequency characteristics will be further validated through angular vibration test.

B. Identification and Modeling of Frequency Characteristics for Acceleration Channel

Taking the specific fiber optic IMU as an example, this study identifies the key links that influence the frequency characteristics of the acceleration channel. The schematic diagram illustrating the signal transmission is presented in Fig. 7.

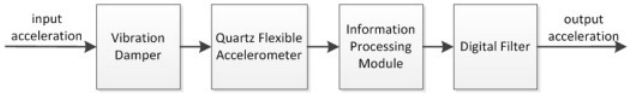


Fig. 7. Schematic diagram of signal transmission in acceleration measurement channel.

From Fig. 7, the key components affecting the frequency characteristics of the acceleration channel are the damper, quartz flexible accelerometer, information processing module, and digital filters. These components are modeled and simulated in the following sections.

a) Development of the Damper Model in Acceleration Channel

Based on classical vibration analysis theory, acceleration vibration systems[9] can be modeled in Fig. 8.

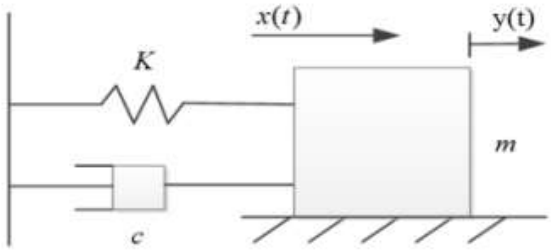


Fig. 8. Linear vibration damping system model.

The equation of motion is given by:

$$m\ddot{y}+c(\dot{y}-\dot{x})+k(y-x)=0 \quad (11)$$

Applying Laplace transformation to the above equation of motion yields the damper's transfer function as:

$$H_{x1}(s)=\frac{Y(s)}{X(s)}=\frac{cs+k}{ms^2+cs+k} \quad (12)$$

Upon normalization,

$$H_{x1}(s)=\frac{Y(s)}{X(s)}=\frac{2\cdot\xi_n\cdot\omega_n\cdot s+\omega_n^2}{s^2+2\cdot\xi_n\cdot\omega_n\cdot s+\omega_n^2} \quad (13)$$

where $\xi_n=\frac{c}{2\sqrt{mk}}$ is the damping ratio; $\omega_n=\sqrt{k/m}$ is the natural frequency.

The natural frequency under acceleration vibration is 110 Hz, with a damping ratio of 0.25. Substituting these characteristic parameters into the transfer function results in:

$$H_{x1}(s)=\frac{3.4558e2\cdot s+4.7769e5}{s^2+3.4558e2\cdot s+4.7769e5} \quad (14)$$

The S-domain transfer function of the damper is converted to the Z-domain using the bilinear transformation with a sampling frequency of 2000Hz, resulting in $H_{x1}(z)$.

b) Development of the Quartz Flexible Accelerometer Model

The quartz accelerometer measures the external specific force using Newton's second law and the piezoelectric effect[10]. After processing by its internal sensor and circuit, it outputs a pulse equivalent to the specific force. The accelerometer consists of two main parts: the meter assembly and the servo circuit, as shown in the functional schematic diagram in Fig. 9.

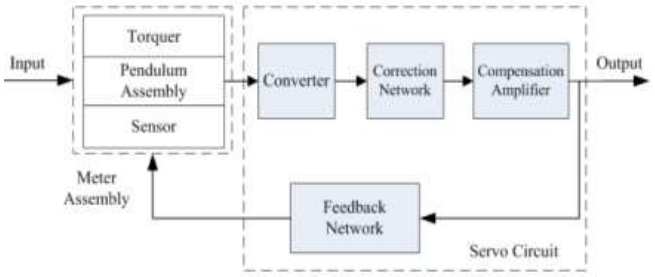


Fig. 9. Schematic diagram of functional principles in quartz flexible accelerometer.

The functional schematic diagram of the accelerometer is transformed into an equivalent closed-loop control model, represented as follows:

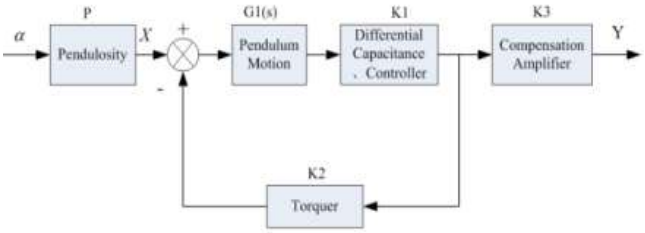


Fig. 10. Closed-loop control system model for quartz flexible accelerometer.

In the figure, $X=\alpha\cdot P$ is the inertial torque acting on the pendulum assembly, and $G_1(s)=\frac{1}{Js^2+Cs+K_c}$ is the transfer function of the accelerometer's pendulum motion, where J is the moment of inertia, C is the damping coefficient, and Kc is the stiffness of the flexible beam. The differential capacitance sensor and controller simplify to a proportional gain K1, the torque actuator simplifies to K2, and the amplitude compensation amplification section simplifies to K3.

By simplifying and analyzing the closed-loop model, the system's transfer function is derived as follows:

$$H_{x2}(s)=\frac{Y(s)}{X(s)}=\frac{K_1\cdot K_3}{Js^2+Cs+K_c+K_1\cdot K_2} \quad (15)$$

The quartz accelerometer's moment of inertia is 1.3E-7 kg•m², damping coefficient is 2.1E-4 N•m•s, flexible beam stiffness is 6.78E-4 N•m/rad, capacitance sensor coefficient is 2.45E3 pF/rad, controller coefficient is 25 mA/pF, torque actuator coefficient is 2.7E-5 N•m/mA, and compensator coefficient is 2.70E-5. Substituting these parameters into the transfer function gives:

$$H_{x2}(s)=\frac{1.2726e7}{s^2+1.6154e3\cdot s+1.2726e7} \quad (16)$$

The transfer function of the quartz accelerometer is transformed from the S-domain to the Z-domain, resulting in $H_{x2}(z)$.

c) Development of the Information Processing Module Model in Acceleration Channel

The information processing module collects and reprocesses the acceleration data from Section b), including temperature error compensation and data smoothing. Its model acts as a phase delay link with the transfer function:

$$H_{x3}(s)=e^{-s\tau} \quad (17)$$

The acceleration and angular velocity channel information is processed simultaneously with consistent phase delays. The transfer function of Z-domain is $H_{x4}(z)$.

d) Development of the Digital Filter Model in Acceleration Channel

Based on the frequency characteristics in Sections a) to c) and aligned with the channel dynamic requirements, the digital filter was designed. The requirements of IMU are as follows: amplitude-frequency (within 150Hz) ensure that the frequency is no less than 80 Hz at -1 dB attenuation and no less than 100 Hz at -3 dB attenuation, with a resonant peak not exceeding 4 dB. Phase-frequency (within 90Hz)ensure a phase delay of no more than 4.5ms.

The digital filter includes low-pass and notch sections. Its transfer function in the Z-domain is $H_{x4}(z)$,the frequency characteristics shown in Fig. 11.

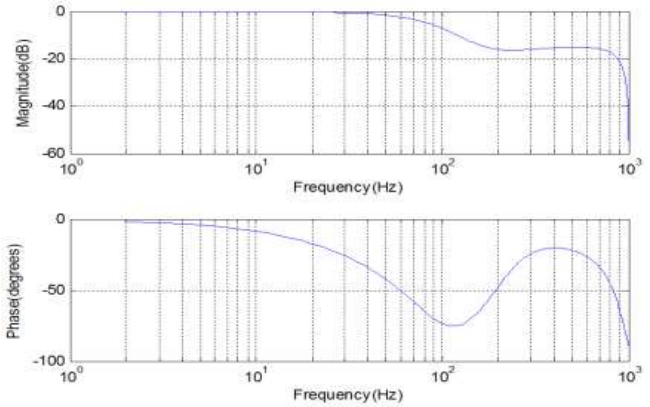


Fig. 11. Frequency characteristics of the digital filter in acceleration measurement channel.

e) Comprehensive Model in Acceleration Channel

The transfer functions of each link in the Acceleration channel from Section a) to d) are cascaded to obtain the overall transfer function.

$$H_x(z) = H_{x1}(z) \cdot H_{x2}(z) \cdot H_{x3}(z) \cdot H_{x4}(z) \quad (18)$$

The frequency characteristic shown in Fig. 12.

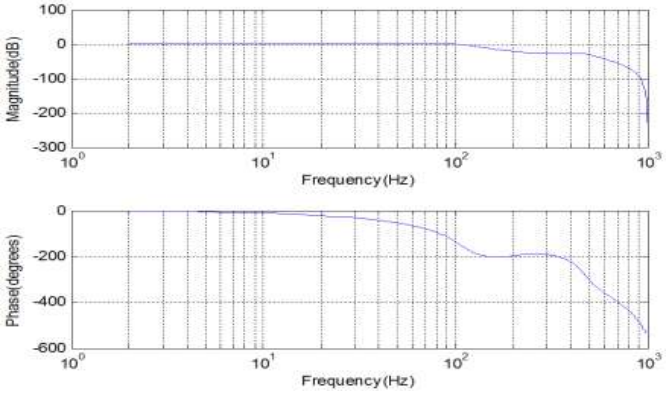


Fig. 12. Frequency characteristics in acceleration measurement channel.

The theoretical frequency characteristics are derived from Fig. 10, with specific data presented in Table II.

TABLE II. THEORETICAL RESULTS OF ACCELERATION CHANNEL FREQUENCY CHARACTERISTICS

Demand of frequency characteristics		Theoretical results	Y/N
freq. at -1dB (Hz)	≥80	107.5722	Y
freq. at -3dB (Hz)	≥100	116.1449	Y
resonant peak (dB)	≤4	0.8024	Y
system delay (ms)	≤4.5	≤4.0968	Y

As shown in Table II, the theoretical modeling and simulation of the acceleration channel confirm that the -1dB and -3dB frequencies, as well as the resonant peak, meet system requirements. The phase-frequency delay time is also within the specified range. The dynamic frequency characteristics will be further validated through a sweep frequency vibration test.

III. VERIFICATION OF FREQUENCY CHARACTERISTICS FOR IMU PRODUCTS

This chapter uses the 001# product of the specified optical fiber IMU to verify the system's dynamic frequency characteristics. The angular velocity channel is tested via angular vibration, and the acceleration channel is tested via sweep frequency vibration.

A. Angular Vibration Test

a) Procedure of Angular Vibration Test

The angular vibration test is performed on a dedicated platform. The control console outputs the required vibration frequency and controls the platform operation. The processing circuit collects, processes, and synchronizes signals from the platform and IMU. The test system connection is shown in Fig. 13, with single-point frequency input as specified in Table III.

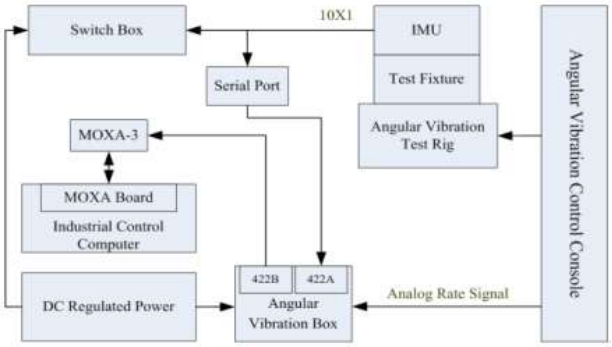


Fig. 13. Fig. 13. Schematic diagram of the angular vibration test.

TABLE III. ANGULAR VIBRATION TEST INPUT

Index	Freq.	Test time	Index	Freq.	Test time
1	0.5Hz	20s	11	70Hz	4s
2	1Hz	20s	12	80Hz	4s
3	5Hz	15s	13	90Hz	3s
4	10Hz	15s	14	100Hz	3s
5	15Hz	15s	15	110Hz	3s
6	20Hz	10s	16	120Hz	3s
7	30Hz	10s	17	130Hz	3s
8	40Hz	10s	18	140Hz	3s
9	50Hz	6s	19	150Hz	3s
10	60Hz	6s	/	/	/

b) Results of Angular Vibration Test

The results of angular vibration frequency characteristics are presented in Table IV.

TABLE IV. ANGULAR VIBRATION TEST RESULTS

Axis	Demand of dynamic characteristics		Practical results	Y/N
X	freq. at -1dB(Hz)	≥50	62.9756	Y
	freq. at -3dB(Hz)	≥70	87.4518	Y
	resonant peak(dB)	≤2	0.0379	Y
	system delay(ms)	≤4.5	≤3.5598	Y
Y	freq. at -1dB(Hz)	≥50	63.4309	Y
	freq. at -3dB(Hz)	≥70	86.0576	Y
	resonant peak(dB)	≤2	0.0619	Y
	system delay(ms)	≤4.5	≤3.5339	Y
Z	freq. at -1dB(Hz)	≥50	64.5923	Y
	freq. at -3dB(Hz)	≥70	87.5391	Y
	resonant peak(dB)	≤2	0.0116	Y
	system delay(ms)	≤4.5	≤3.5283	Y

Table IV shows that in the angular vibration test, the amplitude-frequency and phase-frequency results of the 001# product across the X, Y, and Z axes meet the angular velocity channel's requirements. Specifically, the minimum frequency is 62.9756Hz at -1dB (X-axis), 86.0576Hz at -3dB (Y-axis), with a maximum resonance peak of 0.0619dB (Y-axis) and a maximum system delay of 3.5598ms (X-axis).

B. Sweep Frequency Vibration Test

a) Procedure of Sweep Frequency Vibration Test

The sweep frequency vibration test is conducted on the test platform, the control console outputs the required sweep-frequency and magnitude, controls the platform operation. The system connection is shown in Fig. 14 with frequency input conditions in Table V.

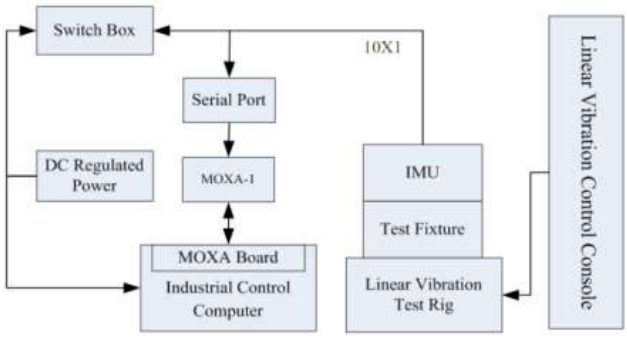


Fig. 14. Schematic diagram of the sweep vibration test.

TABLE V. TABLE V. WEEP VIBRATION TEST INPUT

Index	Content	Condition
1	Sweep range	20Hz~200Hz
2	Sweep mode	Logarithmic
3	Test time	2min
4	Vibration magnitude	2g

b) Results of Sweep Frequency Vibration Test

The results of sweep frequency vibration characteristics are presented in Table VI.

TABLE VI. SWEEP VIBRATION TEST RESULTS

Axis	Demand of dynamic characteristics		Practical results	Y/N
X	freq. at -1dB(Hz)	≥80	107.2958	Y
	freq. at -3dB(Hz)	≥100	116.7245	Y
	resonant peak(dB)	≤4	0.8261	Y
	system delay(ms)	≤4.5	≤4.0489	Y
Y	freq. at -1dB(Hz)	≥80	108.5253	Y
	freq. at -3dB(Hz)	≥100	117.4210	Y
	resonant peak(dB)	≤4	0.8400	Y
	system delay(ms)	≤4.5	≤4.0469	Y
Z	freq. at -1dB(Hz)	≥80	109.6415	Y
	freq. at -3dB(Hz)	≥100	117.7712	Y
	resonant peak(dB)	≤4	1.0521	Y
	system delay(ms)	≤4.5	≤4.0497	Y

Table VI shows that in the sweep vibration test, the amplitude-frequency and phase-frequency results of the 001# product across the X, Y, and Z axes meet the acceleration channel's requirements. Specifically, the minimum frequencies are 107.2958Hz at -1dB and 116.7245Hz at -3dB (both on the X-axis), the maximum resonance peak is 1.0521dB (Z-axis), and the maximum system delay is 4.0497ms (Z-axis).

IV. COMPARATIVE ANALYSIS OF THEORETICAL AND EXPERIMENTAL RESULTS FOR FREQUENCY CHARACTERISTICS

Chapter 2 identifies the key factors affecting the frequency characteristics of the angular velocity and acceleration channels, models these factors, and establishes the channel transfer model. Theoretical simulation analysis predicts that the product meets amplitude-frequency and phase-frequency design requirements. Chapter 3 verifies these characteristics through experiments, confirming that frequency characteristics meet design specifications. This chapter compares and analyzes the theoretical simulation and experimental verification results.

A. Comparative Analysis of Frequency Characteristics for Angular Velocity Channel

The theoretical simulation results of the angular velocity channel were compared with the actual angular vibration test results, the comparison shown in Fig. 15. The figure shows that the amplitude-frequency and phase-frequency curves in the X, Y, and Z axes are highly consistent with the theoretical analysis.

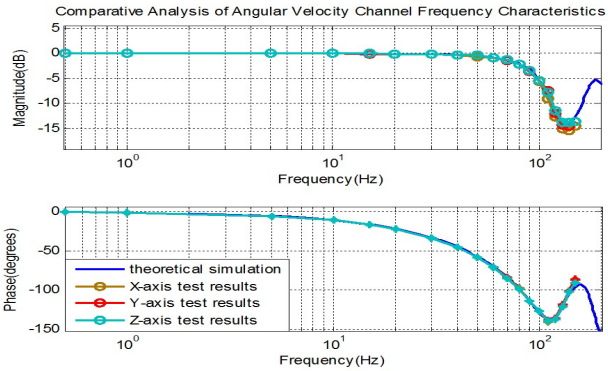


Fig. 15. Comparative analysis of theoretical simulation and test result of angular velocity channel frequency characteristics.

The difference between the theoretical analysis and the angular vibration test results of the three axes was calculated to show the variation in amplitude and phase differences across frequencies. Details are in Fig. 16 and 17.

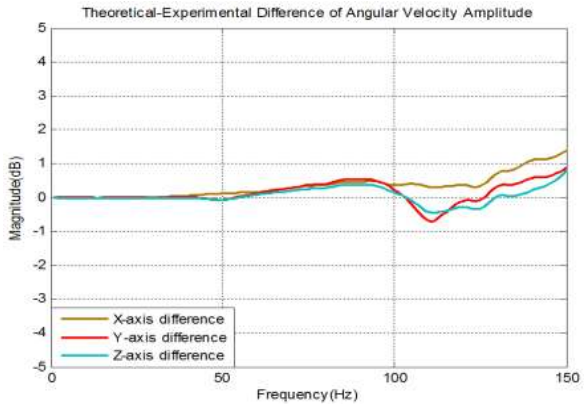


Fig. 16. Difference of theoretical simulation and test result of angular velocity amplitude.

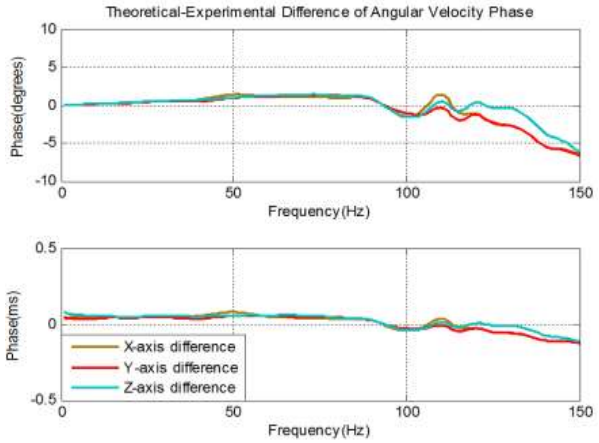


Fig. 17. Difference of theoretical simulation and test result of angular velocity phase.

From Fig. 15 and 16, at 50Hz, the differences between the theoretical amplitude-frequency curve and the X, Y, and Z-axis test curves are 0.12dB, 0.08dB, and 0.10dB, respectively. At 70Hz, the differences are 0.27dB, 0.28dB, and 0.19dB. For frequencies above 100Hz, the differences increase slightly: 1.36dB for the X-axis, 0.88dB for the Y-axis, and 0.79dB for the Z-axis.

From Fig. 15 and 17, the maximum angle differences between the theoretical phase-frequency curve and the X, Y, and Z-axis test curves are 6.51°, 6.76°, and 6.23°, respectively. The corresponding maximum time differences are 0.12ms, 0.13ms, and 0.11ms.

B. Comparative Analysis of Frequency Characteristics for Acceleration Channel

The frequency characteristics of the acceleration channel were compared between the theoretical simulation and the actual sweep vibration test results. As shown in Fig. 18, the amplitude-frequency and phase-frequency curves for the X, Y, and Z axes closely match the theoretical analysis.

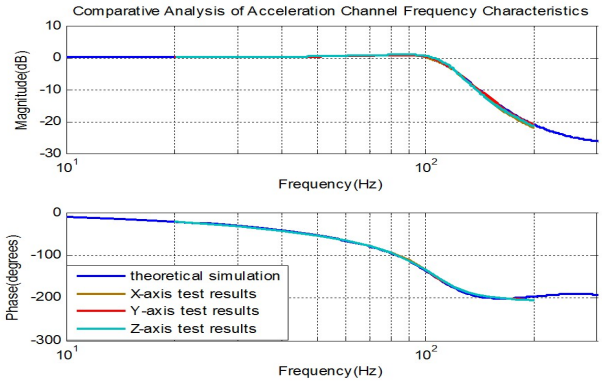


Fig. 18. Comparative analysis of theoretical simulation and test result of acceleration channel frequency characteristics.

The differences in amplitude and phase between the theoretical analysis and sweep frequency test results for the three axes were calculated at various frequencies. Details are in Fig. 19 and 20.

From Fig. 18 and 19, at 80Hz, the discrepancies between the theoretical amplitude-frequency curve and the test curves are 0.06dB (X-axis), 0.08dB (Y-axis), and 0.13dB (Z-axis). At 100Hz, these discrepancies are 0.02dB (X-axis), 0.21dB (Y-axis), and 0.39dB (Z-axis). For frequencies above 140Hz, the discrepancies increase slightly, with maximum values of 1.09dB (X-axis), 0.36dB (Y-axis), and 0.87dB (Z-axis).

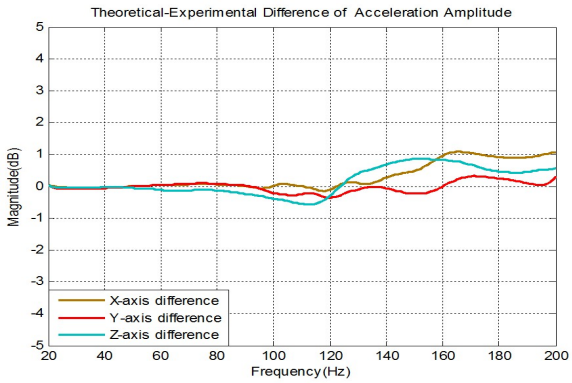


Fig. 19. Difference of theoretical simulation and test result of acceleration amplitude.

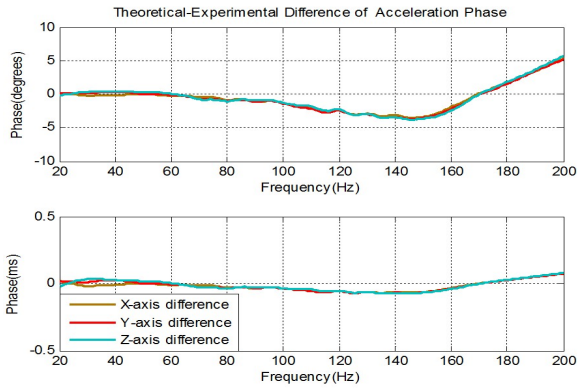


Fig. 20. Difference of theoretical simulation and test result of acceleration phase.

From Fig. 18 and 20, the maximum angular discrepancies between the theoretical phase-frequency curve and the test curves are 5.35° (X-axis), 5.18° (Y-axis), and 5.69° (Z-axis). The corresponding maximum time discrepancies are 0.074ms (X-axis), 0.072ms (Y-axis), and 0.079ms (Z-axis).

The results of Sections A and B show that the dynamic frequency characteristics of the angular velocity and acceleration channels exhibit good consistency between theoretical simulation and practical experiments. This confirms the feasibility of replacing early-stage exploratory tests with theoretical simulation analysis.

V. CONCLUSION

The study addresses the issues of multiple iterations, long test cycles, and low efficiency in traditional methods for testing the dynamic frequency characteristics of IMU. A new analytical method is proposed based on constructing and optimizing digital models of inertial measurement channels. By identifying key links in these channels and establishing precise mathematical models, the frequency characteristics of angular velocity and acceleration measurement channels are analyzed, yielding theoretical results. Based on this, physical tests validate the system's frequency characteristics, with all results meeting design requirements. Comparison of theoretical and experimental data shows that in the angular velocity channel, the amplitude-frequency error is less than 1.36 dB and the phase-frequency error is less than 0.13ms. In the acceleration channel, the amplitude-frequency error is less

than 1.09 dB and the phase-frequency error is less than 0.079ms.

These findings confirm the strong consistency between theoretical analysis and experimental results, validating the accuracy and reliability of the proposed method. The research provides an efficient solution for verifying IMU dynamic frequency characteristics and serves as a valuable theoretical and technical reference for future studies and applications in related fields.

REFERENCES

[1] Y. Y Qin. Inertial Navigation (second edition) [M], Science Press,2014:243.

[2] X. H Liu, C. C Hao, Y. K Wang. Study on the Angular Vibration Frequency Characteristics of Inertial Measurement Devices Based on Wavelet Transform [J]. Information Technology and Informatization,2024,(07):116-121.

[3] J. K Guan. Research on Vibration Control of Single-Axis Electrohydraulic Angular Vibration Table[D]. Harbin Institute of Technology,2009.

[4] M. P ,F. C ,G. Z et al. Spectral analysis of sine-sweep vibration: A fatigue damage estimationmethod[J].Mechanical Systems and Signal Processing,2021,157.

[5] W. Y Wang, P. Qi, H. Cheng et al. Theory Analysis on Resonance Frequencies of Linear and Torsional of Strapdown IMU System[J].Aviation Precision Manufacturing Technology,2012,58(06): 33-35+18.

[6] J. L Liu, K. Z Fu, Z. Q Mai et al. Sensorless Control Strategy of Improved HF Pulsating Voltage Injection Based on Dual Frequency Notch Filter[J]. Proceedings of the CHEE,2021,41(02): 749-759.

[7] G Sagnac.L'ether lumineux demontre par l'effet du vent relative d'ether dans an interferometre[J]. Comptes Rendus de l'Academie des Sciences,1913,157(125): 708-710.

[8] Z. X Zhang, D. W Zhang. Theory analysis on resonance frequencies of linear vibration and torsional vibration of strapdown IMU damping system[J]. Journal of Chinese Inertial Technology,2009,17(06):654-657.

[9] Y. Zhang, T. Z Fang, H. Liu et al. Robust Control Method of Grid-connected Inverter Based on μ -Synthesis Theory and Genetic Algorithm[J]. Proceedings of the CSEE,2024,44(12):4891-4904.

[10] R. L Lin. Study on Quartz Vibrating Beam Accelerometer Technology[D]. University of electronic science and technology, 2006.

Infrared Sensor-Based Attitude Determination System for CubeSat Satellites

D.V.Akulin
Special Technology Center, LLC
Saint Petersburg, Russia
alkuin.work@gmail.com

M. V. Mentyukov
Special Technology Center, LLC
Saint Petersburg, Russia
maksim.mentukov.stc@yandex.ru

K.O.Ivanov
«Saint Petersburg State University of Aerospace Instrumentation»
Saint Petersburg, Russia

Abstract – The article describes an attitude determination system for a small spacecraft, based on infrared focal plane array. It outlines the fundamental operating principles of the IR sensor, as well as the algorithms for determining the nadir and Sun direction vectors. The results of the attitude system's performance during flight development tests are also presented.

Keywords – nanosatellite, CubeSat, Earth infrared radiation, infrared focal plane array, attitude determination system

I. Introduction

The Attitude Determination System (ADS) is a critically important subsystem of a spacecraft (SC), ensuring the proper operation of mission payloads. Currently, sensors based on various physical principles are used for attitude determination: star trackers, Sun sensors, and magnetometers. However, each of these sensor types has significant limitations.

Star trackers provide high-precision attitude determination but are characterized by large mass and dimensions, as well as high cost [1]. Sun sensors only function when there is a direct line of sight to the Sun, making them unusable in shadowed regions of the orbit [2].Magnetometers are subject to the influence of electromagnetic interference created by onboard systems of spacecraft, which reduces the accuracy of measurements [3].

Infrared sensors are a promising alternative to the above sensors. Their advantages include:

- Compact size and low weight;
- Immunity to electromagnetic interference;
- Capability to operate in all orbital regions due to constant IR radiation from Earth.

Additionally, an infrared focal plane array (IRFPA) can detect the Sun's direction, which appears as a point source spanning a few pixels. Thus, an IRFPA can be used to develop a three-axis spacecraft attitude determination system.

This study is based on an analysis of modern methods for attitude determination using Earth and Sun infrared radiation. In [4], the authors describe the principle of Earth's infrared horizon detection. Research in [5-7] discusses the working principle and mathematical model of an IR sensor. The authors examine various sensor configurations capable of determining the nadir vector. Publications [8-9] details the use of low-resolution IRFPA for Earth horizon detection, covering image processing techniques such as subpixel interpolation, lens distortion compensation, and temperature correction. In [10], the authors propose using IRFPA to determine the Sun-direction vector. By combining these approaches, a three-axis attitude determination system for small spacecraft can be developed.

II. Mathematical Model of Attitude Determination System

A. Configuration of ADS

The attitude determination system consists of six IRFPAs located on six faces of the satellite. In Fig. 1, the cones schematically show the fields of view of each IRFPA. Each matrix has a size of 32x24 pixels and a field of view of 110x75 degrees. All sensors are oriented orthogonally to each other. This configuration results in a maximum coverage of the field of view of about 95% (Fig. 2).

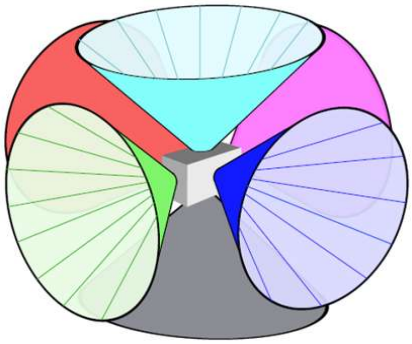


Fig. 1. Configuration of the IRFPAs location on the satellite

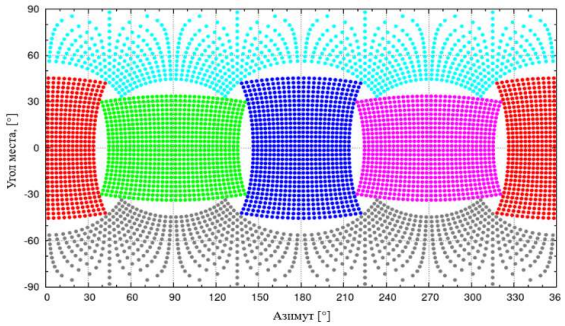


Fig. 2. Fields of view of IRFPA

Figure 3 displays the Earth and Sun imaging pattern across six IRFPAs acquired by the satellite at a 500 km altitude. In this orbit, the angular size of the Earth is approximately 130°, allowing its image to be captured simultaneously by multiple sensors. This improves the accuracy of nadir vector determination. In Fig. 3, the Earth is highlighted in green as displayed on five IRFPAs. The bright red spot on the sensor located on the +Z orientation face of the spacecraft is the Sun's image. The blurring of the spot is due to lens aberrations in the IRFPA and motion smear caused by the spacecraft's rotational movement

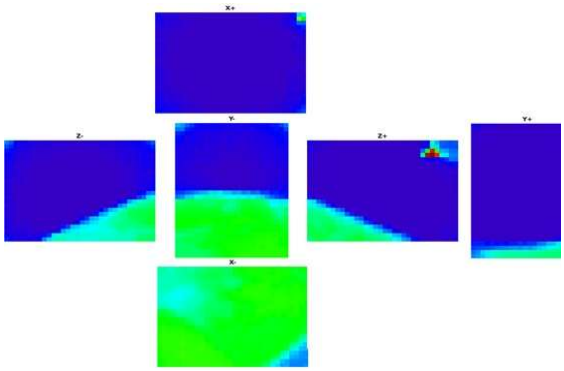


Fig. 3. Image of the Earth and the Sun on the IRFPA

B. Nadir vector determination algorithm

The algorithm consists of four sequential stages:

1. Image Filtering. Threshold filtering is applied using the minimum Earth temperature value in the infrared image. Pixels with values below the specified threshold are discarded. Custom binary masks for each IRFPA are employed to exclude areas obstructed by protruding spacecraft structural elements.

2. Image stitching. Sensor matrix points are mapped onto a unit sphere according to calibration data and the alignment of each sensor's boresight axis on the spacecraft (Figure 3). The resulting set of points on the sphere is designated as S.

3. Approximate Nadir Vector Calculation. The arithmetic mean position (centroid) is computed for all points in set S. The vector from the sphere's center to this calculated point serves as the initial nadir vector estimate (shown as the yellow point in Fig. 4).

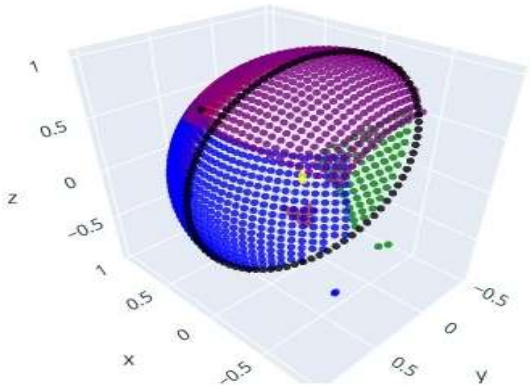


Fig. 4. Visualization of image stitching and approximate nadir vector computation steps

4. Iterative Nadir Vector Refinement. Using the current nadir vector estimate, the expected position of Earth's disk on the unit sphere is calculated (a circle with 130° angular diameter). A set of points S_e belonging to the calculated disk boundary is determined. Experimental data correlation is performed by computing distances between observed Earth edge points (set S) and modeled disk points (S_e). The sum of squared distances is used as the error metric. The least squares method is applied to minimize the error. Upon iteration completion, the final vector n is accepted as the refined nadir vector.

C. Algorithm for Determining the Sun Direction Vector

The algorithm for determining the Sun direction vector involves identifying a bright, isolated spot in the image captured by the IRFPA. Once the coordinates of the spot's center are determined, it is necessary to convert the pixel coordinates into the direction vector of the Sun in the IRFPA's reference frame. The matrix coordinate system is defined as follows: the OX and OY axes lie in the focal plane; the OZ axis is aligned along the optical axis of the IRFPA.To solve this problem, the collinearity equation (1) can be used.

$$\vec{s}_m = \begin{pmatrix} \sin\alpha\cos\beta \\ -\sin\alpha \\ \cos\alpha\cos\beta \end{pmatrix} = \lambda \begin{pmatrix} x_p - x_0 \\ y_p - y_0 \\ f \end{pmatrix}, \quad (1)$$

where α, β are positive rotation angles around the OX and OY axes;

x_p, y_p are coordinates of the Sun spot center in pixels;
x_0, y_0 are coordinates of the image center in pixels;
f is the focal length of the IRFPA lens;
λ is the proportionality coefficient.

The Sun direction vector in the IRFPA reference frame must then be transformed into the spacecraft coordinate system (2). The transformation matrix from the IRM coordinate system to the SC coordinate system is defined separately for each IRFPA.

$$\vec{s}_b = C_m^b \vec{s}_m, \quad (2)$$

D. Algorithm for determining spacecraft attitude

When only the nadir vector $\vec{n}_b$ is available (e.g., when the spacecraft is on the shadowed side of the orbit), it is possible to determine the local vertical angles: roll and pitch. The orbital coordinate frame in which the nadir vector is calculated is defined as follows. Axis *y* is directed along the radius vector; axis *x* lies in the orbit plane in the direction of motion; and axis *z* complements the system making it right-handed; the datum point is located in the spacecraft's center of the mass. The nadir vector in the given frame has the coordinates (0, -1, 0). Considering the direction cosine matrix [11 (page 339, formula 4.4.1)], it is possible to obtain the roll (3) and pitch (4) angles:

$$\gamma = arctg\left(\frac{-n_{bx}}{\sqrt{n_{by}^2 + n_{bz}^2}}\right); \quad (3)$$

$$\theta = arctg\left(-\frac{n_{bz}}{n_{by}}\right), \quad (4)$$

where γ, θ – roll and pitch angles respectively;
n_{bx}, n_{by}, n_{bz} – nadir vector components in body frame.

In the presence of the Sun direction vector in the satellite coordinate system and the Sun direction vector in the orbital coordinate system, which is calculated from information about the current position of the satellite in orbit, the spacecraft's attitude is calculated by the TRIAD method. The description of the TRIAD method is presented in the article [12].

III. Accuracy Assessment of IRFPA Based Attitude Determination

Since the IRFPAs serve as the primary attitude sensors, the performance of the attitude determination system can be evaluated indirectly using the following criteria.

A. Accuracy of Inter-Vector Angle Determination

The accuracy of the attitude determination system can be assessed by comparing the angle between the Sun direction vector and the nadir vector in two reference frames: orbital and body. The angle calculated in the orbital frame serves as the ground truth, as it relies on astrodynamics models and the precise satellite's orbital position. The angle is calculated in body frame is measured by IRFPA. The difference between these angles represents attitude determination error. As shown in Fig. 5, the root-mean-square (RMS) error of this angle's difference reaches 4.5°.

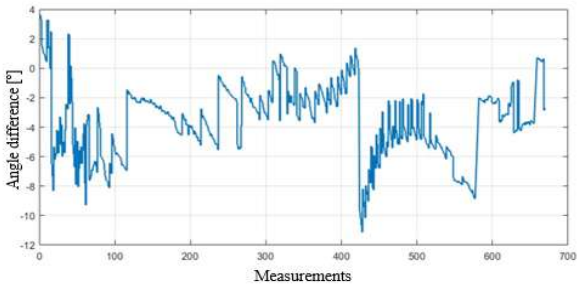


Fig. 5. Error in determining the angle between the nadir and sun direction vectors

B. Verification Using a Technological Camera

A series of experiments was conducted to evaluate attitude determination accuracy by stabilizing the optical axis of a technological camera on the Earth's horizon. The spacecraft was oriented such that the camera's optical axis pointed tangentially to the Earth's horizon (i.e., the horizon was centered in the frame). In the Fig. 6, the center of the frame is marked with a green dot and the Earth's horizon is marked with a red line. The angular deviation between the frame center and the horizon line was calculated using pixel angular resolution (known from camera specifications) and distance in pixels from the center to the horizon. The root-mean-square (RMS) pointing error was 2.5°.



Fig. 6. Image obtained from the process camera

IV. Simplified Two-Axis Attitude Determination Mode

The primary drawback of using IRFPA is the high data throughput (~10 kB) requiring real-time processing. To reduce computational load, a simplified algorithm with data binarization is proposed.

Algorithm Steps:

- Image Binarization. Each pixel is converted to 0 (cold space) or 1 (warm Earth) based on a threshold (minimum Earth temperature). Fig. 7b shows a binarized IRM image.
- Circular Mask Application. A circular mask (Fig. 7c) is applied, centered on the IRM's optical axis.The mask filters out noise and irrelevant regions.
- Nadir Vector Estimation. The ratio of "1" pixels inside the circular region correlates with the angle between the IRM's optical axis and the nadir direction. With at least three such angles, the nadir vector in the spacecraft frame can be computed. This method is analogous to the local vertical determination algorithm described in [13].

The attitude determination accuracy in the simplified mode is approximately 10°.

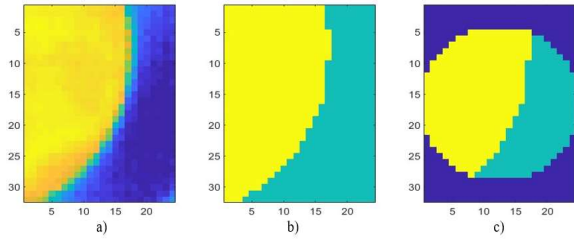


Fig. 7. a) image of the ground in the IR range. b) binarized image. c) circular mask.

V. Conclusions

This study has demonstrated the feasibility of using infrared matrices (IRMs) as a primary attitude determination source for spacecraft. The developed algorithms enable the computation of nadir and Sun-direction vectors, with attitude angles calculated using the TRIAD method. The evaluated attitude determination accuracy, assessed through indirect methods, is approximately 5°.

This research did not include descriptions of several advanced algorithms capable of improving attitude determination accuracy, specifically: precise characterization of IRM lens distortion coefficients during ground testing; Real-time updates to distortion parameters using in-flight data; subpixel interpolation techniques to achieve Sun-direction accuracy beyond single-pixel resolution. These methods were omitted because the spacecraft lacks an alternative attitude reference system required to validate IRM performance. Without independent high-accuracy sensors (e.g., star trackers), the benefits of these algorithms cannot be quantitatively assessed.

Upcoming missions will integrate high-precision star trackers (accuracy: arcminute-level), enabling: IRM output validation; on-orbit calibration of distortion models; performance benchmarking of sub pixel algorithms. Results from these enhanced methods will be published separately.

REFERENCES

[1] M. E. Prokhorov, A. I. Zakharov, A. V. Mironov, F. N. Nikolaev, andM. S. Tuchin, “Modern star trackers” Proc. 38th International Students’ Scientific Conference “Physics of Space”, Ural Federal University, 2009, pp. 170–186.

[2] Rufino, G., & Grassi, M. (2009). Multi-aperture CMOS sun sensor for microsatellite attitude determination. Sensors, 9(6), 4503-4524.K. Elissa, “Title of paper if known,” unpublished.

[3] Psiaki, M. L., Martel, F., & Pal, P. K. (1990). Three-axis attitude determination via Kalman filtering of magnetometer data. Journal of Guidance, Control, and Dynamics, 13(3), 506-514.

[4] Astheimer, R. W. (1966). Infrared Horizon Sensors for Attitude Determination. In Peaceful Uses of Automation in Outer Space: Proceedings of the First IFAC Symposium on Automatic Control in the Peaceful Uses of Space, held June 21–24, 1965, in Stavanger, Norway (pp. 381-397). Boston, MA: Springer US.

[5] Nguyen, T., Cahoy, K., & Marinan, A. (2018). Attitude determination for small satellites with infrared earth horizon sensors. Journal of Spacecraft and Rockets, 55(6), 1466-1475.

[6] Sáez, A. G., Quero, J. M., & Jerez, M. A. (2015). Earth sensor based on thermopile detectors for satellite attitude determination. IEEE Sensors Journal, 16(8), 2260-2271.

[7] Vertát, I., Linhart, R., Masopust, J., Vobornik, A., & Dudácek, L. (2017). Earth’s thermal radiation sensors for attitude determination systems of small satellites. Contrib. Astron. Obs. Skalnaté Pleso, 47, 157-164.

[8] Van Rensburg, H. M. (2008). An infrared earth horizon sensor for a leo satellite (Doctoral dissertation, Stellenbosch: University of Stellenbosch).

[9] Wessels, J. H. (2018). Infrared horizon sensor for cubesat implementation (Doctoral dissertation, Stellenbosch: Stellenbosch University).

[10] Kapás, K., Bozóki, T., Dálya, G., Takátsy, J., Mészáros, L., & Pál, A. (2021). Attitude determination for nano-satellites–I. Spherical projections for large field of view infrasensors. Experimental Astronomy, 51(2), 515-527.

[11] G. I. Emel’yantsevandA. P. Stepanov, Integrirovannyeinertsial’no-sputnikovyesistemyorientatsii i navigatsii (Integrated Itertial-Satellite Systems of Orietations and Navigation), St. Petersburg: Concern SRI Elektroprior, JSC, 2016.

[12] Markley, F. L. (1998). Attitude determination using two vector measurements.

[13] Akulin D.V., Mentyukov M.V., Goncharov R.B. Nadir Finding System with Infrared Temperature Sensors for CubeSat Satellites //XXXI Saint Petersburg International Conference on Integrated Navigation Systems. – 2024. – P. 75-81.

Autonomous Navigation Module for Aircraft

E.S. Minchenkova

LLC «Special Technological Center»,
St. Petersburg, Russia
eminch_work@mail.ru

D.D. Nikeev

LLC «Special Technological Center»,
St. Petersburg, Russia
dnikeev@stc-spb.ru

K.A. Karimulin

State University of Aerospace Instrumentation Saint-Petersburg,
St. Petersburg, Russia
kkarimulin285@gmail.com

P.V. Gaenko

LLC «Special Technological Center»,
St. Petersburg, Russia
pashapost82@mail.ru

***Abstract* — The autonomous aircraft navigation module based on visual navigation system and inertial navigation system is considered. The results of flight tests are given.**

Keywords—integrated navigation system, Kalman filtering, visual navigation system, autonomous navigation

I. INTRODUCTION

The development of algorithms for autonomous navigation of aircraft is an interesting task due to the limitations in obtaining signals from the Global Navigation Satellite System (GNSS) signals, which may be noisy or spoofed. One of the approaches to building autonomous systems is the integration of data from a strapdown inertial navigation system (SINS) and a visual navigation system (VNS).

The module under consideration uses micro-electromechanical (MEMS) inertial sensors of consumer class accuracy, which allows the cost of the navigation module to be significantly reduced. However, such sensors have significant noise characteristics and require continuous correction. Since the VNS operation on weather conditions, time of day and flight altitude; the navigation solution can be lost not only for short periods of time but also for long. In this regard, data fusion similar to classical SINS/GNSS schemes can lead to a rapid increase in errors at times periods when data from the VNS are not available. In the VNS under consideration, the position and attitude of the aircraft are determined by detecting keypoints and comparing them with reference points. A sequence of frames with good visibility is required to determine the ground velocity of the aircraft.

One of the main tasks of the autonomous navigation module is to minimise the error of position determination at the loss of information from the VNS. To improve the prediction accuracy, data from the air data computer can be used [1]. For example, paper [2] presents the results of the navigation complex, which implements obtaining a navigation solution based on data from SINS, GNSS and airspeed meter. The accuracy of position determination in the prediction mode is 7 % of the covered distance. In [3], a method of compensating SINS errors in case of satellite data loss using ground landmarks observations is proposed. It is shown that with a correction period of 300 seconds, the error of position determination does not exceed 200 metres.

This paper presents a data fusion algorithm based on wind speed and airspeed estimation using position and ground speed obtained from the VNS. The VNS uses the data from the SINS to improve the performance and the detection of keypoints in low visibility conditions.

II. MODULE COMPOSITION

The module includes:

- SINS with MEMS sensors (accelerometers, gyroscopes and magnetometers) of consumer accuracy class;
- VNS with visible fisheye camera;
- barometer;
- a computer where the information is integrated.

The attitude angles and vertical velocity are calculated in the attitude and heading reference system (AHRS) algorithm [4, 5]. Airspeed can be obtained using different protocols from the data packet from the on-board computer, which is linked to the Pitot tube. Geographic coordinates and ground speed are computed in the VNS. Based on the barometer measurements, altitude is calculated. Using all this data, wind speed can be calculated and position calculations can be performed in prediction mode. Attitude angles, estimated position and altitude are transmitted to the VNS to reduce image processing time and validate the calculated navigation parameters. In the absence of external airspeed data, the algorithm estimates the value of the longitudinal velocity component.

III. VNS ALGORITMS

The GNSS almanac data are processed by a convolutional neural network and stored as sets of keypoints and their descriptors. Each point has known coordinates, topography altitude and observation altitude. During flight, a region of interest is requested from this dataset, which depends on the current coordinates and the calculated error in their determination. To improve processing speed, the received image from the camera is aligned with the attitude angles from AHRS. The image is fed into a neural network, whose output determines a set of keypoints and their descriptors. In this research, a hybrid neural network structure has been chosen. The SuperPoint model [6] is used for point detection. The descriptors are computed based on the Xfeat model [7]. The Xfeat model allows the generation of shorter length vectors that describe a keypoint more uniquely. By comparing the descriptors of the found keypoints in the camera image with the point descriptors of the reference image, an initial comparison can be made. To filter out the good points, the RANSAC method is further

applied. A rigid-model [8] is used to fit parameters of the matching model. This model only needs two pairs of points from each matched image to determine the transformation parameters. On average, 200 iterations are required to determine the necessary parameters. For comparison, the homography model needs 4 pairs of points, which increases the search time and increases the probability of detection of spurious models. Additional validation of the obtained solution is obtained by comparing the calculated altitude and heading with the real one (which comes from the integration algorithm). To estimate the ground speed, it is necessary to accumulate several images with good visibility and analyse movement of keypoints. The Kannala camera model was used to calibrate the camera parameters [9].

IV. Wind Speed Estimation Algorithm

The algorithm is based on the extended Kalman filter. The state vector includes estimates of the wind speed projection on the geographic frame (ENH) $W=[w_e\ w_n\ w_h]^T$, the longitudinal component of the air velocity on the body-fixed frame V_a, and estimates of latitude φ and longitude λ:

The models of wind and airspeed components are assumed to be Wienerian.

The ground velocity vector in ENH $V=[v_e\ v_n\ v_h]^T$ is determined by the sum of vector W and projection V_a.The equations describing the evolution of the coordinate are defined as follows:

$$\dot{\varphi} = v_n/R_n, \dot{\lambda} = v_e/R_e/cos(\varphi). \quad (1)$$

Integration of the equations (1) is performed at each step of orientation obtaining. Due to the peculiarities of the visual system operation, correction can be performed separately for velocity data and coordinate data.

The measurements for the Kalman filter are the ground speed from the VNS in the body-fixed frame. VNS algorithms do not allow estimation of vertical velocity, so this parameter is included in the measurements from the VNS algorithm. In the case of position correction, the measurements are the latitude and longitude calculated by VNS.

The vertical channel operates independently and is calculated from barometer measurements converted to an ellipsoid. The initial altitude is determined from the Shuttle Radar Topography Mission database using the coordinates of the takeoff location. In addition, the barometric altitude is corrected if altitude measurements from the VNS are available.

The air velocity measurement from the Pitot tube directly corrects for the respective component in the state vector.

V. Flight Tests

To assess accuracy, the module was equipped with an additional GNSS sensor. During the flight, the integrated solution was tested and the VNS was switched off for several minutes. Fig. 1 shows the position determination error in one of the flights.

The flight altitude was 750 meters, and at the end of the flight a landing was performed. The flight was conducted in the evening. In the autonomous flight sections, the position determination error calculated from the estimated parameters after GNSS correction is additionally shown. This result was obtained as a result of post-processing of flight data

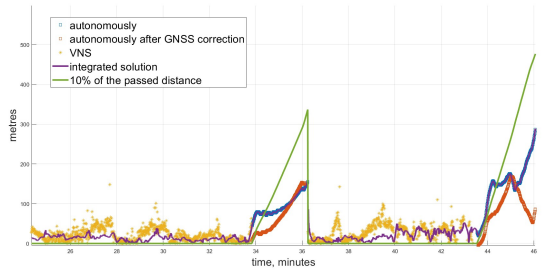


Fig. 1. Position determination error

Fig.2 shows the results from a daytime flight at altitudes of 750-900 meters, with a landing at the end.

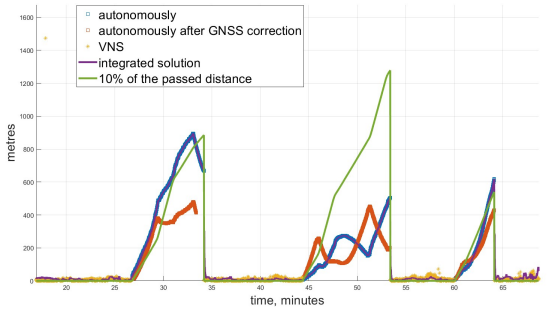


Fig. 2. Coordinate determination error

The error in position estimation in the prediction mode is 10-15 % of the covered distance.

Fig. 3 shows an example of testing the algorithm in flight, which was conducted in cloudy weather. The altitude was changed from 1800 to 1100 meters.

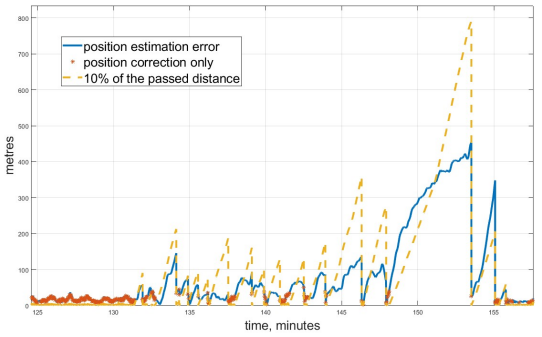


Fig. 3. Position determination error in cloudy weather

As can be seen, velocity measurements appeared less frequently than coordinate measurements. The error of position determination in the autonomous mode may increase if there is a change of altitude and wind speed parameters change significantly. In bad visibility conditions, the search area for the VNS extends. Thus, even if there is a significant position error, after leaving the cloud cover, the VNS is able to determine the location of the aircraft and correct the integrated solution.

VI. Conclusion

The autonomous navigation module provides stable determination of position, speed and orientation of the

aircraft in various meteorological conditions. The accuracy of position determination in the prediction mode is comparable to the accuracy of the same mode when correcting GNSS data. The module has the ability to connect a GNSS receiver and use it in the developed algorithm.

REFERENCES

[1] Mondek M., Hromčík M. On-line on-board wind estimation system for small UAV, 2015, 20th International Conference on Process Control (PC), IEEE, 2015, pp. 369-374.

[2] Pazychev D. B., Bakulev K. S. Navigation Complex for UAV //XXXI Saint Petersburg International Conference on Integrated Navigation Systems, 2024, pp. 200-207.

[3] Starostin N. P., Chernodarov A. V. Optical-inertial positioning of remotely piloted aircraft in the loss of satellite information //XXXI Saint Petersburg International Conference on Integrated Navigation Systems, 2024, pp. 102-105.

[4] Markov A. O., Burdakov S. F. Autonomous attitude determination of a small fixed-wing UAV //Achievements of Modern Radioelectronics, 2019, no. 9, pp. 34-41.

[5] Markov A. O. Autonomous strapdown attitude and heading reference system for a small agile UAV //XXVII Saint Petersburg International Conference on Integrated Navigation Systems, 2020, pp. 193-195.

[6] DeTone D., Malisiewicz T., Rabinovich A. Superpoint: Self-supervised interest point detection and description //Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2018, pp. 224-236.

[7] Potje G. et al. Xfeat: Accelerated features for lightweight image matching //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 2682-2691.

[8] Hartley R., Zisserman A. Multiple view geometry in computer vision, Cambridge university press, 2003.

[9] Kannala J., Brandt S. S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses //IEEE transactions on pattern analysis and machine intelligence, 2006, T. 28, no. 8, pp. 1335-1340.

TRIAD Method for Solving Orientation Problems

V.V. Matveev
Department of control devices,
FSBEI HE Tula State University,
Tula, Russia
matweew.valery@yandex.ru

A.V Kalikanov
Department of control devices,
FSBEI HE Tula State University,
Tula, Russia
kalikanov.aleksei@mail.ru

A.N. Khomyachkova
Department of control devices,
FSBEI HE Tula State University,
Tula, Russia
khoma05_19@mail.ru

***Abstract*— Solution of the orientation problem for stationary and moving objects by the TRIAD method based on measurements of gyroscopic, accelerometric and magnetometric sensors is considered. It is shown that the problem of initial orientation of inertial navigation systems can be solved based on the measurement of two vectors of any of the pairs of sensor triads. A new algorithm for the orientation problem solution based on two kinematic equations with respect to two orthogonal vectors is presented. It is shown that the proposed algorithm is not inferior in accuracy to widely used algorithms in such kinematic parameters as quaternions and Euler vector components.**

Keywords — TRIAD method, new orientation algorithm, gyroscope, accelerometer, magnetometer, integration.

I. INTRODUCTION

Improving methods and algorithms for determining the orientation of moving objects of various locations is a current area of scientific research and development in the field of aerospace instrumentation, robotics and automation. The use of three-axis modules (triads) of gyroscopes, accelerometers and magnetometers allows us to effectively solve problems of initial orientation and current orientation of both stationary and moving objects, based on measurements of angular velocity, specific force and the intensity of the Earth's magnetic field. However, each triad of sensors separately does not always allow us to accurately and effectively determine the angular position of an object relative to its center of mass.

To increase the accuracy and reliability of orientation determination, sensor signals are integrated. In this case, the connecting link of sensor signals is a complementary filter [1], which includes Wiener [2], Kalman [3], Madgwick [4] and other filters.

In addition to the above-mentioned filters, an effective method for integrating data to solve the problem of object orientation is the TRIAD (Tri-Axial Attitude Determination – three-axis determination of orientation [5]) method. The TRIAD method makes it possible to estimate the matrix of direction cosines between two coordinate systems by observing the projections of two support vectors [6]. This method is widely used to determine the orientation of spacecraft [5, 7], where measurements of magnetometric and star sensors are used as support vectors in the problem of studying the spin of an aircraft model using the vectors of geomagnetic and gravitational fields [8], in problems of the initial alignment of inertial navigation systems using the vectors of acceleration of gravity and the angular velocity of the Earth's daily rotation [9].

The aim of this work is to develop the TRIAD (Tri-Axial Attitude Determination) method to solve problems of orientation of stationary and moving objects based on gyroscopic, accelerometric and magnetometric measurements.

The work was carried out with the support of the grant of the Government of the Tula Region in the field of science and technology no. DS/162 dt.28.12.24.

II. SOLUTION TO THE PROBLEM OF ORIENTATION OF A MOVING OBJECT

As a reference, we introduce a geographic coordinate system, the origin O of which we place at the center of mass of the moving object, and we direct its axes to the cardinal points: OX_g - to the north, OY_g - along the true vertical upwards and OZ_g to the east, respectively (Fig. 1) [10].

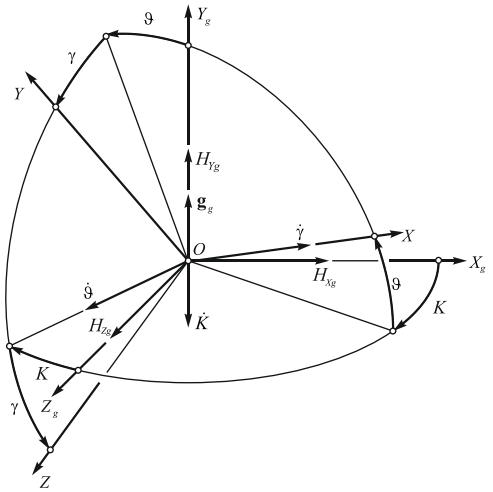


Fig. 1. Coordinate systems

Position of the coordinate system associated with the object $OXYZ$ is defined by three Euler-Krylov angles: heading K, pitch ϑ and roll γ. The problem of orientation of a moving object will be solved if we can find a triple of angles K, ϑ, γ or other kinematic parameters (direction cosines, Rodrigues-Hamilton parameters, components of the Euler vector, etc.) that uniquely determine the position of the associated coordinate system $OXYZ$ relative to the geographic coordinate system $OX_gY_gZ_g$.

Figure 1 also shows the components of the geomagnetic field strength vector $\mathbf{H}_g = \|H_{Xg} \quad H_{Yg} \quad H_{Zg}\|^T$ and the vector of acceleration of gravity $\mathbf{g}_g = \|0 \quad g \quad 0\|^T$, which are input physical quantities for obtaining measuring information from magnetometers and accelerometers, respectively. The measuring information from gyroscopic sensors is the projection of the absolute angular velocity vector $\boldsymbol{\omega}$ on the axis of the associated coordinate system. Thus, the following vectors form the measurement information of gyroscopes, accelerometers and magnetometers used to solve the orientation problem:

$$\begin{aligned}\boldsymbol{\omega} &= \mathbf{A}\mathbf{u}_g + \boldsymbol{\omega}^r,\\ \mathbf{n} &= \mathbf{A}(\mathbf{a}_g + \mathbf{g}_g),\\ \mathbf{H} &= \mathbf{A}\mathbf{H}_g,\end{aligned} \quad (1)$$

Where $\mathbf{A}$ – matrix of direction cosines of transition from the coordinate system $OX_gY_gZ_g$ to the system $OXYZ$, $\mathbf{u}_g$ – vector of absolute angular velocity of the coordinate system $OX_gY_gZ_g$, $\boldsymbol{\omega}^r$ – vector of the relative angular velocity of rotation of the system $OXYZ$ relatively $OX_gY_gZ_g$, $\mathbf{a}_g$ – vector of translational acceleration of the object, n – vector of apparent acceleration.

III. SOLUTION TO THE PROBLEM OF ORIENTATION OF A STATIONARY OBJECT

The task of initial orientation is to determine the matrix of direction cosines $\mathbf{A}$ at the initial moment of time, so that it can subsequently be used when running the algorithms of strapdown inertial navigation systems (SINS). Relations (1) on a fixed base take the form:

$$\begin{aligned}\boldsymbol{\omega} &= \mathbf{A}\left\|U\cos B \quad U\sin B \quad 0\right\|^T,\\ \mathbf{n} &= \mathbf{A}\left\|0 \quad g \quad 0\right\|^T,\\ \mathbf{H} &= \mathbf{A}\left\|H_{Xg} \quad H_{Yg} \quad H_{Zg}\right\|^T,\end{aligned} \quad (2)$$

where $U-$ is the angular velocity of the earth's daily rotation, B is the geodetic latitude.

The determination of the orthogonal matrix $\mathbf{A}$ by the TRIAD method assumes the presence of two support vectors defined by projections in two coordinate systems. Assume that $\bar{\mathbf{a}}_g = \left\|\bar{a}_{Xg} \quad \bar{a}_{Yg} \quad \bar{a}_{Zg}\right\|^T$ and $\bar{\mathbf{b}}_g = \left\|\bar{b}_{Xg} \quad \bar{b}_{Yg} \quad \bar{b}_{Zg}\right\|^T$ are the same two normalized vectors $|\bar{\mathbf{a}}_g| = 1$, $|\bar{\mathbf{b}}_g| = 1$ but defined by their projections in the geographic coordinate system. Here and below, the line over the symbol denotes the normalization operation. Let $\bar{\mathbf{a}} = \left\|\bar{a}_X \quad \bar{a}_Y \quad \bar{a}_Z\right\|^T$ and $\bar{\mathbf{b}} = \left\|\bar{b}_X \quad \bar{b}_Y \quad \bar{b}_Z\right\|^T$ be the same normalized vectors, but defined by projections in the coordinate system associated with the moving object. Let us form orthogonal triplets of unit vectors of coordinate axes from the considered vectors [5]:

$$\bar{\mathbf{a}}_g,\ \mathbf{e}_g = \frac{\bar{\mathbf{a}}_g \times \bar{\mathbf{b}}_g}{|\bar{\mathbf{a}}_g \times \bar{\mathbf{b}}_g|},\ \mathbf{r}_g = \bar{\mathbf{a}}_g \times \mathbf{e}_g. \quad (3)$$

$$\bar{\mathbf{a}},\ \mathbf{e} = \frac{\bar{\mathbf{a}} \times \bar{\mathbf{b}}}{|\bar{\mathbf{a}} \times \bar{\mathbf{b}}|},\ \mathbf{r} = \bar{\mathbf{a}} \times \mathbf{e}. \quad (4)$$

Direction cosine matrix of the transition from the coordinate system $OX_gY_gZ_g$ to the system $OXYZ$ based on relations (3), (4) is determined as follows [7]

$$\mathbf{A}^{TRIAD} = \begin{pmatrix}\bar{a}_X & e_X & r_X\\ \bar{a}_Y & e_Y & r_Y\\ \bar{a}_Z & e_Z & r_Z\end{pmatrix}\begin{pmatrix}\bar{a}_{Xg} & e_{Xg} & r_{Xg}\\ \bar{a}_{Yg} & e_{Yg} & r_{Yg}\\ \bar{a}_{Zg} & e_{Zg} & r_{Zg}\end{pmatrix}^T. \quad (5)$$

If the projections of vectors $\bar{\mathbf{a}}_g$ and $\bar{\mathbf{b}}_g$ are known exactly, but $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ measured without error, then the result of determining the matrix $\mathbf{A}^{TRIAD}$ it does not matter which vector is considered the first and which is the second.

Since the implementation of the TRIAD method requires two support vectors, the following combinations of measurements involved are possible when solving the initial orientation problem:

1) magnetometers and accelerometers;

2) gyroscopes and accelerometers;

3) gyroscopes and magnetometers.

From matrix (5) the heading, pitch and roll angles can be obtained [10].

In the initial alignment tasks of the SINS, normalized vectors of the acceleration of gravity and the angular velocity of the Earth's rotation are used, which in the geographic coordinate system have the following components

$$\bar{g}_g = \left\|0 \quad 1 \quad 0\right\|^T,\ \bar{U}_g = \left\|\cos B \quad \sin B \quad 0\right\|^T. \quad (7)$$

From the structure of vectors (7) it can be established that the matrix factor in calculating the matrix $\mathbf{A}^{TRIAD}$ (5) will have the following components:

$$\left\|\begin{matrix}\bar{a}_{Xg} & e_{Xg} & r_{Xg}\\ \bar{a}_{Yg} & e_{Yg} & r_{Yg}\\ \bar{a}_{Zg} & e_{Zg} & r_{Zg}\end{matrix}\right\|^T = \left\|\begin{matrix}0 & 1 & 0\\ 0 & 0 & -\frac{\cos B}{|\cos B|}\\ -\frac{\cos B}{|\cos B|} & 0 & 0\end{matrix}\right\| = \left\|\begin{matrix}0 & 1 & 0\\ 0 & 0 & -1\\ -1 & 0 & 0\end{matrix}\right\|, \quad (8)$$

for $B \in [-\pi/2, \pi/2]$.

In this case, the evaluation of the direction cosine matrix using the TRIAD method is simplified and takes the form

$$\mathbf{A}^{TRIAD*} = \left\|\begin{matrix}-r_X & \bar{a}_X & -e_X\\ -r_Y & \bar{a}_Y & -e_Y\\ -r_Z & \bar{a}_Z & -e_Z\end{matrix}\right\|. \quad (9)$$

Feature of the matrix $\mathbf{A}^{TRIAD*}$, in contrast to the initial orientation method described in [11], is that the initial alignment problem is solved without using information about the latitude of the object's location, the values of the gravity acceleration vectors, and the angular velocity of the Earth's daily rotation. This method, like any other, is applicable only to gyroscopes sensitive to the angular velocity of the Earth's daily rotation.

IV. A NEW ORIENTATION ALGORITHM BASED ON THE TRIAD METHOD

The TRIAD method considered above on a fixed base assumes measuring the components of two support vectors. The question arises about the possibility of constructing an algorithm for the orientation of a moving object based on the TRIAD method, but using measurements of only a triad of gyroscopes. The solution to the problem is possible from the following considerations.

Let $\bar{\mathbf{a}}_g$ and $\bar{\mathbf{b}}_g$ be abstract normalized vectors defined by their projections in a geographic coordinate system. Vector $\bar{\mathbf{a}}_g$ in the associated coordinate system can be written as:

$$\bar{\mathbf{a}} = \mathbf{A}\bar{\mathbf{a}}_g. \quad (10)$$

Differentiating both parts of equality (10) with respect to time, we have

$$\dot{\bar{\mathbf{a}}} = \dot{\mathbf{A}}\bar{\mathbf{a}}_g + \mathbf{A}\dot{\bar{\mathbf{a}}}_g .$$

Assuming that the vector $\bar{\mathbf{a}}_g$ is invariant in the geographic coordinate system $\dot{\bar{\mathbf{a}}}_g = 0$ and, given that the derivative of the direction cosine matrix can be represented by the Poisson equation [10]

$$\dot{\mathbf{A}} = -[\boldsymbol{\omega}^r\times]\mathbf{A},$$

where $[\boldsymbol{\omega}^r\times] = \begin{Vmatrix} 0 & -\omega_Z^r & \omega_Y^r \\ \omega_Z^r & 0 & -\omega_X^r \\ -\omega_Y^r & \omega_X^r & 0 \end{Vmatrix}$ – skew-symmetric matrix,

we obtain a kinematic equation of the form

$$\dot{\bar{\mathbf{a}}} = -[\boldsymbol{\omega}^r\times]\bar{\mathbf{a}}. \qquad (11)$$

Reasoning similarly, we write the kinematic equation for the second vector $\bar{\mathbf{b}}$:

$$\dot{\bar{\mathbf{b}}} = -[\boldsymbol{\omega}^r\times]\bar{\mathbf{b}}. \qquad (12)$$

Thus, having measurements of the angular velocity vector $\boldsymbol{\omega}^r = \boldsymbol{\omega} - \mathbf{A}\mathbf{u}_g$ and, by solving the kinematic equations (11) and (12), we can obtain the current projections of the vectors $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ in a fixed coordinate system. The solutions of the kinematic equations (11), (12) have the form

$$\bar{\mathbf{a}}(t) = \mathbf{C}(t)\bar{\mathbf{a}}(0),\ \bar{\mathbf{b}} = \mathbf{C}(t)\bar{\mathbf{b}}(0), \qquad (13)$$

where $\mathbf{C}(t) = -\exp(\int_0^t [\boldsymbol{\omega}^r\times]dt)$ transformation matrix characterizing the observed from the moving coordinate system $OXYZ$ change of vectors $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$.

Subsequent application of the TRIAD method allows us to find the current matrix of direction cosines and thus solve the problem of orientation of a moving object.

Let's choose orthogonal vectors as support vectors: $\bar{\mathbf{a}}_g = \|0 \quad 1 \quad 0\|^T$ and $\bar{\mathbf{b}}_g = \|1 \quad 0 \quad 0\|^T$. In this case, relation (8) is valid and, in addition $|\bar{\mathbf{a}}\times\bar{\mathbf{b}}| = 1$. Thus, the direction cosine matrix takes the form

$$\mathbf{A}^{TRIAD*} = \begin{pmatrix} -\left(\bar{\mathbf{a}}\times(\bar{\mathbf{a}}\times\bar{\mathbf{b}})\right)_X & \bar{a}_X & -(\bar{\mathbf{a}}\times\bar{\mathbf{b}})_X \\ -\left(\bar{\mathbf{a}}\times(\bar{\mathbf{a}}\times\bar{\mathbf{b}})\right)_Y & \bar{a}_Y & -(\bar{\mathbf{a}}\times\bar{\mathbf{b}})_Y \\ -\left(\bar{\mathbf{a}}\times(\bar{\mathbf{a}}\times\bar{\mathbf{b}})\right)_Z & \bar{a}_Z & -(\bar{\mathbf{a}}\times\bar{\mathbf{b}})_Z \end{pmatrix}. \qquad (14)$$

Based on relations (13) and (14), it is possible to construct a numerical algorithm for solving the problem of orientation of a moving object using the TRIAD method based only on gyroscopic measurements.

The analysis of the proposed orientation algorithm was carried out on the basis of mathematical modeling with the following laws of change in the angles of course, pitch and roll:

$$\begin{aligned} K(t) &= 3^\circ\sin(2\pi\cdot 0{,}05t), \\ \vartheta(t) &= 3^\circ\cos(2\pi\cdot 0{,}05t), \\ \gamma(t) &= 10^\circ\sin(2\pi\cdot 0{,}04t). \end{aligned} \qquad (15)$$

The initial values of the support vectors in the associated coordinate system were determined by the following component values: $\bar{\mathbf{a}}[0] = \|0{,}052 \quad 0{,}999 \quad 0\|^T$, $\bar{\mathbf{b}}[0] = \|0 \quad 0 \quad 1\|^T$. The number of the last member of the series expansion was taken as $n = 2$. The proposed algorithm was compared with algorithms frequently used in the practice of constructing SINS:

–in quaternions [5]

– with the Euler vector based on the Bortz kinematic equation [14]:

After calculating the current matrices and quaternions, a transition to the Euler-Krylov angles was performed, followed by their comparison with the actual values of the heading, pitch, and bank angles. The graphs of the algorithm errors are shown in Fig. 2–4.

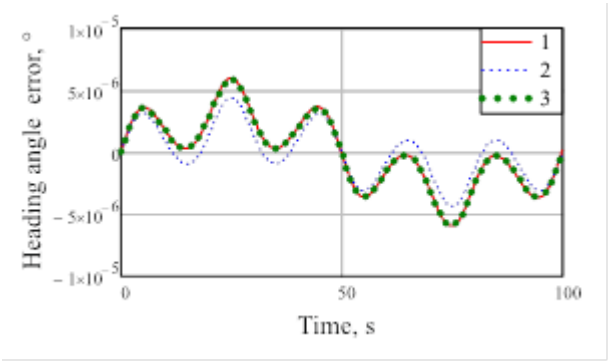


Fig. 2. Algorithm errors for heading angle: 1 – in quaternions; 2 – with Euler vector; 3 – TRIAD method

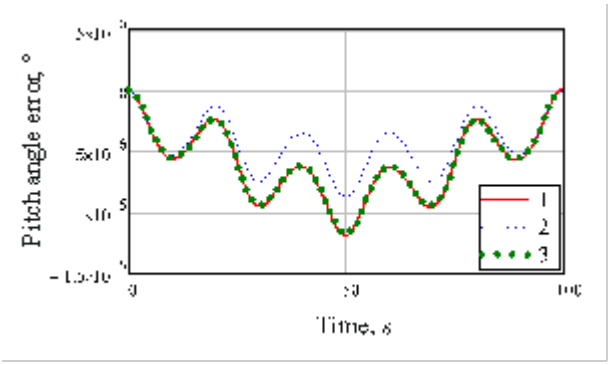


Fig. 3. Errors of algorithms for pitch angle: 1 – in quaternions; 2 – by means of the Euler vector; 3 – TRIAD method

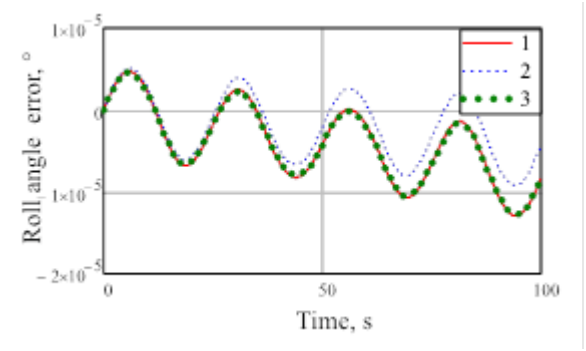


Fig. 4. Errors of the algorithms for the bank angle: 1 – in quaternions; 2 – using the Euler vector; 3 – the TRIAD method

From the simulation results, it can be concluded that the proposed orientation algorithm is comparable in accuracy to

the algorithm in quaternions and can also be used in SINS. The accuracy of the algorithm can be increased by using the corresponding number of the series member n. In addition, the proposed orientation algorithm is easily integrated with accelerometric and magnetometric sensors when constructing strapdown orientation systems. For example, the kinematic equation (11) for the vector $\bar{\mathbf{a}}_g = \|0 \quad 1 \quad 0\|^T$ can be interpreted as an estimate of the normalized vector of gravitational acceleration. If we assume that the second vector $\bar{\mathbf{b}}_g[0]$ at the initial moment of time corresponds to the normalized vector of the geomagnetic field $\bar{\mathbf{H}}_g$, then the kinematic equation (12) will give its estimate during the movement of the object. As a result, the equations of complementary filters, combining the measurements of three triads of sensors, can be written as:

$$\bar{\mathbf{a}}^K[k+1] = \alpha_1 \mathbf{C}[k+1]\bar{\mathbf{a}}^K[k] + (1-\alpha_1)\bar{\mathbf{n}}[k+1], \qquad (16)$$

$$\bar{\mathbf{b}}^K[k+1] = \alpha_2 \mathbf{C}[k+1]\bar{\mathbf{b}}^K[k] + (1-\alpha_2)\bar{\mathbf{H}}[k+1], \qquad (17)$$

where $\bar{\mathbf{a}}^K[k]$, $\bar{\mathbf{b}}^K[k]$ – integrated reference vectors of gravity acceleration and geomagnetic field; $\alpha_1 = [0, 1]$, $\alpha_2 = [0, 1]$ – coefficients that determine the weight of the corresponding measurements; $\bar{\mathbf{n}}[k+1]$, $\bar{\mathbf{H}}[k+1]$ – normalized measurement vectors of accelerometric and magnetometric sensors.

After calculating the vectors $\bar{\mathbf{a}}^K[k+1]$ and $\bar{\mathbf{b}}^K[k+1]$ a matrix of direction cosines is formed $\mathbf{A}^{TRIAD}[k+1]$ based on the relation (5). A special case of matrix representation $\mathbf{A}^{TRIAD*}$ in the form (9) or (13) is not applicable in this case.

V. CONCLUSION

The efficiency of the TRIAD method application for solving the problems of initial orientation and current orientation of objects based on gyroscopic, accelerometric and magnetometric measurements is shown. On a fixed base, the initial matrix of direction cosines can in principle be found based on any of the pairs of sensor triads: accelerometers-magnetometers, accelerometers-gyroscopes, magnetometers-gyroscopes. It is established that the initial orientation by the TRIAD method based on gyroscope and accelerometer measurements does not require information on the latitude of the object's location.

A new orientation algorithm based on gyroscopic measurements and kinematic equations relative to two orthogonal vectors is proposed. The algorithm is not inferior in accuracy to the algorithm in quaternions, and is comparable to the orientation algorithm with the Euler vector, and can be successfully used in the implementation of strapdown inertial navigation systems. Among the advantages of the algorithm is the efficiency of its integration with accelerometric and magnetometric measurements in the implementation of strapdown orientation systems.

REFERENCES

[1] Brown R.G., Hwang Patrick Y.C. Introduction to random signals and applied Kalman filtering, John Wiley & Sons. New York, 1997.

[2] Shvedov A.P. Methods for improving the accuracy of information-measuring systems for the orientation of moving objects. Abstract of a dissertation for the degree of Cand. Sci. (Eng.) in specialty 05.11.16 – Information-measuring and control systems (in industry). Tula: Tula State University Press. 2010. – 20 p.

[3] Gui P., Tang L, Mukhopadhyay S.C. MEMS based IMU for tilting measurement: Comparison of complementary and kalman filter based data fusion. No. 6. 2015. P. 2004–2009. DOI:10.1109/ICIEA.2015.7334442.

[4] Madgwick, S., Harrison, A., Vaidyanathan R. Estimation of IMU and MARG orientation using a gradient descent algorithm. Proceedings of the IEEE Int. Conf. Rehabil. Robot. 2011. pp. 179 – 185.

[5] Markley L., Crassidis J. Fundamentals of Spacecraft Attitude Determination and Control. New York, Springer. 2014. 486 p.

[6] Hashim H.A. Attitude Determination and Estimation using Vector Observations: Review, Challenges and Comparative Results, arXiv preprint, 2020. 51 p.

[7] Belkhiat D.E.C., Fourati H. Multisensor Attitude Estimation: Fundamental Concepts and Applications. Boca Raton, CRC Press, 2016. 607 p.

[8] Vyalkov A.V. Determination of orientation angles from observations of vectors in the problem of studying the spin of an aircraft model, Gyroscopy and Navigation. Vol. 28, No. 3 (110), 2020. Pp. 43 – 59.

[9] Golovan A.A., Parusnikov N.A. Mathematical foundations of navigation systems: Part II: Applications of optimal estimation methods to navigation problems. - 2nd ed., corrected. and add. - Moscow: MAKS Press, 2012. -172 p.

[10] Orientation and navigation of moving objects: modern. information technologies / Ed. B. S. Aleshina, K. K. Veremeenko, A. I. Chernomorskogo. - M.: FIZMATLIT, 2006. - 424 p.

[11] Gyroscopic systems. Gyroscopic devices and systems V.2: Textbook for universities in the specialty "Gyroscopic devices and instruments". / D.S. Pelpor, I.A. Mikhalev, V.A. Bauman et al.; Ed. D.S. Pelpor. - 2nd ed., revised and enlarged. - M. Higher. school, 1988. 424 p.

[12] Savege P. G. Strapdown analytics. Strapdown Association, Part 1. Inc. Maple Plain, Minnesota, 2000. 817 p.

The Adaptation and Application of the SINS Fine Alignment in the Case of a Vessel's Rolling

N.A. Efimov

JSC "CRI "Dolphin"
Moscow, Russia
efimovna83@ya.ru

A.A. Golovan

Lomonosov Moscow State University, Moscow, Russia
aagolovan@yandex.ru

V.L. Hirnaya

JSC "CRI "Dolphin"
Moscow, Russia
zvl001@niidelfin.ru

V.V. Sidelnikova

JSC "CRI "Dolphin"
Moscow, Russia
val-sidelnikova@yandex.ru

I.A. Papusha

JSC "CRI "Dolphin"
Moscow, Russia
ipapusha@yandex.ru

Abstract — The results of testing the algorithm for the initial alignment of shipborne strapdown inertial navigation system (SINS), implemented in the "gyro azimuth horizon compass GAGK1 / 40" product developed by JSC CRI "Dolphin", are presented. For the initial alignment under the conditions of vessel's pitching, the SINS positional aiding was applied using a close loop Kalman filter. The algorithm was tested on test bench.

Key words — strapdown inertial navigation system, initial alignment, alignment in motion, alignment under vessel's pitching, the Kalman filter

I. INTRODUCTION

Shipborne inertial navigation systems are essential for ensuring the accuracy and reliability of navigation on both ships and submarines. As is known, vessel rolling leads to a change in the orientation and position of the vessel in space, which, in turn, affects the accuracy of the accelerometers and gyroscopes readings. The mentioned dynamic motion requires the implementation of the alignment algorithm that can not only compensate for the effect of pitching, but also adapt the operation of inertial sensors.

One way to mitigate the effects of pitching is through the implementation of the SINS fine alignment algorithm. This adapted version of the algorithm has been incorporated into the GAHC1/40 instrument.

The GAHC1/40 gyroazimuth-horizon compass is a ship borne inertial navigation and gyro stabilization system developed by JSC Central Research Institute "Dolphin". It includes the central N22L unit that generates navigation, orientation, and stabilization parameters. The SINS of the GAHC1/40 product utilizes high-accuracy inertial sensors, including three laser gyroscopes with integrated thermal stabilization and three quartz accelermeters with integrated thermistors. In the GAHC1/40, the use of laser-based sensors reduces the startup time because these gyroscopes do not require preheating system. Power supply consumption is also greatly reduced due to the absence of additional thermo control devices.

The GAHC1/40 product is designed for use on mobile vessels, and its application assumes that the initial alignment of SINS will take place in conditions of vessel's pitching. Traditionally, this problem has been solved by developers using the gyrocompass method, which is based on SINS aiding via velocity and position data [1]. As an alternative, one can consider the algorithm proposed in [2], which has also been tested in practical applications. This algorithm is based on approximating the readings from the SINS accelerometers when latter projected onto the axes of the instrument's trihedral frame "frozen" in inertial space. For the initial alignment of the GAHC1/40 instrument, a widely used algorithm in airborne applications has been selected [3].

II. THE FINE ALINGMENT ALGORITHM UNDER VESSEL'S PITCHING

The alignment of the navigation grade SINS, is carried out in two stages: the first stage is the so-called "coarse" or "algebraic" alignment, and the second stage is the fine alignment. The algorithm of the first stage consists in averaging the readings from the inertial sensors to obtain preliminary estimates of the attitude angles of SINS instrument reference frame with respect to the geodetic coordinate system. Thus, the first estimates of the values of pitch and roll are derived from the mean values of the accelerometers, and, in conditions of stillness, the approximate heading angle can be calculated from the mean values of angular rate sensors (ARC). In the presence of uncontrolled angular and translational movements of the SINS body frame, it is impossible to obtain an initial approximation of the heading angle using this method. However, this does not prevent the implementation of the next stage of the alignment. The second stage, also known as the fine alignment stage, involves the operation of SINS in autonomous inertial mode with simultaneous SINS aiding using the starting point's coordinate. The purpose of the SINS aiding is to estimate the errors in the attitude angles of the instrument's reference frame, which were determined during the first stage. The estimation of the state vector for the SINS error model is performed using the Kalman filter algorithm, in its close loop mode [4]. The absence of a rough estimate of the heading angle does not significantly affect the convergence of the algorithm. With properly

configured Kalman filter settings, the second phase of the alignment takes approximately half an hour.

III. The Models of the Estimation Algorithm

When describing the estimation algorithm used in the process of fine alignment, we adhere to the terminology defined in [3]. In this technique, the total error in determining the relative linear velocity is divided into dynamic and kinematic components. The estimated state vector in the axes of a geodetic reference frame when SINS wander azimuth option of orientation is activated has the following components:

$$x = (\Delta r_1, \Delta r_2, \delta V_1, \delta V_2, \alpha_1, \alpha_2, \beta_3)^{\mathrm{T}},$$

where $\Delta r_1, \Delta r_2$ are the positional errors in the horizontal plane,

$\delta V_1, \delta V_2$ are the dynamic components of the total horizontal velocity errors,

α_1, α_2 are the misalignment errors of the instrument vertical, β_3 is the azimuth kinematic error.

The combined form of the SINS error equations:

$$\begin{aligned}
\Delta\dot{r}_1 &= \delta V_1 + \beta_3 V_2' \\
\Delta\dot{r}_2 &= \delta V_2 - \beta_3 V_1' \\
\delta\dot{V}_1 &= 2u_3\delta V_2 - g\alpha_2 + \Delta f_1^s \\
\delta\dot{V}_2 &= -2u_3\delta V_1 + g\alpha_1 + \Delta f_2^s \\
\dot{\alpha}_1 &= -u_3\frac{\Delta r_1}{a} - \frac{\delta V_2}{a} + u_3\alpha_2 - u_2\beta_3 + \nu_1^s \\
\dot{\alpha}_2 &= -u_3'\frac{\Delta r_2}{a} + \frac{\delta V_1}{a} - u_3\alpha_1 + u_1\beta_3 + \nu_2^s \\
\dot{\beta}_3 &= u_1\frac{\Delta r_1}{a} + u_2\frac{\Delta r_2}{a} + u_2\alpha_1 - u_1\alpha_2 + \nu_3^s,
\end{aligned}$$

where g is the nominal gravity value,

a is the semimajor axis of the Earth's ellipsoid,

V_1', V_2' are the components of the relative linear velocity vector in the axes of local geodetic reference frame with wander azimuth option,

$u_x = (u_1, u_2, u_3)^{\mathrm{T}}$ is the vector of the angular rate of the Earth's rotation,

$\Delta f_i^s, \nu_i^s$ are random error components of inertial sensors, such as white noise.

The constant components of the errors of inertial sensors (biases of accelerometers and gyroscopes) at the alignment stage are not among the estimated parameters

Positional measurements for aiding are formed using data about the initial coordinate of the SINS:

$$\begin{aligned}
z_1^{pos} &= \Delta r_\varphi \sin\varepsilon' + \Delta r_\lambda \cos\varepsilon' \\
z_2^{pos} &= \Delta r_\varphi \cos\varepsilon' - \Delta r_\lambda \sin\varepsilon'
\end{aligned}$$

Where

$$\begin{aligned}
\Delta r_\varphi &= (\varphi' - \varphi^0)R_N \\
\Delta r_\lambda &= (\lambda' - \lambda^0)R_E\cos\varphi'
\end{aligned}$$

$$R_E = \frac{a}{\sqrt{1 - e^2\sin^2\varphi'}} \quad R_N = \frac{a(1-e^2)}{(1 - e^2\sin^2\varphi')^{3/2}}$$

Here:

R_N, R_E – radii of curvature of the meridional section and the prime vertical,

e is the eccentricity of the Earth's ellipsoid,

ε' is the angle that characterizes the azimuthal orientation of the SINS geodetic reference frame with respect ot the geodetic grid,

(φ', λ') are the coordinates generated by the SINS navigation algorithm,

(φ^0, λ^0) are the initial geodetic coordinates.

According to the estimates obtained using the Kalman filter, corrections to the autonomous inertial solution are being calculated. At each step, the developed corrections are introduced into the navigation algorithm through feedback. The filter settings are selected based on the parameters of the typical motion of the SINS carrier during the alignment and the passport characteristics of the inertial sensors. In the absence of an initial approximation of the heading angle, the initial azimuth error covariance for β_3 is set to $P_{\beta_3} = \pi$ radians.

IV. The Results of Testing

The algorithm has been tested on the specialized "Captain" stand, which simulates ship pitching with preset parameters. To prepare for the tests, one need to place the instrument on a horizontal plane using a KO-10TS optical digital quadrant to measure the reference angles of the keel and side pitch. One will also need to use 3T3KA theodolites to determine the azimuth of the direction set by the mirror of the mounting plate. This will help to identify the reference heading value using a basic control instrument. On the ship's pitching bench, tests were conducted with the following parameters: onboard pitching with an amplitude of 18 degrees and a period of 5 seconds, keel pitching with an amplitude of 15 degrees and a period of 5 seconds, and heading with an amplitude of 10 degrees and a period of 5 seconds. The attitude parameters obtained using the onboard SINS alignment met the technical specifications, as the heading error after 30 minutes of operation should not exceed 6 degrees, and the roll and pitch should be within 1 degree.

To investigate the issue of reducing the standby time of the SINS in pitching conditions, numerical simulation was conducted based on the recorded data from the inertial sensors of the N22L device. During the experiment, the stand platform remained stationary relative to the ground for 30 minutes. After that, it performed harmonic oscillations for another 30 minutes with a period of 30 seconds. The amplitude of the roll and trim was 8 degrees, and the amplitude of the course was 3 degrees. The data was processed in the following way. The "reference" heading angle was generated by simulating the SINS algorithm in a stationary area and then correcting it in a operation mode. The heading estimation - the result of the pitching simulation - was obtained by initializing the simulation algorithm on a site with a movable platform (see Fig. 1).

The deviation of the heading estimate obtained during the pitching alignment from the "reference" value is shown in Fig. 2. The accuracy range of 6 degrees was achieved within a time interval of 10 minutes, demonstrating the high efficiency of the algorithm proposed.

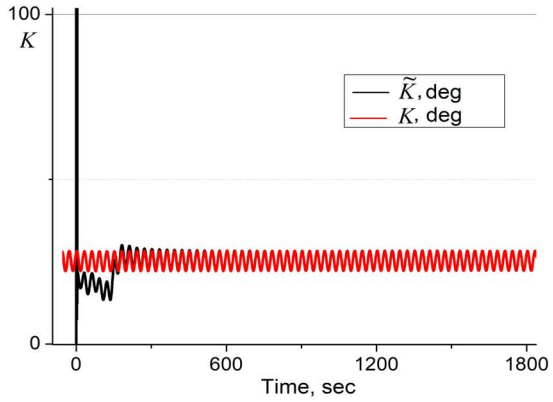


Fig. 1. Estimation of the heading angle (black line) compared to the reference value (red line).

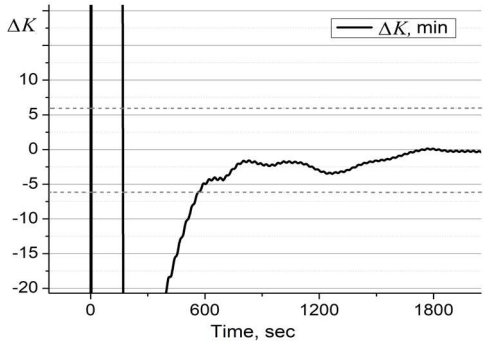


Fig. 2. Error in Estimating the course angle relative to the reference value

V. Conclusion

The paper presents the results of implementing the SINS initial alignment algorithm in marine environments for a navigation grade inertial system. The algorithm is based on the positional aiding of SINS using the Kalman filter in a closed loop mode. The results of the bench tests of the GAHC1/40 instrument demonstrated high accuracy and showed the feasibility of implementing a two-stage algorithm of the alignment in conditions of ship pitching.

References

[1] Emel'yantsev, G.I., Stepanov, A.P. & Blazhnov, B.A. Initial Alignment of Shipborne SINS under Ship Motion. *Gyroscopy Navig.* 2020, **11**, pp. 277–284. https://doi.org/10.1134/S2075108720040045

[2] Barantsev G.O., Golovan A.A., Kuznetsov P.Y. Initial alignment method for a strapdown inertial navigation system on a swing base, in: *Vestnik Moskovskogo universiteta. Seriya 1: Matematika. Mekhanika, Mechanics,* Izdatel'stvo *Moskovskogo universiteta*, 2021, V. 76, pp. 136-141.

[3] Vavilova, N.B., Golovan, A.A. & Parusnikov, N.A. Mathematical foundations of navigation systems (in Russian), Moscow University Press, Moscow, 2020, p. 164.

[4] Vavilova, N.B., Golovan, A.A. & Parusnikov, N.A. Problem of information equivalent functional schemes in aided inertial navigation systems. *Mech. Solids* **43**, 391–399 (2008). https://doi.org/10.3103/S0025654408030114

SESSION III – RELEVANT ISSUES OF THEORY

On the Control of Spatial Elastic Standing Waves in a Complete Spherical Shell

S.E. Perelyaev
Institute of Problems of Mechanics
A.Yu. Ishlinsky RAS
119526, Moscow, Russia
Vernadsky Avenue, 101-1
e-mail: ipm@ipmnet.ru

V.F. Zhuravlev
Institute of Problems of Mechanics
A.Yu. Ishlinsky RAS
119526, Moscow, Russia
Vernadsky Avenue, 101-1
e-mail: ipm@ipmnet.ru

Abstract — **This publication is dedicated to the presentation of the fundamentals of the fundamental theory of the "generalized" Foucault pendulum, implemented based on a fundamentally new 3-D spherical resonance gyroscope, whose basic sensitive element is a high-quality hollow spherical resonator.**

The generalization of the Brian effect [1] to the spatial case is considered and investigated—specifically, an elastic spherically symmetric solid body with a free boundary subjected to mass forces is examined. The planar rotation of the spherical resonator is replaced by spatial rotation. In such an experiment, the generalized effect of Brian's inertia also becomes spatially relevant.

Keywords: Foucault pendulum, semi-spherical wave solid-state gyroscope, 3-D resonance gyroscope, 3-D spherical resonance gyroscope.

I. INTRODUCTION

This work examines the classical equations of elastic vibrations on the surface of a high-quality complete spherical shell. An elastic spherically symmetric solid body with a free boundary subjected to mass forces is investigated. The principal vector of forces acting on the three-dimensional elastic body is assumed to be zero without loss of generality. Under the influence of the principal moment, the spherical elastic body changes its orientation in the considered three-dimensional space. To mathematically describe spatial elastic deformations of the complete spherical shell, a system of three-dimensional coordinates is introduced, which is related to the spherical body itself in such a way that specified initial conditions are satisfied. These initial conditions characterize a coordinate trihedron (a three-faced polyhedron) relative to which the body does not move or rotate on average across all material particles. The problem of eigen vibrations of a complete spherically symmetric high-quality shell admits a group of three-dimensional rotations *SO*(3). This gives rise to the inertial effect of elastic spatial mechanical standing waves on the surface of the rotating complete spherical shell. The phase manifold of spatial elastic mechanical standing waves and, accordingly, the evolutionary basis on this phase manifold are considered.

This work was carried out within the framework of the State Assignment (registration No. 124012500443-0) at the A.Yu. Ishlinsky Institute for Problems in Mechanics of the Russian Academy of Sciences.

II. EFFECT OF INERTIALITY OF ELASTIC FORMS IN A COMPLETE ROTATING SPHERICAL SHELL

An orthogonal coordinate system is considered, related to the original coordinate system by a known orthogonal transformation. The matrix of this orthogonal transformation is a function of time and has a self-adjoint form. It is shown that the three-dimensional differential equation of spatial oscillations admits a solution of the type of mechanical standing wave. Thus, if a spatial standing wave of mechanical elastic vibrations with some pure tone is excited in a stationary spherical elastic body and the body is set into rotation with an arbitrary angular velocity that varies over time, the excited spatial standing wave will rotate relative to the body itself according to a certain law described by an equation similar to Poisson's equation, lagging behind the specified rotation of the body. It can be demonstrated that the angular velocity of precession of the mechanical spatial standing wave relative to the body is proportional to the spatial Brian coefficient and the angular velocity of the body itself relative to absolute space. This study theoretically proves the existence of a fundamentally new spatial effect of inertia of elastic waves on a complete spherical shell, which can be utilized for studying the operation of a three-dimensional (3D) spherical resonance gyroscope or a triaxial wave solid-state gyroscope (WSG). Thus, the complete elastic spherical shell can effectively serve as a spatial 3D sensor for inertial information during arbitrary angular movements and rotations of the solid body itself.

III. EQUATIONS OF ELASTIC VIBRATIONS IN THE BASIS OF EIGEN FORMS

Let us proceed to the consideration of elastic vibrations of a thin spherical shell. We take an elastic spherically symmetric solid body $\boldsymbol{x} = (x_1 x_2 x_3)$ with a free boundary, subjected to mass forces with density f. The resultant vector of all forces acting on the elastic volume body $\int_V \boldsymbol{f} \mathrm{d}m$ will be assumed to be equal to

zero without loss of generality. Under the action of the resultant moment $\int_V \boldsymbol{r} \times \boldsymbol{f} dm$ the considered body changes its orientation in space, where $\boldsymbol{r}$ is the radius vector of an arbitrary point in the body, dm is the mass element, and V is the volume occupied by the body. To describe the elastic deformations of the thin spherical shell, we introduce a coordinate system $x_1x_2x_3$ that is attached to the elastic body under consideration in such a way that the following conditions are satisfied:

$$\int_V \boldsymbol{x} dm = 0, \quad \int_V \boldsymbol{r} \times \boldsymbol{x} dm = 0 \tag{3.1}$$

where $\boldsymbol{x} = (x_1x_2x_3)$ presents the elastic displacement of a material point that occupied position $\boldsymbol{r}$ in the undeformed state. The conditions (4.1) characterize a coordinate tetrahedron with respect to which the elastic body, on average (over all particles), does not translate or rotate. The following problem is posed: knowing the absolute angular velocity $\boldsymbol{\omega}(t)$ of the tetrahedron $x_1x_2x_3$ projected onto its own axes, determine how the waves of elastic deformations behave. We will write down the principle of D'Alembert-Lagrange for the considered body following [2]

$$\int_V \Big[\ddot{x} + \boldsymbol{\omega} \times \big(\boldsymbol{\omega} \times (\boldsymbol{r} + \boldsymbol{x})\big) + \dot{\boldsymbol{\omega}} \times (\boldsymbol{r} + \boldsymbol{x}) + 2\boldsymbol{\omega} \times \dot{\boldsymbol{x}} + \frac{1}{\boldsymbol{\rho}} \boldsymbol{\nabla}\Pi - f\Big] \delta \boldsymbol{x} \times dm = 0 \tag{3.2}$$

Here, ρ is the density that depends only on the magnitude of the radius $|\boldsymbol{r}|$, and $\boldsymbol{\nabla}\Pi$ is the gradient of the quadratic functional in linear elasticity theory. The coordinates that define the angular position of the body as a whole do not vary; it is assumed that the angular velocity $\omega(t)$ is a known function of time. To select generalized coordinates, let us consider the special case where $\omega = 0$. In monograph [2], it is shown that the spectrum of eigenvibrations of a free solid body, under the conditions introduced above (3.1), is discrete. This means that the increasing sequence of eigenfrequencies $\nu_1 \le \nu_2 \le \nu_3$ ... is unbounded, and the eigenfunctions$h_1(r), h_2(r)$, ... corresponding to these sequences of frequencies form an orthonormal system of functions that is complete in the configuration space of the problem under consideration:

$$\int_V h_n(r)\, h_1(r) dm = \delta_n^l \tag{3.3}$$

This allows us to introduce independent Lagrangian coordinates that describe the degrees of freedom during the deformation of the body, in the general case where $\omega(t) \neq 0$, as follows:

$$x = \sum_{n=1}^{\infty} q_n(t) h_n(r) \tag{3.4}$$

The problem of eigen vibrations of a spherically symmetric free body allows for the group *SO*(3), which means that the spectrum of eigenfrequencies is degenerate and consists of a sequence of frequencies that are at least triply degenerate: $\nu_1 = \nu_2 = \nu_3 \le \nu_4 = \nu_5 = \nu_6 \le \cdots$. In this case, the configuration space is represented as a direct product of three-dimensional eigensubspaces:

$$\{h_1, h_2, h_3\} \times \{h_4, h_5, h_6\} \times \ldots$$

Fix the number m of an arbitrary eigensubspace and introduce notations for the corresponding generalized coordinates: $q_{3m-1} = u,\ q_{3m-2} = v,\ q_{3m} = w\ \ (m = 1,2\ldots)$. Substituting formula (3.4) as well as the expression $\delta x = \sum_{n=1}^{\infty} \delta q_n h_n(r)$ into formula (3.2) and setting the coefficients of the independent variations δq_n to zero, we obtain an infinite system of ordinary differential equations of the second order with respect to the coordinates q_n of the following form:

$$\begin{aligned}
&\ddot{u} + au + bv + cw - v(\dot{\omega}, k_3) + w(\dot{\omega}, k_2) - \\
&-2\dot{v}(\omega, k_3) + 2\dot{w}(\omega, k_2) + F_1 + L_1 = 0 \\
&\ddot{v} + bu + bv + ew + u(\dot{\omega}, k_3) - w(\dot{\omega}, k_1) + \\
&+2\dot{u}(\omega, k_3) - 2\dot{w}(\omega, k_1) + F_2 + L_2 = 0 \\
&\ddot{w} + cu + ev + fw - u(\dot{\omega}, k_2) + v(\dot{\omega}, k_1) - \\
&-2\dot{u}(\omega, k_2) + 2\dot{v}(\omega, k_1) + F_3 + L_3 = 0
\end{aligned} \tag{3.5}$$

In the system of differential equations (3.5), the scalar coefficients have the following form:

$$a = \int_V (h_{3m-2}, \omega)^2 dm - \omega^2,$$
$$b = \int_V (h_{3m-2}, \omega)(h_{3m-1}, \omega) dm\,,$$
$$c = \int_V (h_{3m-2}, \omega)(h_{3m}, \omega) dm, d = \int_V (h_{3m-1}, \omega)^2 dm - \omega^2,$$
$$e = \int_V (h_{3m-1}, \omega)(h_{3m}, \omega) dm,\ f = \int_V (h_{3m}, \omega)^2 dm - \omega^2,$$
$$F_1 = \int_V (\nabla\Pi, h_{3m-2})\, dV, \quad F_2 = \int_V (\nabla\Pi, h_{3m-1})\, dV,$$
$$F_3 = \int_V (\nabla\Pi, h_{3m})\, dV.$$

L_1, L_2, L_3 represent linear functions of the generalized coordinates corresponding to other eigensubspaces. The presence of these terms indicates that systems of the type (3.5) for different subspaces are not independent of one another. In deriving the equations (3.5), it was assumed for simplicity that the mass forces $\boldsymbol{f}$ are orthogonal to all eigenfunctions:

$$\int_V \boldsymbol{f} h_n(r) dm = 0.$$

This means that the force $\boldsymbol{f}$ contains only a constant component ($\int_V \boldsymbol{r} \times \boldsymbol{f} dm \neq 0$), which ensures the rotation of the body with an angular velocity $\omega(t)$. In expression (3.5), the vector coefficients k_1, k_2, k_3 take the form:

$$\boldsymbol{k_1} = \int_V h_{3m-1} \times h_{3m} dm, \quad \boldsymbol{k_2} = \int_V h_{3m} \times h_{3m-2} dm,$$
$$\boldsymbol{k_3} = \int_V h_{3m-2} \times h_{3m-1} dm.$$

Due to spherical symmetry, the choice of eigenvectors $(\boldsymbol{h}_{3m-2}, \boldsymbol{h}_{3m-1}, \boldsymbol{h}_{3m})$ can be made such that the following identities hold:

$$k_1 = \upsilon(1,0,0), k_2 = \upsilon(0,1,0), k_3 = \upsilon(0,0,1) \tag{3.6}$$
$$\upsilon = \pm|k_1| = \pm|k_2| = \pm|k_3| = \pm\left|\int_V h_{3m-1} \times h_{3m} dm\right|.$$

Using the Cauchy-Bunyakovsky inequality for (3.6), we obtain $0 \le \upsilon \le 1$. If we now introduce notations:

$$z = \begin{pmatrix} u \\ v \\ w \end{pmatrix},\ \omega = \begin{pmatrix} p \\ q \\ r \end{pmatrix},\ L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix},\ G = \begin{pmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{pmatrix},$$

then the equations (3.5) can be rewritten in the following compact vector form [2]:

$$\ddot{z} + Az + \upsilon\dot{G}z + 2\upsilon G\dot{z} + L = 0 \tag{3.7}$$

where A is a symmetric matrix of positional forces, composed of the coefficients of elastic forces F_1, F_2, F_3 and corresponding scalar coefficients a, b, c, d, e, f.

IV. SPATIAL PRECESSION OF A MECHANICAL STANDING WAVE

The vector differential equation (3.7) defines the spatial evolution of the *m*-th mode of oscillations of a free spherical elastic solid body, induced by the presence of rotation. This

evolution is determined by two circumstances. First, the mode of oscillation itself directly responds to the body's rotation, which is indicated by the presence of terms G and $\dot{G}$. in (3.7). Second, the considered mode of oscillation is influenced by other modes. It is important to note that this influence is negligible since, for example, when solving equations (3.7) using the averaging method, all terms defined by linear functions of generalized coordinates L vanish in the first approximation. Furthermore, there is another noteworthy fact. We observe that there exists a coordinate system $z \to y$: $z = My$, where M is a time-dependent orthogonal transformation matrix, in which equation (3.7) takes on a self-adjoint form when the identity $L = 0$ holds. In this coordinate system, equation (3.7) admits solutions of the standing wave type. Let us demonstrate this. By substituting the equation $z = My$ into the previously obtained expression (3.7), we find

$$\ddot{y} + 2M^T(\dot{M} + \upsilon GM)\dot{y} + M^T(\ddot{M} + 2\upsilon G\dot{M} + \upsilon\dot{G}M + \\ + AM)y = 0 \qquad (4.1)$$

We require the satisfaction of the following first-order differential equation, which is derived from equation (4.1):

$$\dot{M} = -\upsilon GM \qquad (4.2)$$

As a result, we obtain a second-order differential equation of the form:

$$\ddot{M} = -\upsilon\dot{G}M + \upsilon^2G^2M$$

and by substituting it into (4.1), we find the following second-order equation:

$$\ddot{y} + M^T(A - \upsilon^2G^2)My = 0 \qquad (4.3)$$

Thus, if a mechanical standing wave of elastic oscillations is excited in a stationary solid spherical body at some pure tone, and then the body is set into rotation with an arbitrarily time-varying angular velocity $\omega(t)$, the standing wave will rotate relative to the body according to a law of the form (4.2). It should be noted that the first-order differential equation (4.2) is the well-known kinematic equation of Poisson.

Let us compare equation (4.2) with the classical kinematic equation of Poisson for the rigid body, $\dot{N} = -GN$, where the orthogonal matrix N defines the orientation of the rigid body in absolute (inertial) space. From this, it follows that the angular velocity of the mechanical standing wave relative to the body is proportional to its angular velocity relative to space:

$$\Omega_0(t) = -\upsilon\omega(t), \qquad (4.4)$$

Alternatively, for the angular velocity of the mechanical standing wave relative to absolute space, we have the form:

$$\Omega(t) = (1 - \upsilon)\omega(t) \qquad (4.5)$$

Relation (4.4) represents a generalization of the well-known scalar expression of Lynch-Zhuravlev [3] to the spatial case. It is important to emphasize that for a complete spherical elastic shell, the Brian coefficient has a spatial character. It defines the proportionality between the angular velocity vectors, in contrast to the case of a thin ring, where only planar rotational angular velocities were considered.

The spatial mechanical standing wave excited on the surface of an elastic high-Q spherical solid body in the form of a three-dimensional ellipsoid of revolution precesses relative to the body with an angular velocity vector that is proportional to the angular velocity vector of the body itself relative to absolute and/or stationary (inertial) space.

V. CASE OF A COMPLETE SPHERICAL SHELL

The calculation of the Brian coefficient for a complete closed spherical shell is carried out under specific assumptions about the density distribution with respect to the radius *ρ(r)*. Let us consider the case of a thin spherical shell.

In projections onto the traditionally accepted axes in shell theory, *u, v,w*, which are directed along the meridian, parallels, and normal to the surface of the sphere, the eigenvectors $\boldsymbol{U}_1, \boldsymbol{U}_2$ have the following form [3]:

$$\boldsymbol{U}_1 = \begin{pmatrix} A_n\left(\frac{dP_n^k}{d\theta}\right)\cos k\,\varphi \\ -A_n k \sin^{-1}\theta\, P_n^k \sin k\,\varphi \\ P_n^k \cos k\,\varphi \end{pmatrix}; \boldsymbol{U}_2 = \begin{pmatrix} A_n(dP_n^k/d\theta)\sin k\,\varphi \\ A_n \sin^{-1} k\,\varphi \\ 0 \end{pmatrix} \qquad (5.1)$$

where P_n^k are associated Legendre functions, φ, θ are spherical coordinates, and A_n is a coefficient that depends on the number of waves along the meridian and on the frequency of oscillations ν.

For the coefficients A_n and the frequency of oscillations ν in the Lamb approximation, we have the following relations (where μ is Poisson's ratio):

$$A_n = \left[-2(1+\mu) + \frac{(1-\mu)\nu^2}{2}\right][(1+\mu)n(n+1)]^{-1}$$

The frequency equation of Lamb takes the following form:

$$(1-\mu)\nu^4 - 2\nu^2(1 + 3\mu + n(n+1) + \\ +4(1+\mu)(n^2 + n - 2) = 0 \qquad (5.2)$$

The vectors (5.1) represent two eigenvectors out of three that define a three-dimensional eigen subspace (for fixed n and k). The third can be obtained from those written down using a coordinate transformation; however, it is not necessary for the calculation of the coefficient k. The transition from the variables *u, v,w* to the variables adopted in this work, x_1, x_2, x_3, is expressed by the well-known coordinate transformation formula:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\varphi & \cos\theta\sin\varphi & \sin\theta \\ -\sin\varphi & \cos\varphi & 0 \\ -\sin\theta\cos\varphi & -\sin\theta\cos\varphi & \cos\theta \end{pmatrix}\begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

From which we obtain the analytical expression for the first vector:

$$\boldsymbol{h}_1 = \frac{1}{\|\boldsymbol{U}_1\|} \cdot \\ \cdot \begin{pmatrix} A_n\frac{dP_n^k}{d\theta}\cos k\,\varphi\cos\theta\cos\varphi - \frac{A_n k}{\sin\theta}P_n^k\cos k\,\varphi\cos\theta\sin\varphi + \\ +P_n^k\cos k\,\varphi\sin\theta \\ -A_n\frac{dP_n^k}{d\theta}\cos k\,\varphi\sin\varphi - A_n\frac{k}{\sin\theta}P_n^k\sin k\,\varphi\cos\varphi \\ -A_n\frac{dP_n^k}{d\theta}\cos k\,\varphi\sin\theta\cos\varphi + \frac{A_n k}{\sin\theta}P_n^k\sin k\,\varphi\sin\theta\sin\varphi + \\ +P_n^k\cos k\,\varphi\cos\theta \end{pmatrix}$$

A similar analytical expression can be obtained for the second eigenvector $\boldsymbol{h}_2$. After that, we calculate the Brian

coefficient υ for the thin closed (complete) spherical shell to use it as an inertial sensor for the three quasi-angular rotations of the base of the gyroscope relative to inertial space.

$$\upsilon = \left| \int_0^{2\pi} \int_0^{\pi} \boldsymbol{h}_1 \times \boldsymbol{h}_2 \sin\theta \, d\theta d\varphi \right| =$$

$$= \frac{2\pi k A_n}{\|\boldsymbol{U}_1\|^2} \left| \left\{ \int_0^{\pi} \frac{dP_n^k}{d\theta} P_n^k \sin\theta \, d\theta, 0, \frac{dP_n^k}{d\theta} P_n^k \cos\theta \, d\theta \right\} \right|$$

The integral that defines the first component of the written vector is equal to zero. For the second integral, we obtain an analytical expression:

$$\int_0^{\pi} \frac{dP_n^k}{d\theta} P_n^k \cos\theta \, d\theta = -\frac{(n+k)!}{(2n+1)(n-k)!}$$

Considering that the norm $\boldsymbol{U}_1$ has the form:

$$\|\boldsymbol{U}_1\|^2 = \frac{2\pi}{(2n+1)} [1 + n(n+1)A_n^2] \frac{(n+k)!}{(n-k)!} \quad (k \neq 0)$$

we ultimately find the analytical expression for the Brian coefficient:

$$\upsilon = kA_n^2 [1 + n(n+1)A_n^2]^{-1}$$

Let's consider, as an example, the case when $k = n = 2$ and $\mu = 0$. According to the formulas derived above (5.2), we have:

$$\nu^2 = 7 + \sqrt{33}, \qquad A_n = (7 + \sqrt{33})/12 - 1/3$$

For a thin complete spherical shell, the Brian coefficient [3], which is approximately one quarter (υ=*0,25*), turns out to be less than that obtained for a ring (or cylinder), which is equal to 2/5. It should be emphasized that for a complete spherical shell, the Brian coefficient has a spatial character. It defines the proportionality between the angular velocity vectors of the form (4.4), in contrast to the case of a ring, where planar rotational angular velocities were considered. The mechanical standing wave in the form of a three-dimensional ellipsoid precesses relative to the body itself with an angular velocity vector proportional to the angular velocity vector of the body relative to absolute (inertial) space.

VI. THREE-AXIS WAVE SOLID GYROSCOPE

In publications [2,3,4], it has been shown that a complete elastic spherical shell can serve as an inertial information sensor. We will demonstrate the use of wave phenomena in spatial elastic media for determining angular movements of solid bodies using a specific example of a three-axis wave solid gyroscope (WSG). Indeed, let's assume that a moving object rotates with an angular velocity $\omega_0(t)$ relative to inertial space. This angular velocity is projected onto axes associated with the moving object and can be determined by inertial means located on that object. Let there be an elastic spherical shell on the object, rotating with an angular velocity $\omega_1(t)$ relative to the object (the speed ω_1 is defined in the object's axes). In addition to the observed (known) speed $\omega_1(t)$, there is also observed the precession speed of the standing wave projected onto its own axes -$\omega_2(t)$. We denote the angular velocity of such a wave relative to the shell as $\omega_2^*(t)$. Based on the results obtained in the previous section, we will find the relationship between angular velocities $\omega_0(t)$, $\omega_1(t)$, $\omega_2(t)$. Recall that the coordinate transformation from axes associated with the object to axes associated with the spherical shell is given by an orthogonal matrix $N(t)$, which defines the position of the solid body in inertial space. The coordinate transformation from axes associated with the spherical shell to axes associated with the elastic wave is given by matrix $M(t)$. From formula (4.4), it follows that the precession speed of a spatial mechanical standing wave relative to the spherical shell in its own axes is proportional to both the speed of the shell and its speed relative to itself in terms of its own axes:

$$MN\omega_2^* = -\chi N(\omega_1 + \omega_2)$$

From this, in the axes associated with the moving object, we have the following expression:

$$\omega_2^* = -\chi N^T M^T N(\omega_1 + \omega_2)$$

The angular velocity of the standing wave relative to the moving object is given by:

$$\omega_2 = \omega_1 + \omega_2^* = \omega_1 - \chi N^T M^T N(\omega_1 + \omega_0)$$

Thus, we can find the angular velocity of the moving object

$$\omega_0 = \frac{1}{\chi} N^T \mathrm{M} N(\omega_1 - \omega_2) - \omega_1 \qquad (6.1)$$

To compute the angular velocity of the object ω_0 using formula (6.1), in addition to the assumed known speeds $\omega_1(t)$ and $\omega_2(t)$, it is necessary to know the matrices $N(t)$ and $M(t)$, which are related to angular velocities by the following well-known kinematic equations of Poinsot:

$$\dot{M} = -M[N(\omega_2 - \omega_1)]^{\nu}, \quad \dot{N} = -N\omega_1^{\nu} \qquad (6.2)$$

where ω_1^{ν} denotes the matrix representation of angular velocity. The sequence for solving the kinematic differential equations of the form (6.1) and (6.2) is as follows:

a) First, solve the following equation:

$$\dot{N} = -N\omega_1^{\nu}$$

b) Then, find the desired orthogonal matrix $N(t)$.

Substituting this solution back into the equation

$$\dot{M} = -M[N(\omega_2 - \omega_1)]^{\nu},$$

allows us to find matrix $M(t)$. After that, substituting into expression (6.1) allows us to find the desired speed ω_0. If the system outputs not the angular velocities $\omega_1(t)$ and $\omega_2(t)$, but rather matrices $N(t)$ and $S(t) = M(t)N(t)$, then first we find the matrix of angular velocities: $\omega_1^{\nu} = -N^T\dot{N}$ and after substituting into (6.2), we find the desired angular velocity ω_0. The problem of determining angular velocity ω_0 is most easily solved if we ensure that the spherical shell has zero angular velocity ($\omega_1 \equiv 0$). Then, the orthogonal matrix becomes матрица $N = E$ — the identity matrix. Consequently, from expression (6.1), we obtain:

$$\omega_0 = -\frac{1}{\chi} M\omega_2$$

The condition $\omega_1 \equiv 0$ can be ensured by fixing the sphere at points where it touches the vertices of an inscribed tetrahedron. Then, if we consider the form of elastic oscillations in which the sphere transforms into a three-axis ellipsoid, the distances

between these specified points change by quantities of second order with respect to the amplitude of the elastic oscillations, and in a linear formulation, fixing them does not constrain the oscillations of the shell. All of the above was discussed under the assumption that the shell possesses spherical symmetry and that there are no external disturbances. If deviations from these assumptions are small, then it is possible to compensate for the new effects that arise (for example, destruction of the standing wave and/or reduction of its amplitude) by introducing appropriate control mechanisms. Physically, such control can be achieved by applying electrical forces to the surface of the shell through specially positioned electrodes. The calculation of these forces, depending on the shape and number of electrodes and on the potential distribution across them, is detailed in monograph [3]. In this publication, we will limit ourselves to the remark that after calculating the necessary forces and projecting onto the chosen eigenform (in the sense of calculating the integral $\int_v f\, h_n(r)dm$, terms will appear on the right-hand sides of the differential equations governing the elastic oscillations of the spherical shell [2,3] that can be considered dependent on the phase state of the examined form. By knowing this state and using the principle of feedback, we can control the magnitudes of the potentials on the corresponding electrodes at our discretion. Thus, we reduce the task posed above to determining the evolutions of the elastic oscillation form that arise when perturbing the ideal scheme discussed earlier.

The information-rich 3D-WSG represents a three-dimensional inertial sensor based on the following fundamental principles:

• The resonator has perfect axial symmetry to ensure high performance during balancing.

• The control of the elastic standing wave is achieved through electrostatic forces generated by electrodes that are in close proximity to the surface of the spherical resonator.

• The resonator is designed in the shape of a complete (closed) sphere, where this construction allows each gram of the resonator material to participate most effectively in maintaining the energy of bending oscillations.

• A high-Q resonator made of amorphous fused quartz provides optimal isotropy (absence of crystallographic axes) and minimal energy dissipation (low internal friction).

As a result of the optimization performed, the calculated characteristics of the triaxial WSG have been obtained, which, surprisingly, do not depend on the dimensions of the resonator. They are related to the quality of the resonator, indicated by its Q-factor (or decay time), as well as to the quality of wave pattern control, meaning the characteristics of the control electronics.

The main drawback of the triaxial WSG is its high cost, primarily due to the extremely high precision required for manufacturing the spherical resonator and assembling complex components and assemblies that must precisely match the geometry of the spherical shell of the resonator to form a uniform gap between these two precision parts of the gyroscope.

VII.PHASE MULTIPLICITY OF ELASTIC SPATIAL STANDING WAVES

All further analysis is conducted within the framework of the first approximation of the averaging method. Therefore, we will consider the following assumptions to be valid regarding the three second-order differential equations of elastic vibrations of a spherical shell obtained in publications [3,4].

The angular velocity of the body relative to space, $\boldsymbol{\omega}(t)$, is considered small compared to the frequency of the natural elastic vibrations of the spherical shell. Simultaneously, we also consider the coupling coefficients that appear in the terms $L(L_1, L_2, L_3)$, which represent linear functions of generalized coordinates corresponding to other natural subspaces, to be small. These terms arise as a result of calculating integrals of the type

$$\int_v h_k \times h_i dm = 0$$

where h_k , h_i belong to different natural subspaces, and they are indeed small compared to similar integrals where h_k and h_i are taken from the same natural subspace. Furthermore, we will add to the system of differential equations of the elastic vibrations of the sphere terms defined by various perturbations and control influences, which we will also consider to be small.

The presence of these terms characterizes the fact that the differential systems of elastic vibrations of the complete spherical shell for different subspaces are not independent of each other. In deriving these equations, it was assumed that the mass forces f are orthogonal to all eigenfunctions.

$$\int_v f\, h_n(r)dm = 0$$

This condition implies that the forces f consist only of a constant component given by

$$\int_v f \times h_n(r)dm \neq 0,$$

which ensures the rotation of the body itself with an angular velocity $\omega(t)$. This restriction will be lifted later. Considering the above, the infinite system of ordinary second-order differential equations with respect to generalized coordinates obtained in the monographs [2, 3] can be expressed as:

$$\begin{aligned} \ddot{u} + u &= \varepsilon U \\ \ddot{v} + v &= \varepsilon V \\ \ddot{w} + w &= \varepsilon W \end{aligned} \tag{7.1}$$

The general solution of the system of differential equations (7.1) in the absence of corresponding control ($(\varepsilon = 0)$ has the following analytical form:

$$\begin{aligned} u &= x_1 cost + x_4 sint \\ v &= x_2 cost + x_5 sint \\ w &= x_3 cost + x_6 sint \end{aligned} \tag{7.2}$$

The velocities along the trajectories (7.2) — the changes in the original coordinates u,v,w — can be obtained by differentiating the expressions (7.2), yielding:

$$\begin{aligned} \dot{u} &= -x_1 sint + x_4 cost \\ \dot{v} &= -x_2 sint + x_5 cost \\ \dot{w} &= -x_3 sint + x_6 cost \end{aligned} \tag{7.3}$$

Next, we will consider a change of variables

$(u, v, w, \dot{u}, \dot{v}, \dot{w}) \rightarrow (x_1, x_2, x_3, x_4, x_5, x_6)$ in the system of equations of the form (7.1), leading it (after averaging over time *t* to the following general form of the equation:

$$\dot{x}=\varepsilon X(x) \tag{7.4}$$

where x is a new phase variable, and the system (7.4) defines the evolution in phase space of those forms of oscillations observed in the system (7.1) under the condition of no control $\varepsilon = 0$. When the condition $\varepsilon = 0$ is satisfied, any trajectory in the system (7.1) corresponds to a stationary point in the system (7.4). The trajectories themselves are defined by parametric equations (7.2) and represent a planar curve that is generally an ellipse. The equation of the plane in which this ellipse lies can be easily derived by noting that to eliminate $cost$ and $sint$ from the system (7.2), it is necessary to satisfy conditions of linear dependence:

$$det\begin{Vmatrix} u & x_1 & x_4 \\ v & x_2 & x_5 \\ w & x_3 & x_6 \end{Vmatrix} = 0 \tag{7.5}$$

Based on (7.5), we obtain the desired equation:

$$u(x_2x_6 - x_3x_5) + v(x_3x_4 - x_1x_6) + \\ +w(x_1x_5 - x_2x_4) = 0 \tag{7.6}$$

Standing waves on the surface of a spherical shell are characterized by all points passing through the equilibrium position simultaneously during oscillations, which corresponds to the simultaneous vanishing of functions of the form (7.2). This, in turn, means that the elliptical trajectory in the parameter space (u, v, w) transforms into a segment of a straight line.

In this case, the plane defined by (7.6) is not well-defined, leading to the following conditions on the phase vector:

$$k_1 = (x_3x_4 - x_1x_6) = 0,\ k_2 = (x_1x_5 - x_2x_4) = 0, \\ k_3 = (x_2x_6 - x_3x_5) = 0 \tag{7.7}$$

Thus, the set of elastic standing waves of an arbitrarily chosen form of oscillation in the spherical shell, which is in one-to-one correspondence with the set of linear forms of oscillation in system (7.1), is represented in the phase space of system (7.4) as a quadratic variety (7.7), which represents the intersection of three spatial conical cylinders.

VIII. EVOLUTIONARY BASIS ON THE PHASE VARIETY OF STANDING WAVES

If in a system of the form (7.1) we choose initial conditions such that, in the absence of control $\varepsilon = 0$, a linear form of oscillation is realized (these conditions correspond to the position of the initial point in the system (7.4)) on the variety (7.7), then when the condition $\varepsilon \neq 0$ is met, this linear form will undergo one of the following evolutions: shape destruction, shape precession, change in oscillation frequency, and change in oscillation amplitude. Each of these types of evolution of the linear form corresponds to specific directions in phase space, which we pay special attention to in this section of the article.

The direction of the fastest destruction of a standing wave is determined by the normal to the variety (7.7):

$$\boldsymbol{e_1} = \frac{dk_1}{dx} = \{-x_6, 0, x_4, x_3, 0, -x_1\} \\ \boldsymbol{e_2} = \frac{dk_2}{dx} = \{x_5, -x_4, 0, -x_2, x_1, 0\} \\ \boldsymbol{e_3} = \frac{dk_3}{dx} = \{0, x_6, -x_5, 0, -x_3, x_2\} \tag{8.1}$$

If the right-hand side of system (7.4) belongs to the linear span of three vectors of the form (8.1), then the evolution of the standing wave consists of its destruction.

To determine the directions in the phase space of system (7.4) along which the standing wave undergoes precession, we introduce a spherical coordinate system in the space of the original variables (u, v, w):

$$u = r\cos\varphi\cos\psi \\ v = r\cos\varphi\sin\psi \\ w = r\sin\varphi \tag{8.2}$$

Let us consider the variation in position corresponding to a change in angle ψ:

$$\delta u = -r\cos\varphi\sin\psi\delta\psi = -v\delta\psi \\ \delta v = r\cos\varphi\cos\psi\delta\psi \\ \delta w = 0 \tag{8.3}$$

Then, on the other hand, the variation in position can be computed based on formulas (7.2):

$$\delta u = \delta x_1 cost + \delta x_4 sint \\ \delta v = \delta x_2 cost + \delta x_5 sint \\ \delta w = \delta x_3 cost + \delta x_6 sint \tag{8.4}$$

By equating these relationships from (8.4) with the analytical expressions from (7.7), we obtain:

$$-v\delta\psi\,\delta u = \delta x_1 cost + \delta x_4 sint = -(x_2 cost + x_5 sint)\delta\psi \\ u\delta\psi = \delta x_2 cost + \delta x_5 sint = (x_1 cost + x_4 sint)\delta\psi \\ 0 = \delta x_3 cost + \delta x_6 sint$$

Taking $\delta\psi \to 0$ and equating separately in the obtained equalities coefficients at $sint$ and $cost$, we find:

$$\frac{dx_1}{d\psi} = -x_2;\ \frac{dx_2}{d\psi} = x_1;\ \frac{dx_3}{d\psi} = 0;\ \frac{dx_6}{d\psi} = 0;\ \frac{dx_5}{d\psi} = x_4;\ \frac{dx_4}{d\psi} = -x_5$$

In other words, the desired direction of precession is defined by the following vector:

$$\boldsymbol{e_4} = \{-x_2, x_1, 0, -x_5, x_4, 0\}$$

For the two other directions of precession (noting that there are three directions of elementary rotations for an ellipse as for a rigid body), we introduce a spherical coordinate system using cyclic permutations of axes. To find direction $\boldsymbol{e_5}$:

$$u = r\cos\varphi\sin\psi \\ v = r\sin\varphi \\ w = r\cos\varphi\cos\psi$$

And correspondingly, to find direction $\boldsymbol{e_6}$:

$$u = r\sin\varphi \\ v = r\cos\varphi\cos\psi \\ w = r\cos\varphi\sin\psi$$

By performing similar calculations sequentially both times, we find the desired unit vectors:

$$\boldsymbol{e_5} = \{x_3, 0, -x_1, x_6, 0, -x_4\} \\ \boldsymbol{e_6} = \{0, -x_3, x_2, 0, -x_6, x_5\}$$

To construct the direction of frequency variation, we will equate the variations from (8.4) with those obtained by varying over time in the general analytical expression of system (7.2):

$$\delta u = -x_1 sint\delta t + x_4 cost\delta t$$
$$\delta v = -x_2 sint\delta t + x_5 cost\delta t$$
$$\delta w = -x_3 sint\delta t + x_6 cost\delta t$$

As a result, we obtain the following desired unit vector:

$$\boldsymbol{e_7} = \{x_4, x_5, x_6, -x_1, -x_2, -x_3\}$$

To find the direction of amplitude variation, we introduce the equation of a six-dimensional sphere of unit radius:

$$S = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 - 1) = 0$$

Then the direction of amplitude change is determined by the gradient of S:

$$\boldsymbol{e_8} = \frac{dS}{dx} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

Now let us summarize all the results obtained above. The directions in phase space that define the main evolutions of the linear form of oscillations in configuration space are given by the following vectors:

$$\begin{aligned} \boldsymbol{e_1} &= \{-x_6, 0, x_4, x_3, 0, -x_1\} \\ \boldsymbol{e_2} &= \{x_5, -x_4, 0, -x_2, x_1, 0\} \\ \boldsymbol{e_3} &= \{0, x_6, -x_5, 0, -x_3, x_2\} \end{aligned} \quad (8.5)$$

These vectors define the destruction of the linear form of elastic oscillations:

$$\begin{aligned} \boldsymbol{e_4} &= \{-x_2, x_1, 0, -x_5, x_4, 0\} \\ \boldsymbol{e_5} &= \{x_3, 0, -x_1, x_6, 0, -x_4\} \\ \boldsymbol{e_6} &= \{0, -x_3, x_2, 0, -x_6, x_5\} \end{aligned} \quad (8.6)$$

These vectors define the precession of the linear form of elastic oscillations.

The change in oscillation frequency is represented by the unit vector:

$$\boldsymbol{e_7} = \{x_4, x_5, x_6, -x_1, -x_2, -x_3\}$$

The change in oscillation amplitude is represented by the following unit vector:

$$\boldsymbol{e_8} = \{x_1, x_2, x_3, x_4, x_5, x_6\}.$$

IX. ALGEBRAIC PROPERTIES OF THE EVOLUTIONARY BASIS

Let us compute the determinant of the Gram matrix for the system of vectors (8.5) and (8.6), assuming that the Gram matrix has the following form:

$$\widetilde{\boldsymbol{\Gamma}} = \begin{Vmatrix} (\boldsymbol{e}_1\boldsymbol{e}_1) & \dots & (\boldsymbol{e}_1\boldsymbol{e}_8) \\ \vdots & & \vdots \\ (\boldsymbol{e}_8\boldsymbol{e}_1) & \dots & (\boldsymbol{e}_8\boldsymbol{e}_8) \end{Vmatrix} \quad (9.1)$$

where $(\boldsymbol{e}_i\boldsymbol{e}_k)$ denotes the scalar product of the vectors.

The computation of scalar products shows that the Gram matrix exhibits a block structure, in which the Gram matrix of the vectors corresponding to the destruction bundle of the standing wave coincides with the Gram matrix of the vectors corresponding to the precession bundle.

On the considered manifold of standing waves, we have

$$K_1 = K_2 = K_3 = 0$$

In addition to its block structure, it has a block-diagonal form with a determinant equal to the following expression:

$$det\widetilde{\boldsymbol{\Gamma}} = \|\mathrm{x}\|^4 \cdot det^2\boldsymbol{\Gamma}$$

For the matrix $\boldsymbol{\Gamma}$, we have $det\,\boldsymbol{\Gamma}$ and the following analytical expression:

$$\begin{Vmatrix} x_1^2+x_3^2+x_4^2+x_6^2 & -(x_2x_3+x_5x_6) & -(x_1x_2+x_4x_5) \\ & x_1^2+x_2^2+x_4^2+x_5^2 & -(x_1x_3+x_4x_6) \\ & & x_2^2+x_3^2+x_5^2+x_6^2 \end{Vmatrix} = (K_1^2 + K_2^2 + K_3^2) \cdot \|\mathrm{x}\|^2 \quad (9.2)$$

Thus, we ultimately have:

$$det\widetilde{\boldsymbol{\Gamma}} = (K_1^2 + K_2^2 + K_3^2) \cdot \|\mathrm{x}\|^8 \quad (9.3)$$

On the manifold of mechanical standing waves, we have $det\boldsymbol{\Gamma} = 0$. This means that the three destruction vectors $\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}$ are linearly dependent. The same can be said for the three precession vectors $\boldsymbol{e_4}, \boldsymbol{e_5}, \boldsymbol{e_6}$. In this case, the first bundle is orthogonal to the second, and each of them is orthogonal to the vectors $\boldsymbol{e_7}$ and $\boldsymbol{e_8}$, which are themselves orthogonal to each other. Thus, from each bundle of the aforementioned vectors, it is possible to select two mutually orthogonal vectors, allowing us to construct an orthogonal basis on the manifold K of standing waves consisting of six vectors.

Now let us examine the structure of the algebra of operators generated by the evolutionary basis. The eight vector fields of the forms (8.5) and (8.6), along with the vectors $\boldsymbol{e_7}$ and $\boldsymbol{e_8}$, generate eight one-parameter subgroups of a non-commutative eight-parameter group. The basis of the algebra of this group is defined by the following analytical operators:

$$\begin{aligned} U_1 &= -x_6\frac{\partial}{\partial x_1} + x_4\frac{\partial}{\partial x_3} + x_3\frac{\partial}{\partial x_4} - x_1\frac{\partial}{\partial x_6} \\ U_2 &= x_5\frac{\partial}{\partial x_1} - x_4\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_4} + x_1\frac{\partial}{\partial x_5} \\ U_3 &= x_6\frac{\partial}{\partial x_2} - x_5\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_4} + x_2\frac{\partial}{\partial x_6} \\ U_4 &= -x_2\frac{\partial}{\partial x_1} + x_1\frac{\partial}{\partial x_2} - x_5\frac{\partial}{\partial x_4} + x_4\frac{\partial}{\partial x_5} \\ U_5 &= x_3\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_3} + x_6\frac{\partial}{\partial x_4} - x_4\frac{\partial}{\partial x_6} \\ U_6 &= -x_3\frac{\partial}{\partial x_2} + x_2\frac{\partial}{\partial x_3} - x_6\frac{\partial}{\partial x_5} + x_5\frac{\partial}{\partial x_6} \\ U_7 &= x_4\frac{\partial}{\partial x_1} + x_5\frac{\partial}{\partial x_2} + x_6\frac{\partial}{\partial x_3} - x_1\frac{\partial}{\partial x_4} - x_2\frac{\partial}{\partial x_5} - x_3\frac{\partial}{\partial x_6} \\ U_8 &= x_1\frac{\partial}{\partial x_1} + x_2\frac{\partial}{\partial x_2} + x_3\frac{\partial}{\partial x_3} + x_4\frac{\partial}{\partial x_4} + x_5\frac{\partial}{\partial x_5} + x_6\frac{\partial}{\partial x_6} \end{aligned} \quad (9.4)$$

The structural constants of the considered group are obtained by calculating the following commutators:

$$\begin{aligned} &[U_1, U_2] = U_6, \quad [U_1, U_3] = -U_4, \quad [U_1, U_4] = -U_3, \\ &[U_1, U_5] = 0, \quad [U_1, U_6] = U_2, \quad [U_1, U_7] = 0, \\ &[U_1, U_8] = 0, \quad [U_2, U_3] = -U_5, \quad [U_2, U_4] = 0, \\ &[U_2, U_5] = 0, \quad [U_2, U_6] = -U_1, \quad [U_2, U_7] = 0, \\ &[U_2, U_8] = 0, \quad [U_3, U_4] = U_1, \quad [U_3, U_5] = -U_2, \\ &[U_3, U_6] = 0, \quad [U_3, U_7] = 0, \quad [U_3, U_8] = 0, \end{aligned} \quad (9.5)$$

$$[U_4, U_5] = U_6, \quad [U_4, U_6] = -U_5, \quad [U_4, U_7] = 0,$$
$$[U_4, U_8] = 0, \quad [U_5, U_6] = U_4, \quad [U_5, U_7] = 0,$$
$$[U_5, U_8] = 0, \quad [U_6, U_7] = 0,$$
$$[U_6, U_8] = 0, \quad [U_7, U_8] = 0.$$

From the results presented above, it follows that the frequency variation operators and amplitude variation operators commute with all other operators of the evolution group, from which a six-dimensional subgroup is distinguished, generated by the wave destruction bundle and the precession bundle.

We should also note the following important properties. The commutator of two operators from the precession bundle is an operator from the same bundle, which is not surprising since the precession bundle generates a three-dimensional rotation group $SO(3)$. Additionally, the commutator of two operators, one from the destruction bundle and the other from the precession bundle, results in an operator from the destruction bundle. Thus, unlike the one-dimensional case (the classical half-spherical WSG), where the dimension of the evolution algebra corresponded to the dimension of phase space (four) and was commutative, in the spatial case (hollow spherical shell), the dimension of the evolution algebra (eight) does not match the dimension of phase space (six) and is also non-commutative [2].

X. STABILIZATION EQUATION FOR SPATIAL STANDING WAVE

To ensure the asymptotic stability of the excited spatial mechanical standing wave in the hollow spherical shell when ε=0, we introduce feedback into the right-hand side of system (7.4) in the following manner:

$$\dot{x} = -\sum_{i,j}(a_{i,j})e_i k_j - e_8 S \quad (i,j = 1,2,3) \qquad (10.1)$$

The stationary state $\dot{x} \equiv 0$ for the system is achieved on the manifold $k_j = 0,\ S = 0,$, which defines a standing wave with an equal unit amplitude of oscillation. The equality of oscillation amplitudes being one is related to the choice of measurement scale for lengths and their normalization to a dimensionless form. The 3×3 matrix of coefficients $a_{i,j}$ should be chosen such that the state $k_j = 0, S = 0$ is asymptotically stable. If this condition is met, the stabilization problem for the standing wave will be fully resolved.

The formation of the right-hand sides in system (10.1) in terms of the original forces acting in system (7.1) requires applying spherical dissipative forces, or hyperbolic dissipative forces, and the amplitude of oscillations should be taken equal to the following analytical expression:

$$S = \tfrac{1}{2}(\|x\|^2 - 1) \qquad (10.2)$$

According to the conducted theoretical studies, it is most effective to use dissipative spherical forces, as their effectiveness does not depend on the location of the point on the full sphere but only on its radius. This means that in the considered system (7.1), the term $e_8 S$ will be represented as follows:

$$\begin{aligned} \ddot{u} + u &= -d\dot{u}S \\ \ddot{v} + v &= -d\dot{v}S \\ \ddot{w} + w &= -d\dot{w}S \end{aligned} \qquad (10.3)$$

These equations are analogous to the classical Van der Pol equations that describe self-oscillations. To form the term equal to

$$\sum_{i,j}(a_{i,j})\, e_i k_j$$

it is necessary to use either positional hyperbolic or circular forces. Their amplitudes should be taken proportional to k_j with proportionality coefficients determined by the condition of asymptotic stability. It is worth noting that it is more advantageous to use circular forces, as they do not lead to a change in the frequency of resonant oscillations.

It should be specifically emphasized that the principle of stabilizing spatial standing waves discussed in this study fundamentally differs from the one currently accepted in modern developments of WSG. The difference lies in the fact that, in our considered principle, feedbacks are formed based on the observation of the phase state within the framework of autonomous right-hand sides. In the existing approach, control based on amplitude in semi-spherical resonators of WSG is non-autonomous, as it utilizes parametric pumping.

Now let us consider how to choose the coefficients $a_{i,j}$. To do this, we will compute the total time derivatives of the system (10.1) with respect to k_l and S.

$$\frac{dK_l}{dt} = \frac{dK_l}{dx}\dot{x} = e_l\left(-\sum_{i,j}(a_{i,j})\, e_i k_j - e_8 S\right) \qquad (10.4)$$

It should be noted that the products $e_l \cdot e_i$ represent the coefficients of the Gram matrix. In this case, we choose the matrix

$$\sum_{i,j}(a_{i,j})\, e_i e_l = \sum_{i}(a_{i,j})\, \gamma_{i,l}$$

based on the following condition:

$$\sum_i (a_{i,j})\gamma_{i,l} = \delta_{j,l} \cdot det\Gamma \qquad (10.5)$$

where $\delta_{j,l}$ is the Kronecker delta: $\delta_{j,l} = 1$ when $j = l$, and $\delta_{j,l} = 0$ when $j \neq l$. Since it follows from the condition above that the matrix

$$\left\{\tfrac{1}{\det\boldsymbol{\Gamma}}\ a_{i,j}\right\}$$

is the inverse of the Gram matrix, and considering the rule for calculating the inverse of a matrix, we conclude that the matrix

$$\{\, a_{i,j}\}$$

represents the matrix of algebraic complements of the Gram matrix:

$$\{\, a_{i,j}\} = \{\boldsymbol{\Gamma}_{i,j}\}$$

If condition (10.5)

$$\sum_i (a_{i,j})\gamma_{i,l} = \delta_{j,l} \cdot det\boldsymbol{\Gamma}$$

holds, then the first term in (10.4) takes the form:

$$e_l\left(-\sum_{i,j} a_{i,j}\, e_i k_j\right) = -K_l \det\boldsymbol{\Gamma}$$

It can be shown that the product $e_l e_8$ in the matrix $\mathbf{\Gamma}$ takes the following form:

$$e_l e_8 = 2K_l \tag{10.6}$$

As a result, equation (10.4) takes the form:

$$\frac{dK_l}{dt} = -K_l \det \mathbf{\Gamma} - 2K_l S \tag{10.7}$$

where $l = 1,2,3$.

Now let's compute the derivative of the variable S:

$$\frac{dS}{dt} = \frac{dS}{dx}\dot{x} = e_8\left(-\sum_{i,j} a_{i,j}\, e_i k_j - e_8 S\right) \tag{10.8}$$

Due to the equality $e_l e_8 = 2K_l$, the first term on the right-hand side of expression (10.8) takes the following form:

$$e_8\left(-\sum_{i,j} a_{i,j}\, e_i k_j\right) = -\sum_{i,j} a_{i,j}\, e_i k_j$$

In this case, for the second term, $(e_8 e_{8)} = \|x\|^2 = (2S + 1)$. Therefore, we can express the sought equation for the derivative with respect to S as follows:

$$\frac{dS}{dt} = -\sum_{i,j} \Gamma_{i,j}\, k_i k_j - S(2S + 1) \tag{10.9}$$

Thus, the stabilization equation for a spatial standing wave of given amplitude, considering (9.2), takes on the following analytical form:

$$\frac{dK_l}{dt} = -K_l\,(K_1^2 + K_2^2 + K_3^2)(2S + 1) - 2K_l S \tag{10.10}$$

$$\frac{dS}{dt} = -\sum_{i,j} \Gamma_{i,j}\, k_i k_j - S(2S + 1) \tag{10.11}$$

We introduce the following notations:

$$y=(K_1^2 + K_2^2 + K_3^2)$$

taking this into account, we can rewrite system (10.11) as follows:

$$\frac{dy}{dt} = -2y^2\,(2S + 1) - 4yS$$

$$\frac{dS}{dt} = -\sum_{i,j} \Gamma_{i,j}\, k_i k_j - S(2S + 1)$$

According to the first approximation of the normal form method, we retain only resonant terms in the resulting system:

$$\frac{dy}{dt} = -2\mathrm{y}^2;\ \frac{dS}{dt} = -S$$

From which we derive the following analytical expressions:

$$y = y_0/(1 + 2\,\mathrm{y}_0 t)$$
$$S = S_0 e^{-t}$$

This signifies the asymptotic stability of the manifold y=0, $S = 0$, albeit not exponential stability.

XI. CONCLUSION

1. From the theoretical studies presented in this work, it follows that a three-axis inertial sensor (3D-WSG) is technically feasible, as there exists a fundamental possibility for the excitation and stabilization of a spatial mechanical standing wave while maintaining the effective functioning of the information channel.

2. For numerous practical applications, we have, in one new gyroscopic device, a very compact three-dimensional inertial sensor for measuring angular displacement and/or angular velocity. This device allows for simultaneous measurement of increments of three quasi-coordinates—angles of pseudo-orientation—which are incorporated into well-known kinematic differential equations (KDE). By integrating these equations, we solve the classical Darboux problem of determining the spatial orientation of a moving object in three-dimensional space using specified or known angular velocities.

REFERENCES

[1] Bryan G. H. On the beats in the vibrations of a revolving cylinder or bell // Proc. Camb. Phil. Soc., Vol VII, Nov. 24, pp. 101–111, 1890.

[2] Zhuralyev V.F., Klimov D.M. "Applied Methods in the Theory of Oscillations." ISBN: 5-02-006627-3, - Moscow: Nauka, 1988. – 325 p.

[3] Klimov D.M., Zhuravlev V.F., Zbanov Yu.K. Quartz Semi-Spherical Resonator (Wave Solid-State Gyroscope). Moscow: Kim L.A., 2017. 194 p.

[4] Zhuravlev V.F., Klimov D.M. The Spatial Effect of Inertia of Elastic Waves on a Sphere // Izvestiya RAN. MTT. 2021, No. 3, pp. 3-6.

[5] V.F. Zhuravlev. On the Formation of Feedbacks in the Spatial Oscillator of Van der Pol // Applied Mathematics and Mechanics, 2020, Vol. 84, No. 2, pp. 151–157.

[6] V. Zhuravlev, S. Perelyaev. The Generalized Foucault Pendulum as a 3D Integrating Gyroscope Using the Three-Dimensional Precession of Standing Waves in a Rotating Spherically Symmetric Elastic Solid. DGON Inertial Sensors and Systems Symposium. IEEE, Publication Year: 2019, pp. 1–12.

New Quaternion Algorithms of SINS in Euler Parameters for Determining the Inertial Orientation of Moving Objects: using the Structures of Runge-Kutta Methods

Yu.N. Chelnokov
Institute of Precision Mechanics and Control Problems of the Russian Academy of Sciences
Saratov, Russia
ChelnokovYuN@gmail.com

Annotation — **The report presents new quaternion algorithms for SINS of the third and fourth orders of accuracy, designed to determine the orientation of moving objects in an inertial coordinate system. The algorithms implement numerical integration on the on-board computer in real time of the classical quaternion equation of object orientation in Euler (Rodrigues-Hamilton) parameters based on information from the sensitive elements of the SINS. The input information for the first class of algorithms is instantaneous information about the rotational motion of the object (projections of the absolute angular velocity vector of the object onto the coordinate axes associated with it), for the second class of algorithms – integral information (increments of integrals from projections of the absolute angular velocity vector of the object onto the coordinate axes associated with it (quasi-coordinates)). Output information – Euler parameters (inertial orientation quaternion) and aircraft-type angles characterizing the orientation of the object in the inertial coordinate system. The orientation algorithms are constructed according to a recurrent scheme using the structures of classical Runge–Kutta methods of the third and fourth orders of accuracy. The results of mathematical modeling of the proposed algorithms are presented, demonstrating their high accuracy characteristics.**

Keywords: Quaternion algorithms, inertial orientation, moving object, Euler (Rodrigues–Hamilton) parameters, instantaneous and integral information about the rotational motion of an object, structures of Runge–Kutta methods.

I. INTRODUCTION

The report presents new quaternion algorithms for SINS of the third and fourth orders of accuracy, designed to determine the orientation of moving objects in an inertial coordinate system, implementing numerical integration on an on-board computer in real time of the classical quaternion differential equation of object orientation in four-dimensional Euler (Rodrigues-Hamilton) parameters based on information from the sensitive elements of the SINS (gyroscopes).

Classical Runge–Kutta methods of numerical integration of quaternion differential equation of orientation and the first class of algorithms for integrating this equation proposed by us in the report use the values of the coefficients of these equations (the values of the projections of the absolute angular velocity vector of the object onto the coordinate axes associated with it) as input information. Therefore, in order to apply the structures of Runge–Kutta methods of numerical integration of quaternion differential equation of orientation in the case of using integral primary information about the rotational motion of the object (increments of integrals from the projections of the absolute angular velocity vector of the object onto the coordinate axes associated with it (quasi-coordinates)), we proposed various algorithms for forming projections of the absolute angular velocity vector of the object onto the coordinate axes associated with it based on measured quasi-coordinates, using either the Simpson and trapezoid methods, or second-degree polynomials, or Taylor series. These algorithms are included in the general quaternion structures of various versions of the second class of orientation algorithms.

II. ORIENTATION ALGORITHMS IN EULER (RODRIGUE–HAMILTON) PARAMETERS IMPLEMENTING THE STRUCTURES OF RUNGE-KUTTA METHODS OF 4 AND 3 ORDERS OF ACCURACY USING INSTANTANEOUS INFORMATION ON THE ANGULAR MOTION OF AN OBJECT

The quaternion differential equation of object orientation in four-dimensional Euler (Rodrigues-Hamilton) parameters has the well-known form (2.1):

$$\frac{d\boldsymbol{\lambda}}{dt} = \frac{1}{2}\boldsymbol{\lambda} \circ \boldsymbol{\omega}_x(t), \qquad (2.1)$$

$$\boldsymbol{\lambda} = \lambda_0 + \boldsymbol{\lambda}_v = \lambda_0 + \lambda_1\mathbf{i} + \lambda_2\mathbf{j} + \lambda_3\mathbf{k}, \quad \boldsymbol{\omega}_x = \omega_1\mathbf{i} + \omega_2\mathbf{j} + \omega_3\mathbf{k},$$

where $\boldsymbol{\lambda}$ is the orientation quaternion of the object in the inertial coordinate system; $\boldsymbol{\omega}$ is the vector of the absolute angular velocity of the object, $\boldsymbol{\omega}_x$ is its mapping on the axes of the coordinate system X associated with the object; λ_j (j = 0, 1, 2, 3) are the Rodrigues–Hamilton (Euler) parameters, ω_k (k = 1, 2, 3) are the projections of the vector $\boldsymbol{\omega}$ on the axes of the associated coordinate system; **i**, **j**, **k** are the vector imaginary Hamilton units.

2.1. The obtained orientation algorithms in increments of the Euler (Rodrigues–Hamilton) parameters $\lambda_j^* = \Delta\lambda_j$,

implementing the structures of the Runge-Kutta methods of the 4th order of accuracy using instantaneous information about the absolute angular motion of the object, have the following form:

1) Algorithm for calculating the orientation quaternion at the current integration step:

$$\boldsymbol{\lambda}_n = \boldsymbol{\lambda}_{n-1} \circ \boldsymbol{\lambda}^*. \tag{2.2}$$

2) Algorithm for calculating the increment $\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda}$ of the orientation quaternion $\boldsymbol{\lambda}$ at the integration step (the upper line is the conjugation symbol, the central circle is the quaternion multiplication symbol):

$$\begin{aligned}\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda} = \overline{\boldsymbol{\lambda}}(t_{n-1}) \circ \boldsymbol{\lambda}(t_n) = \overline{\boldsymbol{\lambda}}_{n-1} \circ \boldsymbol{\lambda}_n = \\ = 1 + (1/12)(\mathbf{K}_1 + 2\mathbf{K}_2 + 2\mathbf{K}_3 + \mathbf{K}_4).\end{aligned} \tag{2.3}$$

3) Algorithms for calculating quaternion coefficients $\mathbf{K}_i$ included in the quaternion modified formula (2.3) of the Runge-Kutta method of the 4th order of accuracy ($h = t_n - t_{n-1}$ – integration step):

$$\begin{aligned}\mathbf{K}_1 &= h\boldsymbol{\omega}(t_{n-1}),\\ \mathbf{K}_2 &= h(1 + 0.25\mathbf{K}_1) \circ \boldsymbol{\omega}(t_{n-1} + 0.5h),\\ \mathbf{K}_3 &= h(1 + 0.25\mathbf{K}_2) \circ \boldsymbol{\omega}(t_{n-1} + 0.5h),\\ \mathbf{K}_4 &= h(1 + 0.5\mathbf{K}_3) \circ \boldsymbol{\omega}(t_{n-1} + h).\end{aligned} \tag{2.4}$$

2.2. The obtained orientation algorithms in increments of the Euler (Rodrigues–Hamilton) parameters $\lambda_j^* = \Delta\lambda_j$, implementing the structures of the Runge-Kutta methods of the 3rd order of accuracy using instantaneous information about the absolute angular motion of the object, have the following form:

1) The algorithm for calculating the orientation quaternion $\boldsymbol{\lambda}$ at the current integration step has the form (2.2).

2) Algorithm for calculating the increment $\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda}$ of the orientation quaternion at the integration step:

$$\begin{aligned}\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda} = \overline{\boldsymbol{\lambda}}(t_{n-1}) \circ \boldsymbol{\lambda}(t_n) = \overline{\boldsymbol{\lambda}}_{n-1} \circ \boldsymbol{\lambda}_n = \\ = 1 + (1/12)(\mathbf{K}_1 + 4\mathbf{K}_2 + \mathbf{K}_3).\end{aligned} \tag{2.5}$$

3) Algorithms for calculating quaternion coefficients $\mathbf{K}_i$ included in the quaternion modified formula (2.5) of the Runge-Kutta method of the 3rd order of accuracy ($h = t_n - t_{n-1}$ – integration step):

$$\begin{aligned}\mathbf{K}_1 &= h\boldsymbol{\omega}(t_{n-1}),\\ \mathbf{K}_2 &= h(1 + 0.25\mathbf{K}_1) \circ \boldsymbol{\omega}(t_{n-1} + 0.5h),\\ \mathbf{K}_3 &= h(1 - 0.5\mathbf{K}_1 + \mathbf{K}_2) \circ \boldsymbol{\omega}(t_{n-1} + h).\end{aligned} \tag{2.6}$$

III. ORIENTATION ALGORITHMS IN EULER (RODRIGUE–HAMILTON) PARAMETERS IMPLEMENTING THE STRUCTURES OF RUNGE-KUTTA METHODS OF 4 AND 3 ORDERS OF ACCURACY USING INTEGRAL INFORMATION ON THE ANGULAR MOTION OF AN

The first version of the orientation algorithms in increments of the Euler (Rodrigues–Hamilton) parameters $\lambda_j^* = \Delta\lambda_j$, implementing the structures of the Runge-Kutta methods of the 4th order of accuracy using integral information about the absolute angular motion of the object have the following form (it is assumed that the initial value of the absolute angular velocity vector of the object $\boldsymbol{\omega}(t_0)$ at the initial moment of time t_0 is known (given):

1) Algorithm for calculating the coefficients $\mathbf{K}_i$ included in the formulas of the Runge-Kutta method of the 4th order of accuracy, at the first step of integration:

$$\mathbf{K}_1 = \boldsymbol{\gamma}_1 = h\boldsymbol{\omega}(t_0), \tag{3.1}$$

$$\boldsymbol{\gamma}^* = 5\boldsymbol{\gamma}' + \boldsymbol{\gamma}'' - \boldsymbol{\gamma}_1, \tag{3.2}$$

where $\boldsymbol{\gamma}'$ and $\boldsymbol{\gamma}''$ are the input integral information of the algorithm, measured by gyroscopes:

$$\boldsymbol{\gamma}' = \int_{t_0}^{t_0+h/2} \boldsymbol{\omega}(t)dt, \quad \boldsymbol{\gamma}'' = \int_{t_0+h/2}^{t_0+h} \boldsymbol{\omega}(t)dt, \quad t_0 + h = t_1;$$

$$\mathbf{K}_2 = 0.5(1 + 0.25\mathbf{K}_1) \circ \boldsymbol{\gamma}^*, \tag{3.3}$$

$$\mathbf{K}_3 = 0.5(1 + 0.25\mathbf{K}_2) \circ \boldsymbol{\gamma}^*, \tag{3.4}$$

$$\begin{aligned}\boldsymbol{\Gamma}_1 &= \boldsymbol{\gamma}_1 + 4(\boldsymbol{\gamma}'' - \boldsymbol{\gamma}'),\\ \mathbf{K}_4 &= (1 + 0.5\mathbf{K}_3) \circ \boldsymbol{\Gamma}_1.\end{aligned} \tag{3.5}$$

2) Algorithm for calculating the increment $\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda}$ of the orientation quaternion at the first step of integration has the form (2.3)

3) The algorithm for calculating the increment of the orientation quaternion $\boldsymbol{\lambda}$ at the first step of integration has the form (2.3).

$$\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_0 \circ \boldsymbol{\lambda}^*, \quad \boldsymbol{\lambda}_0 = \boldsymbol{\lambda}(t_0). \tag{3.6}$$

At the second and subsequent steps of integration, the calculation is carried out according to the following algorithms:

1) Algorithm for calculating the coefficients $\mathbf{K}_i$ included in the formulas of the Runge-Kutta method of the 4th order of accuracy, at the second and subsequent steps of integration:

$$\mathbf{K}_1 = h\boldsymbol{\omega}(t_{n-1}) = \boldsymbol{\gamma}_n = \boldsymbol{\Gamma}_{n-1}, \quad n = 2,3...,$$

where the quaternion $\boldsymbol{\gamma}_n$ is assigned the value of the quaternion $\boldsymbol{\Gamma}_{n-1} = \boldsymbol{\gamma}_{n-1} + 4(\boldsymbol{\gamma}'' - \boldsymbol{\gamma}')$ calculated at the previous integration step when the coefficient $\mathbf{K}_4$ is found at this step.

$$\boldsymbol{\gamma}^* = 5\boldsymbol{\gamma}' + \boldsymbol{\gamma}'' - \boldsymbol{\gamma}_n,$$
$$\mathbf{K}_2 = 0.5\left(1 + 0.25\mathbf{K}_1\right) \circ \boldsymbol{\gamma}^*,$$
$$\mathbf{K}_3 = 0.5\left(1 + 0.25\mathbf{K}_2\right) \circ \boldsymbol{\gamma}^*,$$
$$\boldsymbol{\Gamma}_n = \boldsymbol{\gamma}_n + 4\left(\boldsymbol{\gamma}'' - \boldsymbol{\gamma}'\right),$$
$$\mathbf{K}_4 = \left(1 + 0.5\mathbf{K}_3\right) \circ \boldsymbol{\Gamma}_n,$$

where $\boldsymbol{\gamma}'$ and $\boldsymbol{\gamma}''$ are the input integral information of the algorithm, measured by gyroscopes at the 2nd, 3rd and subsequent integration steps:

$$\boldsymbol{\gamma}' = \int_{t_{n-1}}^{t_{n-1}+h/2} \boldsymbol{\omega}(t)dt, \quad \boldsymbol{\gamma}'' = \int_{t_{n-1}+h/2}^{t_{n-1}+h} \boldsymbol{\omega}(t)dt,$$
$$t_{n-1} + h = t_n, \quad n = 2,3\ldots.$$

2) The algorithm for calculating the increment $\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda}$ of the orientation quaternion at the second and subsequent integration steps has the form (2.3).

3) Algorithm for calculating the orientation quaternion $\boldsymbol{\lambda}$ at the second and subsequent integration steps:

$$\boldsymbol{\lambda}_n = \boldsymbol{\lambda}_{n-1} \circ \boldsymbol{\lambda}^*, \quad n = 2,3\ldots.$$

Algorithms 2) and 3) are the same as the algorithms used in the first step of integration.

The second version of orientation algorithms in increments of Euler (Rodrigues–Hamilton) parameters $\lambda_j^* = \Delta\lambda_j$, implementing the structures of Runge-Kutta methods of the 4th order of accuracy using integral information about the absolute angular motion of the object.

In the first algorithm described above, at the initial moment of time t_0 it is necessary to know the value of the absolute angular velocity vector of the object $\boldsymbol{\omega}(t_0)$. But this requirement can be eliminated by using three measured values of the increments of the vector quasi-angles.

This proposed second version of orientation algorithms in Euler parameter increments, implementing the structures of the Runge-Kutta methods of the 4th order of accuracy using integral information about the absolute angular motion of the object, is as follows:

1) At the first step of the second algorithm for three integrals $\boldsymbol{\gamma}_i$:

$$\boldsymbol{\gamma}_0 = \int_{t_0}^{t_0+h/2} \boldsymbol{\omega}(t)dt, \quad \boldsymbol{\gamma}_1 = \int_{t_0+h/2}^{t_0+h} \boldsymbol{\omega}(t)dt, \quad \boldsymbol{\gamma}_2 = \int_{t_0+h}^{t_0+3h/2} \boldsymbol{\omega}(t)dt,$$

measured at the first three time intervals, the value of the quaternion $\boldsymbol{\gamma}$ is determined by the formula (obtained by Ya. G. Sapunkov):

$$\boldsymbol{\gamma} = h\boldsymbol{\omega}(t_0) = \frac{1}{6}\left(11\boldsymbol{\gamma}_0 - 7\boldsymbol{\gamma}_1 + 2\boldsymbol{\gamma}_2\right).$$

Next, at this step, we use algorithm (3.1)-(3.6), (2.3).

At the second and subsequent steps of integration, we carry out the calculations using the same algorithms that were described above in the first version of the algorithms.

Orientation algorithms in increments of Euler (Rodrigues–Hamilton) parameters $\lambda_j^* = \Delta\lambda_j$, implementing the structures of Runge-Kutta methods of the 3rd order of accuracy using integral information about the absolute angular motion of the object

$$\boldsymbol{\gamma}' = \int_{t_{n-1}}^{t_{n-1}+h/2} \boldsymbol{\omega}(t)dt, \quad \boldsymbol{\gamma}'' = \int_{t_{n-1}+h/2}^{t_{n-1}+h} \boldsymbol{\omega}(t)dt,$$
$$t_{n-1} + h = t_n, \quad n = 1,2,3\ldots.$$

have the following appearance:

1) Algorithm for calculating the coefficients $\mathbf{K}_i$ included in the modified formula of the Runge-Kutta method of the 3rd order of accuracy:

At the first step of integration:

$$\mathbf{K}_1 = \boldsymbol{\gamma} = \boldsymbol{\gamma}_1 = h\boldsymbol{\omega}(t_0).$$

At the second and subsequent steps of integration

$$\mathbf{K}_1 = \boldsymbol{\gamma}_{n-1} + 4\left(\boldsymbol{\gamma}''_{n-1} - \boldsymbol{\gamma}'_{n-1}\right), \quad n = 2,3\ldots,$$
$$\mathbf{K}_2 = 0.5\left(1 + 0.25\mathbf{K}_1\right) \circ \left(5\boldsymbol{\gamma}' + \boldsymbol{\gamma}'' - \boldsymbol{\gamma}\right),$$
$$\mathbf{K}_3 = \left(1 + \mathbf{K}_2 - 0.5\mathbf{K}_1\right) \circ \left(\boldsymbol{\gamma} + 4\left(\boldsymbol{\gamma}'' - \boldsymbol{\gamma}'\right)\right).$$

2) Algorithm for calculating the increment $\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda}$ of the orientation quaternion at the at the integration step:

$$\boldsymbol{\lambda}^* = \Delta\boldsymbol{\lambda} = 1 + \left(1/12\right)\left(\mathbf{K}_1 + 4\mathbf{K}_2 + \mathbf{K}_3\right).$$

3) Algorithm for calculating the increment of the orientation quaternion $\boldsymbol{\lambda}$ at the integration step:

$$\boldsymbol{\lambda}_n = \boldsymbol{\lambda}_{n-1} \circ \boldsymbol{\lambda}^*.$$

We note that all the orientation algorithms we obtained are constructed using a convenient recurrent scheme and that a rational quaternion computational scheme in increments of Euler parameters was used to calculate the inertial orientation of the object, implementing the quaternion formula for adding up the rotations of the object.

IV. MODELING ORIENTATION ALGORITHMS

For modeling orientation algorithms that use integral primary information about the rotational motion of an object (increments of integrals from projections of the absolute angular velocity vector of an object onto the coordinate axes (quasi-coordinates) associated with it), we have proposed various algorithms for forming projections of the absolute angular velocity vector of an object onto the coordinate axes associated with it based on measured quasi-coordinates, which use either the Simpson and trapezoid methods, or second-degree polynomials, or Taylor series. These algorithms are

included in the general quaternion structures of various versions of the second class of orientation algorithms.

Chelnokova L.A. performed mathematical modeling of the above orientation algorithms using the program she developed in order to evaluate their methodological errors. When modeling the object's motion, it is assumed that it performs harmonic oscillations for each of the three angular variables (such as aircraft angles). Harmonic oscillations of the object with different amplitudes and frequencies were modeled, and different integration steps were used. The object's motion time (integration) was 600 sec (10 min). Modeling showed high accuracy characteristics of the proposed algorithms.

We present the results of modeling the orientation algorithms that implement the structures of the Runge-Kutta methods of the 4th order of accuracy using instantaneous or integral information about the absolute angular motion of the object, for one of the options for modeling the angular motion of the object.

The maximum methodical errors in determining the orientation angles for the amplitudes of angular harmonic oscillations of an object equal to 5-15 degrees and oscillation frequencies equal to 0.5-1 Hz,

when using instantaneous information:

for an integration step $h = 0.001$ s, they lie in the range of 10^{-9}-10^{-8} degrees,

for an integration step $h = 0.01$ s – in the range of 10^{-7}-10^{-6} degrees;

when using integral information:

for the integration step $h = 0.001$ s they lie in the range of 10^{-7}-10^{-6} degrees,

for the integration step $h = 0.01$ s – in the range of 10^{-5}-10^{-3} degrees (errors are different for different angular variables).

The errors of the second class of orientation algorithms using input integral information about the object's motion include methodological errors of forming this information based on instantaneous information about the object's motion specified during modeling, and also include methodological errors of the proposed algorithms for forming instantaneous information based on integral information included in the second class of algorithms. Therefore, the errors of the second class of algorithms are significantly greater than the errors of the first class of algorithms.

A comparison of new 4th-order-accurate orientation algorithms, constructed by us on the basis of the classical quaternion differential equation of object orientation in four-dimensional Euler parameters and implementing the structures of 4th-order-accurate Runge-Kutta methods using integral information about the absolute angular motion of the object, with 4th-order-accurate orientation algorithms, constructed by us earlier on the basis of the new quaternion differential equation of orientation in associated quaternions (in quaternions with zero scalar parts) proposed by the speaker using the Picard successive approximation method and using integral information about the absolute angular motion of the object [1-3], showed that these algorithms are comparable in accuracy (with a slight advantage of the algorithms constructed by us on the basis of the quaternion differential equation of orientation in associated quaternions).

REFERENCES

[1] Y.N. Chelnokov, S.E. Perelyaev, “New Equations and Algorithms for Orientation and Navigations of SDINS with Four-Dimensional Skew-Symmetric Operators”, 21st Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2014 – Proceedings 21, pp. 365-369, 2014.

[2] S.E. Perelyaev, Yu.N.Chelnokov, “New algorithms for determining the inertial orientation of an objekt”, Journal of Applied Mathematics and Mechanics, vol. 78, issue 6, pp. 560-567, 2014.

[3] Yu.N. Chelnokov1, S.E. Perelyaev, L.A. Chelnokova, “An Investigation of Algorithms for Estimating the Inertial Orientation of a Moving Object”, Izv. Sarat. Un-ta. New Ser. Ser.: Mathematics. Mechanics. Informatics, vol. 16, issue 1, pp. 80-95, 2016.

Time-Quasi-Optimal Spacecraft Angular Acceleration Constructed According to the Poinsot Kinematic Interpretation

A.V. Molodenkov
Institute of Precision Mechanics and Control of RAS
Saratov, Russia
molalexei@yandex.ru

Ya.G. Sapunkov
Institute of Precision Mechanics and Control of RAS
Saratov, Russia
SapunkovYaG@mail.ru

T.V. Molodenkova
Yu.A. Gagarin Saratov State Technical University
Saratov, Russia
moltw@yandex.ru

***Abstract*— The kinematic problem of finding the time-optimal program angular acceleration of spacecraft under arbitrary boundary conditions for angular position and angular velocity of spacecraft is considered. Based on the Poinsot kinematic interpretation, which represents an arbitrary angular motion of a rigid body around a fixed point as a generalized conical motion, using the Pontryagin maximum principle, the quasi-optimal analytical solution of the problem is obtained, which is brought to the algorithm. This report presents the theoretically and numerically justified analytical quasi-optimal algorithm for constructing program angular acceleration of spacecraft, which, according to the results, is close to optimal.**

Keywords—spacecraft, rigid body, angular acceleration, Poinsot kinematic interpretation, arbitrary boundary conditions, quasi-optimal control, algorithm

I. INTRODUCTION

When solving many problems of guidance and control of spacecraft, it is necessary to know the optimal angular acceleration of spacecraft in one sense or another under arbitrary given boundary conditions. Since the exact solution of the problem is unknown, as a rule, in the literature such acceleration is either calculated numerically or constructed analytically on the basis of polynomials (splines) by representing the orientation quaternion of a spacecraft by polynomials and expressing the angular velocity vector through this quaternion. However, there are no guarantees (proven theorems or considerations from theoretical mechanics) that on the whole set of angular motions of a spacecraft, under any boundary conditions, these analytical solutions will approximate the optimal trajectory of the angular motion of a spacecraft well enough.

The report investigates the kinematic problem of finding the optimal program angular acceleration of spacecraft under arbitrary boundary conditions of the angular position and angular velocity of spacecraft with limited control. It is difficult to obtain an optimal analytical solution to this problem under arbitrary boundary conditions, since the exact solution to the problem of finding the orientation of a rigid body by its known angular velocity (the Darboux problem) in general case is not known. Within the bounds of the Poinsot classical kinematic interpretation of angular motion of a rigid body around a fixed point as a generalized conical motion, the problem of finding the optimal angular acceleration of spacecraft is reformulated in this class of motions. The trajectory of spacecraft is given by explicit expressions that contain arbitrary quaternionic and scalar constants and two arbitrary scalar functions that are parameters of generalized conical motion. In relation to these parameters and their derivatives, an optimization problem is formulated and solved, in which the second derivatives of these parameters act as controls. At the same time, the generality of the original problem is practically not violated, since the known exact solution to the classical problem of finding optimal angular acceleration in the case of plane Eulerian rotation of spacecraft or the exact solution in the new special case of regular precession, which is obtained in the report, and similar solutions to the problem within the bounds of the concept under consideration completely coincide (or, in other words, the known analytical optimal solutions of the problem under their boundary conditions are "embedded" in the proposed approach); in other cases, in numerical calculations of the solution of the initial problem and the proposed analytical solution, relative error between values of the optimization functional (time performance index), which is the defining characteristic of the problem, is no more than one percent including turns of spacecraft at large angles. Therefore, the proposed analytical solution of the problem can be interpreted as quasi-optimal with respect to the classical problem of finding the optimal angular acceleration of a spacecraft. Explicit expressions for the orientation quaternion and the angular velocity vector of a spacecraft are given, an explicit formula for the angular acceleration of a spacecraft is obtained by differentiating the expression of the angular velocity vector, an analytical algorithm for solving the problem is provided. This algorithm can be used on board of spacecraft.

The developed analytical solution method has previously been successfully applied in full dynamic problems of optimal spacecraft turns of various configurations with different optimization criteria [1-7]. In the report at the XXXI ICINS-2024, the authors presented a simpler smooth quasi-optimal solution to the problem of finding for optimal angular acceleration of spacecraft under arbitrary boundary conditions with a fixed reorientation time and unlimited control [8].

II. STATEMENT OF THE CLASSICAL PROBLEM

The angular motion of a spacecraft as a rigid body around the center of mass is described by equations [9]:

$$2\dot{\mathbf{\Lambda}} = \mathbf{\Lambda} \circ \boldsymbol{\omega}(t) \tag{2.1}$$

$$\dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon} \tag{2.2}$$

Here $\mathbf{\Lambda}(t) = \lambda_0(t) + \lambda_1(t)i_1 + \lambda_2(t)i_2 + \lambda_3(t)i_3$ is the quaternion describing the position of a spacecraft in an

inertial space; $\boldsymbol{\omega}(t)=\omega_1(t)\mathbf{i}_1+\omega_2(t)\mathbf{i}_2+\omega_3(t)\mathbf{i}_3$ is the angular velocity vector of a spacecraft specified by its projections onto spacecraft-fixed coordinate axes; i_1,i_2,i_3 – the units of the hypercomplex space (imaginary Hamiltonian units), which can be identified with the vectors of a three-dimensional vector space $\mathbf{i}_1,\mathbf{i}_2,\mathbf{i}_3$; the symbol "$\circ$" stands for the quaternion product; $\boldsymbol{\varepsilon}(t)=[\varepsilon_1(t),\varepsilon_2(t),\varepsilon_3(t)]^T$ – the vector of angular acceleration. Phase coordinates $\boldsymbol{\Lambda},\boldsymbol{\omega}$ and control $\boldsymbol{\varepsilon}$ satisfy the requirements of the optimal control problem [8] ($\boldsymbol{\Lambda}(t),\boldsymbol{\omega}(t)$ are continuous functions and $\boldsymbol{\varepsilon}(t)$ is a piecewise continuous function); quaternion $\boldsymbol{\Lambda}(t)$ is normalized, i.e. $\|\boldsymbol{\Lambda}\|=\lambda_0{}^2+\lambda_1{}^2+\lambda_2{}^2+\lambda_3{}^2=1$.

The angular acceleration vector is modulo bounded:

$$|\boldsymbol{\varepsilon}|\le\varepsilon_{\max} \tag{2.3}$$

Arbitrary boundary conditions for the angular position and angular velocity of a spacecraft are given:

$$\boldsymbol{\Lambda}(0)=\boldsymbol{\Lambda}_0,\ \boldsymbol{\Lambda}(T)=\boldsymbol{\Lambda}_T \tag{2.4}$$

$$\boldsymbol{\omega}(0)=\boldsymbol{\omega}_0,\ \boldsymbol{\omega}(T)=\boldsymbol{\omega}_T \tag{2.5}$$

The task is to find the optimal control $\boldsymbol{\varepsilon}^{\text{opt}}(t)$ that provides a minimum of the functional

$$J=T \tag{2.6}$$

III. Transition to dimensionless variables

Since the main variable of the problem quaternion $\boldsymbol{\Lambda}$ is a dimensionless quantity, using the formulas

$$t^{\text{dimls}}=t\,\varepsilon_{\max}^{1/2},\ \boldsymbol{\omega}^{\text{dimls}}=\boldsymbol{\omega}\,\varepsilon_{\max}^{-1/2},\ \boldsymbol{\varepsilon}^{\text{dimls}}=\boldsymbol{\varepsilon}\,\varepsilon_{\max}^{-1}$$

we reformulate the problem in dimensionless variables, while all expressions of the problem statement except limitation on the modulus of the spacecraft angular acceleration vector

$$|\boldsymbol{\varepsilon}|\le 1 \tag{3.1}$$

will not change. Next, we consider the dimensionless formulation of problem (2.1), (2.2), (2.4)–(2.6), and (3.1).

IV. Application of the maximum principle

Applying the Pontryagin maximum principle [9, 10], Applying the Pontryagin maximum principle [8, 9], we introduce the conjugate variables $\boldsymbol{\Psi}(t)$ (quaternion) and $\boldsymbol{\varphi}(t)$ (vector) to the phase variables $\boldsymbol{\Lambda}(t),\boldsymbol{\omega}(t)$, respectively. The Hamilton-Pontryagin function

$$H=(\boldsymbol{\Psi},\boldsymbol{\Lambda}\circ\boldsymbol{\omega})/2+(\boldsymbol{\varphi},\boldsymbol{\varepsilon}) \tag{4.1}$$

where "(. , .)" is the scalar product of vectors.

Equations for conjugate variables

$$\begin{cases}2\dot{\boldsymbol{\Psi}}=\boldsymbol{\Psi}\circ\boldsymbol{\omega},\\ \dot{\boldsymbol{\varphi}}=-\,\mathbf{vect}\left(\widetilde{\boldsymbol{\Lambda}}\circ\boldsymbol{\Psi}\right)/2\end{cases} \tag{4.2}$$

where "**vect**(.)" is the vector part of a quaternion and the symbol "~" means quaternion conjugation.

Quaternion equations for variables $\boldsymbol{\Psi}$ and $\boldsymbol{\Lambda}$ coincide; therefore, their solutions differ by a constant quaternion $\mathbf{C}$:

$$\boldsymbol{\Psi}=\mathbf{C}\circ\boldsymbol{\Lambda} \tag{4.3}$$

Due to this, and taking into account the designation [9]

$$\mathbf{p}=\mathbf{vect}(\widetilde{\boldsymbol{\Lambda}}\circ\boldsymbol{\Psi})=\widetilde{\boldsymbol{\Lambda}}\circ\mathbf{c}_v\circ\boldsymbol{\Lambda} \tag{4.4}$$

where $\mathbf{c}_v$ is the vector part of quaternion $\mathbf{C}$, expressions (4.2) take the form

$$\begin{cases}\mathbf{p}=\widetilde{\boldsymbol{\Lambda}}\circ\mathbf{c}_v\circ\boldsymbol{\Lambda},\\ \dot{\boldsymbol{\varphi}}=-\,\mathbf{p}/2\end{cases} \tag{4.5}$$

Thus, due to the self-adjointness of the linear differential system of equations (2.1), the dimension of the boundary value problem of the maximum principle is reduced by four [9]. From the maximum condition (4.1), we obtain the structure of the optimal control:

$$\boldsymbol{\varepsilon}^{\text{опт}}=\boldsymbol{\varphi}/|\boldsymbol{\varphi}| \tag{4.6}$$

From (2.1), (2.2), (4.5), (4.6) we have:

$$\dot{\mathbf{p}}=[\mathbf{p},\boldsymbol{\omega}],\quad \mathbf{p}=-2|\boldsymbol{\varphi}|\frac{d^2\boldsymbol{\omega}}{dt^2}-2\frac{d\boldsymbol{\omega}}{dt}\frac{d|\boldsymbol{\varphi}|}{dt}$$

where "[. , .]" denotes the vector product.

Therefore, on the entire interval of motion, the optimal angular velocity of a spacecraft must satisfy the third-order differential vector equation

$$\frac{d^3\boldsymbol{\omega}}{dt^3}=\left[\frac{d^2\boldsymbol{\omega}}{dt^2},\boldsymbol{\omega}\right]+\frac{1}{|\boldsymbol{\varphi}|}\left(\frac{d|\boldsymbol{\varphi}|}{dt}\left(\left[\frac{d\boldsymbol{\omega}}{dt},\boldsymbol{\omega}\right]-2\frac{d^2\boldsymbol{\omega}}{dt^2}\right)-\frac{d^2|\boldsymbol{\varphi}|}{dt^2}\frac{d\boldsymbol{\omega}}{dt}\right) \tag{4.7}$$

Thereby, the solution of the optimal control problem is reduced to the solution of boundary value problem (2.1), (4.7), (2.4), and (2.5).

Given variable $\mathbf{p}$ (4.4), the Hamilton-Pontryagin function (4.1) will be written as follows:

$$H=(\mathbf{p},\boldsymbol{\omega})/2+(\boldsymbol{\varphi},\boldsymbol{\varepsilon}) \tag{4.8}$$

V. The exact solution of the problem of finding the optimal angular acceleration of spacecraft in the particular case

We will seek the solution of equations (2.1), (4.7) in the class of conical motions with constant $|\boldsymbol{\varphi}|=c$. To do this, we represent the optimal angular velocity of spacecraft as

$$\boldsymbol{\omega}=\widetilde{\mathbf{K}}\circ(\mathbf{i}_1\alpha\sin\Omega t+\mathbf{i}_2\alpha\cos\Omega t+\mathbf{i}_3\Omega)\circ\mathbf{K} \tag{5.1}$$

where $\mathbf{K}$ (quaternion) and α,Ω are undetermined constants; $\|\mathbf{K}\|=1$. The quaternion $\mathbf{K}$ makes it possible to rotate the vector in parentheses in (5.1) about a constant axis passing through the fixed point of spacecraft. Successively differentiating (5.1) three times with respect to the variable t and substituting the results in (4.7), one can verify that

equality holds. At the same time, we note that $\ddot{\boldsymbol{\omega}} = [\ddot{\boldsymbol{\omega}}, \boldsymbol{\omega}] = (\ddot{\boldsymbol{\omega}} \circ \boldsymbol{\omega} - \boldsymbol{\omega} \circ \ddot{\boldsymbol{\omega}})/2$.

The trajectory of spacecraft motion with angular velocity (5.1) in (2.1) can be found explicitly and has the form of the regular precession

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}} \circ \mathbf{exp}\{\mathbf{i}_2 \alpha t/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega t/2\} \circ \mathbf{K} \qquad (5.2)$$

where "$\mathbf{exp}\{.\}$" denotes the quaternion exponential function [9]. Expressions (5.1), (5.2) involve five arbitrary constants $K_0, K_1, K_2, \alpha, \Omega$. The restriction on the modulus of the angular acceleration vector $\boldsymbol{\varepsilon}$ (3.1) is fulfilled under the condition $\alpha^2\Omega^2 = 1$. Let us satisfy the boundary conditions (2.3), (2.4). Because of the insufficient number of arbitrary constants in solution (5.1), we impose additional requirements on the quantities $|\boldsymbol{\omega}_0|$ and $\boldsymbol{\omega}_T$. The direction of the unit vector of the initial angular velocity $\boldsymbol{\omega}_0^e = \boldsymbol{\omega}_0/|\boldsymbol{\omega}_0|$ is arbitrary and given. At t=0, we obtain from (5.1):

$$\boldsymbol{\omega}_0 = |\boldsymbol{\omega}_0|\boldsymbol{\omega}_0^e = \widetilde{\mathbf{K}} \circ (\mathbf{i}_2\alpha + \mathbf{i}_3\Omega) \circ \mathbf{K} \qquad (5.3)$$

$$\|\boldsymbol{\omega}_0\| = \|\widetilde{\mathbf{K}} \circ (\mathbf{i}_2\alpha + \mathbf{i}_3\Omega) \circ \mathbf{K}\| = \|\widetilde{\mathbf{K}}\|\|\mathbf{i}_2\alpha + \mathbf{i}_3\Omega\|\|\mathbf{K}\| = \alpha^2 + \Omega^2 \qquad (5.4)$$

At the right end of the spacecraft trajectory at t=T from (5.2) we have

$$\boldsymbol{\Lambda}_T = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}} \circ \mathbf{exp}\{\mathbf{i}_2 \alpha T/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega T/2\} \circ \mathbf{K} \qquad (5.5)$$

$$\mathrm{scal}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T) = \mathrm{scal}(\mathbf{exp}\{\mathbf{i}_2 \alpha T/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega T/2\}) \qquad (5.6)$$

where "scal(.)" denotes the scalar part of the quaternion. From (5.2), (5.3) – (5.5) we find | T, $|\boldsymbol{\omega}_0|$, α, Ω, and $\mathbf{K}$.

Let us represent (5.3), (5.5) in the form:

$$(\mathbf{i}_2\alpha + \mathbf{i}_3\Omega) \circ \mathbf{K} - \mathbf{K} \circ \boldsymbol{\omega}_0 = 0$$

$$\mathbf{exp}\{\mathbf{i}_2 \alpha T/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega T/2\} \circ \mathbf{K} - \mathbf{K} \circ \widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T = 0$$

or the same notation based on m- and n-matrices isomorphic to quaternions [11]:

$$\left(\begin{bmatrix} 0 & 0 & -\alpha & -\Omega \\ 0 & 0 & -\Omega & \alpha \\ \alpha & \Omega & 0 & 0 \\ \Omega & -\alpha & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -\omega_{01} & -\omega_{02} & -\omega_{03} \\ \omega_{01} & 0 & \omega_{03} & -\omega_{02} \\ \omega_{02} & -\omega_{03} & 0 & \omega_{01} \\ \omega_{03} & \omega_{02} & -\omega_{01} & 0 \end{bmatrix} \right) \begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (5.7)$$

$$\left(\begin{bmatrix} m_0 & -m_1 & -m_2 & -m_3 \\ m_1 & m_0 & -m_3 & m_2 \\ m_2 & m_3 & m_0 & -m_1 \\ m_3 & -m_2 & m_1 & m_0 \end{bmatrix} - \begin{bmatrix} n_0 & -n_1 & -n_2 & -n_3 \\ n_1 & n_0 & n_3 & -n_2 \\ n_2 & -n_3 & n_0 & n_1 \\ n_3 & n_2 & -n_1 & n_0 \end{bmatrix} \right) \begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (5.8)$$

where the elements of the coefficient matrix of the system of linear algebraic equations (SLAE) (5.8) are determined by the components of quaternions:

$$\mathbf{m} = \mathbf{exp}\{\mathbf{i}_2 \alpha T/2\} \circ \mathbf{exp}\{\mathbf{i}_3 \Omega T/2\} \qquad (5.9)$$

$$\mathbf{n} = \widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T \qquad (5.10)$$

At the same time

$$\|\mathbf{m}\| = 1 \qquad (5.11) \qquad\qquad \|\mathbf{n}\| = 1 \qquad (5.12)$$

Note that the ranks of the coefficient matrices of systems (5.7) and (5.8) are 2.

Choosing two linearly independent equations in each of systems (5.7) and (5.8), we obtain the SLAE

$$\begin{bmatrix} 0 & \omega_{01} & \omega_{02}-\alpha & \omega_{03}-\Omega \\ -\omega_{01} & 0 & -(\omega_{03}+\Omega) & \omega_{02}+\alpha \\ m_1-n_1 & 0 & -(m_3+n_3) & m_2+n_2 \\ 0 & n_1-m_1 & n_2-m_2 & n_3-m_3 \end{bmatrix} \begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (5.13)$$

For the existence of nontrivial solutions of the homogeneous SLAE, the determinant of the coefficient matrix of SLAE must be zero. Thus, taking into account (5.4), (5.5), (5.18), and (5.19), we obtain

$$|\boldsymbol{\omega}_0| = (m_2\alpha + m_3\Omega)/(n_1\omega_{01}^e + n_2\omega_{02}^e + n_3\omega_{03}^e) \qquad (5.14)$$

From (5.4), (5.6), and (5.14) with regard to (5.9), (5.10), we obtain two equations to determine α, Ω, and T:

$$\begin{cases} \alpha^2\Omega^2 = 1, \\ (\alpha^2+\Omega^2)(n_1\omega_{01}^e + n_2\omega_{02}^e + n_3\omega_{03}^e)^2 - (m_2\alpha + m_3\Omega)^2 = 0, \\ \mathrm{scal}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T) - \cos(\alpha T/2)\cos(\Omega T/2) = 0 \end{cases} \qquad (5.15)$$

The unknown $|\boldsymbol{\omega}_0|$ can be found by (5.14) after solving (5.15). In the system (5.17), we replace the last equation with condition $\|\mathbf{K}\| = 1$. From the first three equations in (5.13), K_0, K_1, K_2 are expressed in terms of K_3:

$$K_0 = A_0K_3/D, \quad K_1 = A_1K_3/D, \quad K_2 = A_2K_3/D \qquad (5.16)$$

$$A_0 = \omega_{01}\left[-(\omega_{03}+\Omega)(m_2+n_2) + (\omega_{02}+\alpha)(m_3+n_3)\right] \qquad (5.17)$$

$$A_1 = \omega_{01}\left[(\Omega-\omega_{03})\frac{D}{\omega_{01}^2} - (m_2+n_2)(\omega_{02}-\alpha) - (m_1+n_1)\frac{\omega_{02}^2-\alpha^2}{\omega_{01}}\right] \qquad (5.18)$$

$$A_2 = \omega_{01}\left[(m_1-n_1)(\omega_{02}+\alpha) + (m_2+n_2)\omega_{01}\right] \qquad (5.19)$$

$$D = \omega_{01}\left[(m_1-n_1)(\omega_{03}+\Omega) + (m_3+n_3)\omega_{01}\right] \qquad (5.20)$$

From (5.17)–(5.20) and $\|\mathbf{K}\| = 1$ we have

$$K_3 = \left[1 + (A_0/D)^2 + (A_1/D)^2 + (A_2/D)^2\right]^{1/2} \qquad (5.21)$$

After determining K_3 from (5.21), K_0, K_1, K_2 are calculated using (5.16). The angular velocity boundary condition at the t=T must have the form

$$\boldsymbol{\omega}(T) = \boldsymbol{\omega}_T = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1\alpha\sin\Omega T + \mathbf{i}_2\alpha\cos\Omega T + \mathbf{i}_3\Omega) \circ \mathbf{K} \qquad (5.22)$$

Thus, in cases when the angular velocity boundary conditions satisfy constraints (5.14), (5.22) (this implies that the vector $\boldsymbol{\omega}$ belongs, on the whole time interval of motion, to a conical surface that is defined by the arbitrary spacecraft attitude boundary conditions $\boldsymbol{\Lambda}_0, \boldsymbol{\Lambda}_T$ and the arbitrary direction of the angular velocity initial value $\boldsymbol{\omega}_0^e$), the trajectory of the angular motion of spacecraft is in the class of conical motions, and it can be calculated by explicit analytic expressions (5.1) and (5.2).

The optimal angular acceleration from (2/2), (5.1)

$$\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} = \alpha\Omega\widetilde{\mathbf{K}} \circ (\mathbf{i}_1\cos\Omega t - \mathbf{i}_2\sin\Omega t) \circ \mathbf{K} \qquad (5.23)$$

$$|\boldsymbol{\varepsilon}|^2 = \alpha^2\Omega^2 = 1 \qquad (5.24)$$

The conjugate variables $\boldsymbol{\varphi}$ and $\mathbf{p}$ can be found using (5.23), (4.5), and (4.4). Thus, the problem is completely solved subject to the given constraints.

An algorithm for solving the problem of finding the optimal acceleration of spacecraft in the class of regular conical motions:

1) according to the given quaternions $\mathbf{\Lambda}_0, \mathbf{\Lambda}_T$ (2.4), and the unit vector $\boldsymbol{\omega}_0^e$ from (2.5), (5.15), and (5.14) (taking into account (5.19), and (5.10)) to calculate α, Ω, $|\boldsymbol{\omega}_0|$, and T;

2) using $\mathbf{\Lambda}_0, \mathbf{\Lambda}_T$, α, Ω, $|\boldsymbol{\omega}_0|$, and T calculate the components of the quaternion $\mathbf{\mathcal{K}}$ by formulas (5.16)–(5.21);

3) using the formula $\boldsymbol{\omega}_0^{\text{calc}} = |\boldsymbol{\omega}_0| \boldsymbol{\omega}_0^e$ and (5.22)

$$\boldsymbol{\omega}_T^{\text{calc}} = \widetilde{\mathbf{\mathcal{K}}} \circ (\mathbf{i}_1 \alpha \sin \Omega T + \mathbf{i}_2 \alpha \cos \Omega T + \mathbf{i}_3 \Omega) \circ \mathbf{\mathcal{K}}$$

are calculated the vectors $\boldsymbol{\omega}_0^{\text{calc}}, \boldsymbol{\omega}_T^{\text{calc}}$;

4) the resultant $\boldsymbol{\omega}_0^{\text{calc}}, \boldsymbol{\omega}_T^{\text{calc}}$ are compared with $\boldsymbol{\omega}_0, \boldsymbol{\omega}_T$ defined in (2.5);

5) if the equation at step 4 of the algorithm holds true, then the optimal solution is found in the class of conical motions. The angular velocity of the spacecraft, the trajectory of its angular motion, the vector of the control torque, and the value of the performance index are estimated by (5.1), (5.2), and (5.23), respectively;

6) the conjugate variables of the problem $\boldsymbol{\varphi}$ and $\mathbf{p}$ are calculated by (4.4), (4.5), and (5.23).

The optimal solution of the problem obtained in the class of regular conical motions and the previously known [9] exact solution in the class of plane Eulerian rotations for special cases of boundary conditions for the angular velocity of a spacecraft will be used as analytical confirmations in constructing the quasi-optimal solution to the problem of finding the angular acceleration of spacecraft in the class of generalized conical motions.

VI. An analytical quasi-optimal solution

Generalizing formulas (5.1) and (5.2), we obtain expressions for the angular velocity vector and trajectory of spacecraft:

$$\boldsymbol{\omega} = \widetilde{\mathbf{\mathcal{K}}} \circ (\mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t)) \circ \mathbf{\mathcal{K}} \qquad (6.1)$$

$$\begin{aligned} \mathbf{\Lambda}(t) = \mathbf{\Lambda}_0 \circ \widetilde{\mathbf{\mathcal{K}}} \circ \mathbf{exp}\{-\mathbf{i}_3 g(0)/2\} \circ \mathbf{exp}\{-\mathbf{i}_2 f(0)/2\} \circ \\ \circ \mathbf{exp}\{\mathbf{i}_2 f(t)/2\} \circ \mathbf{exp}\{\mathbf{i}_3 g(t)/2\} \circ \mathbf{\mathcal{K}} \end{aligned} \qquad (6.2)$$

where $f(t)$ and $g(t)$ are arbitrary functions of time. Formulas (6.1), (6.2) include expressions for the angular velocity and trajectory of spacecraft of two known exact solutions to the problem of finding the optimal angular acceleration of spacecraft specified in section V. In addition, the proposed solution structure well correlates with Poinsot's kinematic interpretation of arbitrary angular motion of a rigid body around a fixed point as some generalized conical motion [12]. From formulas (2.2), (6.1) we find $\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}}$.

Using the maximum principle [10], solving the time-optimal problem $J = T \to \min$ under the conditions:

$$\dot{f} = f_1, \quad \dot{g} = g_1, \quad \dot{f}_1 = u_1, \quad \dot{g}_1 = u_2 \qquad (6.3)$$

$$\sqrt{u_1^2 + u_2^2} \le u_*, \ 0 < u_* < 1 \text{ for execution (2.3)} \qquad (6.4)$$

$$f = f(0), \ f_1 = f_1(0), \ g = g(0), \ g_1 = g_1(0) \qquad (6.5)$$

$$f = f(T), \ f_1 = f_1(T), \ g = g(T), \ g_1 = g_1(T) \qquad (6.6)$$

where f, f_1, g, g_1 are phase coordinates, u_1, u_2 are controls, we obtain the following quasi-optimal algorithm for solving the initial problem of finding the optimal angular acceleration of spacecraft.

1. Controls u_1, u_2 are defined as:

$$\begin{gathered} u_1 = u_*(-c_1 t + c_3)/U(t), \quad u_2 = u_*(-c_2 t + c_4)/U(t), \\ U(t) = \sqrt{(-c_1 t + c_3)^2 + (-c_2 t + c_4)^2} \end{gathered} \qquad (6.7)$$

2. If $c_1 c_4 - c_2 c_3 \ne 0$, then u_1, u_2 are continuous parameters and from (2.1) $\boldsymbol{\varepsilon}(t)$ is a continuous control function. Explicit expressions for f, f_1, g, g_1 containing arbitrary constants $c_1, \ldots, c_5, c_7, c_8$:

$$\begin{aligned} f &= -c_1 u_* \{(t/2 - A/2 - Bc_2/c_1) F(t) + \\ &\quad + B[B/2 + (t-A)c_2/c_1] \ln\big(t - A + F(t)\big)\}/C + \\ &\quad + c_5 t + c_6, \\ g &= -c_2 u_* \{(t/2 - A/2 + Bc_1/c_2) F(t) - \\ &\quad - B[-B/2 + (t-A)c_1/c_2] \ln\big(t - A + F(t)\big)\}/C + \\ &\quad + c_7 t + c_8, \\ f_1 &= -c_1 u_* [F(t) + B \ln\big(t - A + F(t)\big) c_2/c_1]/C + c_5 \\ g_1 &= -c_2 u_* [F(t) - B \ln\big(t - A + F(t)\big) c_1/c_2]/C + c_7 \\ A &= (c_1 c_3 + c_2 c_4)/C^2, \quad B = (c_1 c_4 - c_2 c_3)/C^2 \\ C &= \sqrt{c_1^2 + c_2^2}, \ F(t) = \sqrt{(t-A)^2 + B^2} \end{aligned} \qquad (6.8)$$

3. If $c_1 c_4 - c_2 c_3 = 0$, then $c_3/c_1 = c_4/c_2 = t_*$ and u, u_2 take on the appearance of piecewise-constant

$$\begin{aligned} u_1 &= -c_1 u_* \operatorname{sign}(t - t_*)/C, \\ u_2 &= -c_2 u_* \operatorname{sign}(t - t_*)/C \end{aligned} \qquad (6.9)$$

and in this case $\boldsymbol{\varepsilon}(t)$ is a piecewise-continuous function. Then:

$$\begin{aligned} f &= -c_1 u_* (t - t_*)^2 \operatorname{sign}(t - t_*)/2C + c_5 (t - t_*) + c_6 \\ g &= -c_2 u_* (t - t_*)^2 \operatorname{sign}(t - t_*)/2C + c_7 (t - t_*) + c_8 \\ f_1 &= -c_1 u_* (t - t_*) \operatorname{sign}(t - t_*)/C + c_5 = \\ &= -c_1 u_* |t - t_*|/C + c_5 \\ g_1 &= -c_2 u_* (t - t_*) \operatorname{sign}(t - t_*)/C + c_7 = \\ &= -c_2 u_* |t - t_*|/C + c_7 \end{aligned} \qquad (6.10)$$

4. Satisfaction of the boundary conditions (2.4), (2.5):

$$\begin{aligned} \widetilde{\mathbf{K}}_1 \circ \widetilde{\mathbf{K}}_2 \circ (\mathbf{i}_1 f_1(0) \sin g(0) + \mathbf{i}_2 f_1(0) \cos g(0) + \\ + \mathbf{i}_3 g_1(0)) \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\omega}_0 \end{aligned} \qquad (6.11)$$

$$\begin{aligned} \widetilde{\mathbf{K}}_1 \circ \widetilde{\mathbf{K}}_2 \circ (\mathbf{i}_1 f_1(T) \sin g(T) + \mathbf{i}_2 f_1(T) \cos g(T) + \\ + \mathbf{i}_3 g_1(T)) \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\omega}_T \end{aligned} \qquad (6.12)$$

$$\begin{aligned} \mathbf{\Lambda}_0 \circ \widetilde{\mathbf{K}}_1 \circ \widetilde{\mathbf{K}}_2 \circ \mathbf{exp}\{-\mathbf{i}_3 g(0)/2\} \circ \mathbf{exp}\{\mathbf{i}_2 (f(T) \\ - f(0))/2\} \circ \mathbf{exp}\{\mathbf{i}_3 g(T)/2\} \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \mathbf{\Lambda}_T \end{aligned} \qquad (6.13)$$

where the quaternion $\mathbf{\mathcal{K}}$ is represented as the product

$$\mathbf{K} = \mathbf{K}_2 \circ \mathbf{K}_1, \quad \mathbf{K}_1 = \mathbf{exp}\{\mathbf{i}_1 \alpha_1 / 2\}, \quad \mathbf{K}_2 = \mathbf{exp}\{\mathbf{i}_2 \alpha_2 / 2\} \tag{6.14}$$

α_1, α_2 are some constants.

5. From formulas (6.10)–(6.13) and expression for the Hamilton-Pontryagin function of problem (6.3)–(6.6) for $t = T$

$$H = c_1 f_1 + c_2 g_1 + c_3 u_1 + c_4 u_2 - (c_1 u_1 + c_2 u_2) T = 1$$

taking into account (6.7), (6.8) or (6.9), (6.10) constants $c_1, ..., c_5, c_7, c_8, \alpha_1, \alpha_2$ and time T are determined, and functions f, f_1, g, g_1 are constructed.

6. The angular velocity vector, the orientation quaternion, and the angular acceleration vector of spacecraft are calculated by formulas (6.1), (6.2), (2.2).

The presented material of the report in the most complete form, including supporting numerical examples, is presented in the article [13].

REFERENCES

[1] A.V. Molodenkov, Y.G. Sapunkov, "Analytical quasi-optimal solution for the problem on turn maneuver of an arbitrary solid with arbitrary boundary conditions," Mechanics of Solids, vol. 54, no. 3, 2019, pp. 474-485.

[2] A.V. Molodenkov, Y.G. Sapunkov, " Analytical Solution of the Optimal Slew Problem for an Axisymmetric Spacecraft in the Class of Conical Motions," Journal of Computer and Systems Sciences International, vol. 55, no. 6, 2016, pp. 969-985.

[3] A.V. Molodenkov, Y.G. Sapunkov, "Analytical Solution of the Minimum Time Slew Maneuver Problem for an Axially Symmetric Spacecraft in the Class of Conical Motions," Journal of Computer and Systems Sciences International, vol. 57, no. 2, 2018, pp. 302-318.

[4] Y.G. Sapunkov, A.V. Molodenkov, "Analytical solution of the problem on an axisymmetric spacecraft attitude maneuver optimal with respect to a combined functional," Automation and Remote Control, vol. 82, no. 7, 2021, pp. 1183–1200.

[5] A.V. Molodenkov, Y.G. Sapunkov, "Analytical quasi-optimal solution of the problem of the time-optimal rotation of a spacecraft," Journal of Computer and Systems Sciences International, vol. 60, no. 4, 2021, pp. 639-653.

[6] A.V. Molodenkov, Y.G. Sapunkov, "Analytical Quasi-Optimal Algorithm for the Programmed Control of the Angular Motion of a Spacecraft," Journal of Computer and Systems Sciences International, vol. 62, no. 3, 2023, pp. 569-580.

[7] Y.G. Sapunkov, A.V. Molodenkov, "Quasioptimal Spacecraft Attitude Control Constructed According to the Poinsot Concept," Aerospace, vol. 10, no. 5, 2023, p. 402.

[8] A.V. Molodenkov, Y.G. Sapunkov, T.V. Molodenkova, "Quasi-Optimal Angular Acceleration of Spacecraft Obtained on the Basis of the Poinsot Concept," 31st Saint Petersburg International Conference on Integrated Naviogation Systems, Collection of plenary and poster paper: Saint Petersburg, 2024, pp. 443-449.

[9] V.N. Branetz, I.P. Shmyglevskij, Primenenie kvaternionov v zadachah orientacii tverdogo tela (Application of quaternions to rigid body orientation problems). Nauka, Moscow, 1973

[10] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, The Mathematical Theory of Optimal Processes. Nauka, Moscow, 1961; Gordon and Breach, New York, 1986.

[11] Yu.N. Chelnokov, Kvaternionnye i bikvaternionnye modeli i metody mehaniki tverdogo tela i ih prilogeniya (Quaternion and Biquaternion Models and Methods in Mechanics of Solids and Their Applications). Fizmatlit, Moscow, 2006.

[12] A.V. Molodenkov, S.E. Perelyaev, "Solution of approximate equation for modified rodrigues vector and attitude algorithm design," Journal of Guidance, Control, and Dynamics, vol. 44, no. 6, 2021, pp. 1224-1227.

[13] Y.G. Sapunkov, A.V. Molodenkov, "Quasi-Optimal Angular Acceleration of a Spacecraft Based on the Poinsot Concept," Automation and Remote Control, vol. 86, no. 3, 2025, pp. 217–234.

The Problem of Spacecraft Optimal Flight to the Orbit of a Moon Satellite Using Limited Thrust and New Regular Quaternion Equations of the Perturbed Spatial Restricted Three-Body Problem in KS-variables

I.A. Pankratov
National Research Saratov State University,
Institute for Problems of Precision Mechanics
and Control RAS
Saratov, Russia
PankratovIA.mechanic@gmail.com

Yu.N. Chelnokov
Institute for Problems of Precision Mechanics
and Control RAS
Saratov, Russia
ChelnokovYuN@gmail.com

***Abstract* — In the report, using new regular quaternion differential equations of the perturbed spatial restricted three-body problem in *KS*-variables, the time-optimal control problem for a spacecraft transfer to a lunar satellite orbit was solved in a nonlinear formulation. The control vector (perturbing acceleration of the spacecraft center of mass) is bounded in magnitude. During the solution of the problem, it is necessary to determine the optimal spatial orientation of this vector. Regular quaternion differential equations of orbital motion were used constructed within the framework of the perturbed spatial restricted three-body problem (Earth, Moon, spacecraft) employing *KS*-variables to describe the motion of the SC's center of mass. The stated problem was solved using the Pontryagin maximum principle. The problem of constructing optimal transfer trajectories to a lunar satellite orbit was reduced to a boundary value problem with a free right endpoint, described by a system of twentieth-order nonlinear differential equations. Examples of numerical solutions to the problem are given. We have constructed graphs for the variation of quaternion components characterizing the spacecraft position and velocity, Kepler energy and real time, optimal control. An analysis of the obtained solutions is given. The features and regularities of optimal spacecraft transfers to the orbits of Moon satellite are established.**

Keywords — spacecraft, spatial motion, motion control, restricted three-body problem, Kustaanheimo-Stiefel variables, maximum principle, quaternion, perturbed motion.

I. STATEMENT OF THE PROBLEM

The quaternion differential equations of spacecraft orbital motion, developed by Yu.N. Chelnokov within the framework of the perturbed spatial restricted three-body problem (Earth, Moon, spacecraft) using *KS*-variables, have the following form [1, 2]:

$$\frac{d\boldsymbol{u}_1}{d\tau} = \boldsymbol{s}_1, \tag{1}$$

$$\frac{d\boldsymbol{s}_1}{d\tau} = \frac{1}{2}h_1^*\boldsymbol{u}_1 - \frac{1}{2}r_1\left\{fm_0 r_1 r_0^{-3}\boldsymbol{u}_1 + \right. \\ \left. + \boldsymbol{i}\circ\boldsymbol{u}_1\circ\left[fm_0\left(r_{01}^{-3} - r_0^{-3}\right)\boldsymbol{R}_{01}\right]\right\} - \left(r_1/2\right)\boldsymbol{i}\circ\boldsymbol{u}_1\circ\boldsymbol{p}, \tag{2}$$

$$\frac{dh_1^*}{d\tau} = -fm_0 r_0^{-3} r_1\frac{dr_1}{d\tau} + fm_0\left(r_{01}^{-3} - r_0^{-3}\right)\left(\frac{d\boldsymbol{r}_1}{d\tau}\cdot\boldsymbol{r}_{01}\right) + \frac{d\boldsymbol{r}_1}{d\tau}\cdot\boldsymbol{p}, \tag{3}$$

$$\frac{dt}{d\tau} = r_1, \tag{4}$$

$$\boldsymbol{R}_{01} = x_{01}\boldsymbol{i} + y_{01}\boldsymbol{j} + z_{01}\boldsymbol{k},$$
$$r_0 = \sqrt{x_0^2 + y_0^2 + z_0^2}, \qquad r_1 = \sqrt{x_1^2 + y_1^2 + z_1^2}, \tag{5}$$
$$r_{01} = \sqrt{\left(x_{01} - x_0\right)^2 + \left(y_{01} - y_0\right)^2 + \left(z_{01} - z_0\right)^2}.$$

In these equations, x_0, y_0, z_0 (or x_1, y_1, z_1) are the Cartesian coordinates of the spacecraft in the coordinate system $M_0X_0Y_0Z_0$ ($M_1X_1Y_1Z_1$), whose origin coincides with the Earth center of mass M_0 (or Moon center of mass M_1), and coordinate axes are parallel to those of the inertial reference frame; r_0 is the magnitude of the radius vector $\boldsymbol{r}_0$ of the spacecraft drawn from point M_0, i.e., the distance from the Moon's center of mass to the spacecraft; r_1 is the magnitude of the radius vector $\boldsymbol{r}_1$ of the spacecraft drawn from point M_1, i.e., the distance from the Moon's center of mass to the spacecraft; r_{01} is the magnitude of the radius vector $\boldsymbol{r}_{01}$ of the Moon drawn from point M_0, i.e., the distance between the Earth and Moon centers of mass**;** f is the gravitational constant; m_0, m_1 are the masses of the Earth and Moon, respectively; x_{01}, y_{01}, z_{01} are the Cartesian coordinates of the Moon center of mass in the $M_0X_0Y_0Z_0$ coordinate system. Furthermore, $\boldsymbol{u}_1$ is a quaternion whose components u_{10}, u_{11}, u_{12}, u_{13} are *KS*-variables and they characterize the spacecraft position in the Moon-centered coordinate system $M_1X_1Y_1Z_1$; $\boldsymbol{s}_1$ is a quaternion with components s_{10}, s_{11}, s_{12}, s_{13}, characterizing the spacecraft velocity in the $M_1X_1Y_1Z_1$ coordinate system; t is the time, τ is the fictitious time (independent variable); h_1^* is a spacecraft Kepler energy (additional variable); $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ are Hamilton imaginary units in vector form; $\boldsymbol{p}$ is the bounded in magnitude control vector (perturbing acceleration of the spacecraft center of mass), $|\boldsymbol{p}| \le p_{\max}$.

The radius vector $\boldsymbol{r}_1$, describing the spacecraft position in the $M_1X_1Y_1Z_1$ coordinate system, its magnitude and the velocity vector $\boldsymbol{v}_1$ of the spacecraft motion in this reference frame are determined from the variables $\boldsymbol{u}_1$ and $d\boldsymbol{u}_1/d\tau = \boldsymbol{s}_1$ using quaternion relations [3] (here and further, the wave above quaternion means the conjugate quaternion).

$$\boldsymbol{r}_1 = \widetilde{\boldsymbol{u}}_1 \circ \boldsymbol{i} \circ \boldsymbol{u}_1, \qquad r_1 = \widetilde{\boldsymbol{u}}_1 \circ \boldsymbol{u}_1 = u_{10}^2 + u_{11}^2 + u_{12}^2 + u_{13}^2,$$

$$\boldsymbol{v}_1 = \frac{d\boldsymbol{r}_1}{dt} = 2\widetilde{\boldsymbol{u}}_1 \circ \boldsymbol{i} \circ \frac{d\boldsymbol{u}_1}{dt} = 2r_1^{-1}\widetilde{\boldsymbol{u}}_1 \circ \boldsymbol{i} \circ \frac{d\boldsymbol{u}_1}{d\tau} = 2r_1^{-1}\widetilde{\boldsymbol{u}}_1 \circ \boldsymbol{i} \circ \boldsymbol{s}_1.$$

For the quantities $dr_1/d\tau$, $(d\boldsymbol{r}_1/d\tau)\cdot\boldsymbol{r}_{01}$ and $(d\boldsymbol{r}_1/d\tau)\cdot\boldsymbol{p}$ appearing in equation (3), the following relations hold

$$\frac{dr_1}{d\tau} = 2(u_{10}s_{10} + u_{11}s_{11} + u_{12}s_{12} + u_{13}s_{13}),$$

$$\begin{aligned}\frac{d\boldsymbol{r}_1}{d\tau}\cdot\boldsymbol{r}_{01} = {} & 2x_{01}(u_{10}s_{10} + u_{11}s_{11} - u_{12}s_{12} - u_{13}s_{13}) + \\ & + 2y_{01}(u_{12}s_{11} + u_{11}s_{12} - u_{13}s_{10} - u_{10}s_{13}) + \\ & + 2z_{01}(u_{13}s_{11} + u_{11}s_{13} + u_{12}s_{12} + u_{13}s_{13}),\end{aligned}$$

$$\begin{aligned}\frac{d\boldsymbol{r}_1}{d\tau}\cdot\boldsymbol{p} = {} & 2p_1(u_{10}s_{10} + u_{11}s_{11} - u_{12}s_{12} - u_{13}s_{13}) + \\ & + 2p_2(u_{12}s_{11} + u_{11}s_{12} - u_{13}s_{10} - u_{10}s_{13}) + \\ & + 2p_3(u_{13}s_{11} + u_{11}s_{13} + u_{12}s_{12} + u_{13}s_{13}).\end{aligned}$$

The use of differential equations (1)-(5) of the perturbed spatial restricted three-body problem enables the development of regular analytical and numerical methods for studying the motion of a negligibly small mass body in the vicinity of two other bodies with finite masses, as well as the construction of a regular algorithm for integrating these equations, where equations (1)-(5) are employed for investigation of the spacecraft motion near the Moon (when the distances and satisfy the inequality $m_0 r_1^2 < m_1 r_0^2$).

We note that M.Yu. Loginov and Yu.N. Chelnokov previously conducted [4, 5] a comparative study of the numerical integration accuracy between the classical Newtonian differential equations of the spatial restricted three-body problem (Earth, Moon, and spacecraft) in Cartesian coordinates and the regular quaternion differential equations of this problem in four-dimensional Kustaanheimo-Stiefel variables, developed by Chelnokov [1].

The regular quaternion equations in *KS*-variables demonstrated significantly higher accuracy compared to the equations in Cartesian coordinates: for circular orbits, the accuracy was improved by two orders of magnitude; for perturbed elliptical orbits with medium eccentricity – by four orders of magnitude; and for perturbed elliptical orbits with high eccentricity – by seven orders of magnitude.

The restricted three-body problem constitutes [6] a limiting case of the unrestricted three-body problem. This formulation has found widespread application in both classical celestial mechanics (such as lunar motion theory) and spacecraft flight mechanics (particularly in the lunar mission design problem).

We consider a special case of the restricted three-body problem – the circular restricted three-body problem. Let us assume that material point M_0 represents the Earth, around which material point M_1 (the Moon) moves along a circular orbit according to Kepler laws. We will also consider that the plane of the Moon circular orbit coincides with the coordinate plane $M_0X_0Y_0$. Then, for the Cartesian coordinates x_{01}, y_{01}, z_{01} of the Moon center of mass in the coordinate system $M_0X_0Y_0Z_0$, which appear in the regular quaternion equations (1)-(5), the following relations hold:

$$x_{01} = a\cos(nt), \qquad y_{01} = a\sin(nt), \qquad z_{01} = 0,$$
$$a = 3.85\cdot 10^8\,\text{m}, \qquad n = \sqrt{f(m_0 + m_1)/a^3}.$$

In this formulation, it is assumed that the axis M_0X_0 passes through the initial position of the Moon (this moment in time is taken as the initial epoch for time reckoning).

Let us assume that it is necessary to transfer the spacecraft, whose center of mass motion is described by equations (1)-(5), from a given initial state

$$\begin{gathered}\tau = 0, \qquad t = 0, \\ \boldsymbol{u}_1(0) = \boldsymbol{u}_1^0, \qquad \boldsymbol{s}_1(0) = \boldsymbol{s}_1^0, \qquad h_1^*(0) = h_1^0\end{gathered} \tag{6}$$

to the final state

$$\begin{gathered}\tau = \tau^*, \qquad \left\|\widetilde{\boldsymbol{u}}_1(\tau^*)\circ\boldsymbol{i}\circ\boldsymbol{u}_1(\tau^*)\right\| = (R_{\text{Moon}} + R_{\text{SC}})^2, \\ \left(\widetilde{\boldsymbol{u}}_1(\tau^*)\circ\boldsymbol{i}\circ\boldsymbol{u}_1(\tau^*)\right)\cdot\left(\widetilde{\boldsymbol{u}}_1(\tau^*)\circ\boldsymbol{i}\circ\boldsymbol{s}_1(\tau^*)\right) = 0, \\ \left\|2\widetilde{\boldsymbol{u}}_1(\tau^*)\circ\boldsymbol{i}\circ\boldsymbol{s}_1(\tau^*)\right\| = fm_1r_1^2/(R_{\text{Moon}} + R_{\text{SC}}).\end{gathered} \tag{7}$$

in minimum time; R_{Moon} is the mean radius of the Moon, R_{SC} is the specified radius of the spacecraft circular orbit.

The first condition in (7) specifies the radius of the spacecraft circular orbit; the second condition corresponds to the requirement that the spacecraft velocity vector during its motion along a circular orbit around the Moon must be orthogonal to its radius vector relative to the Moon. The third condition specifies the magnitude of the spacecraft's velocity (it equals to the first cosmic velocity for the Moon).

Also, we consider the case when at the final time moment, in addition to conditions (7), the orientation of the spacecraft orbital plane around the Moon was specified, for example, when it coincides with the plane of the Moon revolution around the Earth.

In the formulated problem, the initial values of the spacecraft phase coordinates $\boldsymbol{u}_1$, $\boldsymbol{s}_1$, h_1^*, t are specified, along with the radius R_{SC} of the spacecraft circular orbit at the end of motion.

The final values of the phase coordinates belong to manifold (7). The final value of fictitious time τ^* is not fixed and it must be determined as part of the solution. Therefore, this is a problem with a free right endpoint.

II. Transition to the dimensionless variables

For the numerical solution of the formulated optimal control boundary value problem, its equations were expressed in dimensionless variables. The dimensionless variables $\boldsymbol{u}_1^{dl}$, $\boldsymbol{s}_1^{dl}$, h_1^{dl}, t^{dl}, τ^{dl} and control components p_k^{dl} are related

to the dimensional variables $\boldsymbol{u}_1$, $\boldsymbol{s}_1$, h_1^*, t, τ and controls p_k $\left(k=\overline{1,3}\right)$ through the following relations:

$$\boldsymbol{u}_1=\sqrt{R}\cdot\boldsymbol{u}_1^{dl},\qquad \boldsymbol{s}_1=\sqrt{fm_0}\cdot\boldsymbol{s}_1^{dl},\qquad h_1^*=\sqrt{fm_0/R}\cdot h_1^{dl},$$
$$t=R\sqrt{R/(fm_0)}\cdot t^{dl},\qquad \tau=\sqrt{R/(fm_0)}\cdot\tau^{dl},\qquad p_k=p_{\max}\cdot p_k^{dl},$$

where R is the characteristic distance.

We note that when we use dimensionless variables in the equations for phase and conjugate variables, a characteristic dimensionless parameter $N^b=p_{\max}\cdot R^2/(fm_0)$. appears in these equations.

Thus, the system of phase equations in dimensionless variables takes the following form (here and subsequently, the superscripts for dimensionless variables are omitted):

$$\begin{gathered}\frac{d\boldsymbol{u}_1}{d\tau}=\boldsymbol{s}_1,\\ \frac{d\boldsymbol{s}_1}{d\tau}=\frac{1}{2}h_1^*\boldsymbol{u}_1-\frac{1}{2}r_1\left\{r_1r_0^{-3}\boldsymbol{u}_1+\right.\\ \left.+\boldsymbol{i}\circ\boldsymbol{u}_1\circ\left[\left(r_{01}^{-3}-r_0^{-3}\right)\boldsymbol{R}_{01}\right]\right\}-N^b\left(r_1/2\right)\boldsymbol{i}\circ\boldsymbol{u}_1\circ\boldsymbol{p},\\ \frac{dh_1^*}{d\tau}=-r_0^{-3}r_1\frac{dr_1}{d\tau}+\left(r_{01}^{-3}-r_0^{-3}\right)\left(\frac{d\boldsymbol{r}_1}{d\tau}\cdot\boldsymbol{r}_{01}\right)+\frac{d\boldsymbol{r}_1}{d\tau}\cdot\boldsymbol{p},\\ \frac{dt}{d\tau}=r_1.\end{gathered}\tag{8}$$

Here, the Cartesian coordinates of the Moon's center of mass are given by:

$$x_{01}=(a/R)\cos(nt),\qquad y_{01}=(a/R)\sin(nt),\qquad z_{01}=0,$$

where the dimensionless oscillation frequency is: $n=\sqrt{(m_0+m_1)/m_0\cdot(R/a)^3}$.

We should note that with this choice of scaling factors, the ratio of distances a/R will also appear in the conjugate equations and in the Hamiltonian-Pontryagin function.

III. The Law of Optimal Control

We have solved this problem using the maximum principle of L.S. Pontryagin. Let us introduce additional dimensionless variables $\psi_1,\ldots,\psi_{10}$, conjugate to the dimensionless components of the quaternions $\boldsymbol{u}_1$, $\boldsymbol{s}_1$, Kepler energy h_1 and time t. A system of tenth-order conjugate equations was constructed for the variables $\psi_1,\ldots,\psi_{10}$. These equations are not presented here due to their cumbersome form.

The time-optimal control law (i.e., the control law satisfying the necessary optimality conditions) is derived from the maximization conditions of the Hamiltonian-Pontryagin function with respect to variables p_k and takes the form:

$$\boldsymbol{p}_{opt}=\frac{\left(\frac{\partial H}{\partial p_1},\ \frac{\partial H}{\partial p_2},\ \frac{\partial H}{\partial p_3}\right)}{\sqrt{\left(\frac{\partial H}{\partial p_1}\right)^2+\left(\frac{\partial H}{\partial p_2}\right)^2+\left(\frac{\partial H}{\partial p_3}\right)^2}},\tag{9}$$

where

$$\begin{aligned}\frac{\partial H}{\partial p_1}&=\frac{N^b}{2}\cdot\left[4\cdot\left(s_{10}\cdot u_{10}+s_{11}\cdot u_{11}-s_{12}\cdot u_{12}-s_{13}\cdot u_{13}\right)\cdot\psi_9+\right.\\&\left.+\|\boldsymbol{u}_1\|\cdot\left(u_{10}\cdot\psi_5+u_{11}\cdot\psi_6-u_{12}\cdot\psi_7-u_{13}\cdot\psi_8\right)\right],\\ \frac{\partial H}{\partial p_2}&=\frac{N^b}{2}\cdot\left[-4\cdot\left(s_{10}\cdot u_{13}-s_{11}\cdot u_{12}-s_{12}\cdot u_{11}+s_{13}\cdot u_{10}\right)\cdot\psi_9+\right.\\&\left.+\|\boldsymbol{u}_1\|\cdot\left(-u_{10}\cdot\psi_8+u_{11}\cdot\psi_7+u_{12}\cdot\psi_6-u_{13}\cdot\psi_5\right)\right],\\ \frac{\partial H}{\partial p_3}&=\frac{N^b}{2}\cdot\left[4\cdot\left(s_{10}\cdot u_{12}+s_{11}\cdot u_{13}+s_{12}\cdot u_{10}+s_{13}\cdot u_{11}\right)\cdot\psi_9+\right.\\&\left.+\|\boldsymbol{u}_1\|\cdot\left(u_{10}\cdot\psi_7+u_{11}\cdot\psi_8+u_{12}\cdot\psi_5+u_{13}\cdot\psi_6\right)\right].\end{aligned}$$

Transversality conditions that do not contain indefinite Lagrange multipliers were also obtained.

Thus, the time-optimal spacecraft transfer problem to lunar satellite orbit has been reduced in the considered formulation to a boundary value problem with free right endpoint, described by a system of twentieth-order nonlinear differential equations and thirteen boundary conditions (6), (7), which must be supplemented with seven transversality conditions. Also, the Hamiltonian is equal to zero at the final time.

IV. Examples of Numerical Solution of the Problem

The algorithm proposed by us for the numerical solution of the problem is based on the Levenberg-Marquardt algorithm, which is a combination of the modified Newton method and the gradient descent method [7]. Also, the Runge-Kutta method was used for integration of phase and conjugate differential equations.

Let the initial values of the coordinates and velocities of the spacecraft center of mass are given as follows (they correspond to one of the lunar capture orbit families presented in [8]):

– in the geocentric coordinate system $M_0X_0Y_0Z_0$

$$\begin{aligned}x_0^0&=-108150050.8\text{ m},& y_0^0&=-258738627.0\text{ m},\\ z_0^0&=-107466936.4\text{ m};& v_{x0}^0&=834.134525\text{ m/s},\\ v_{y0}^0&=-526.477398\text{ m/s},& v_{z0}^0&=178.491815\text{ m/s};\end{aligned}$$

– in the selenocentric coordinate system $M_1X_1Y_1Z_1$

$$\begin{aligned}x_1^0&=-493150050.8\text{ m},& y_1^0&=-258738627.0\text{ m},\\ z_1^0&=-107466936.4\text{ m};& v_{x1}^0&=834.134525\text{ m/s},\\ v_{y1}^0&=-1550.208181\text{ m/s},& v_{z1}^0&=178.491815\text{ m/s}.\end{aligned}$$

The characteristic distance $R=37000000\text{ m}\approx$ a/10 was chosen as the scaling factor for length, following [9].

We have constructed optimal transfer trajectories for spacecraft to lunar satellite orbits with various circular orbit radii lying in the plane of the Moon revolution around Earth. When the final orbit radius decreases from 90000 km to 65000 km, the transfer time increases by a factor of 1.67.

All components of the optimal control vector have local extrema at the beginning of the motion, while its second and third components oscillate approximately according to a harmonic law.

The points of local extrema for the components of the optimal control vector correspond to the points of local extrema for the vector components of quaternion $\boldsymbol{s}_1$, which characterizes the spacecraft velocity in the selenocentric coordinate system. Simultaneously, the local extremum point for component s_{13} corresponds to the local extremum point for the Kepler energy.

Each component of quaternion $\boldsymbol{u}_1$, characterizing the spacecraft position in the selenocentric coordinate system, approaches an asymptote at the end of the motion.

We should note non-uniqueness of the numerical solution of the boundary value problem of optimal spacecraft transfer to lunar satellite orbit, associated with the nonlinearity of the differential equations of the problem. Under the same boundary conditions, different solutions for the laws of motion, control and behavior of conjugate variables were obtained in the formulation of the boundary value optimal control problem. The one solution that corresponds to the minimum-time transfer was chosen.

V. Conclusion

To solve the time-optimal transfer problem of a spacecraft to lunar orbit using finite-thrust propulsion, we employed the Pontryagin maximum principle and novel regular quaternion differential equations of orbital motion in four-dimensional Kustaanheimo-Stiefel variables, developed by Chelnokov Yu.N. within the framework of the perturbed spatial restricted three-body problem (Earth-Moon-spacecraft system). These equations offer qualitative advantages over classical models. Their application eliminates the singularity inherent in classical Cartesian equations of spacecraft motion that occurs during close approaches to gravitational bodies, while simultaneously improving convergence in numerical solutions of optimal control boundary value problems for low-thrust spacecraft trajectories. The latter is due in particular from the fact that when thrust is small compared to gravitational forces, the regular quaternion equations for phase and conjugate variables, formulated using Kustaanheimo-Stiefel variables, become nearly linear. We have given examples of numerical solutions for this optimal control problem involving transfers to lunar orbits of varying radii. The characteristic properties and regularities of its solutions have been revealed. The combination of two methods for solving boundary value problems (modified Newton method and gradient descent method) has made it possible to increase the accuracy of the numerical solution of the corresponding boundary value problem of optimal control.

References

[1] Yu.N. Chelnokov, “Quaternion regularization of the equations of the perturbed spatial restricted three-body problem: I,” Mechanics of Solids, 2017, vol. 52, no. 6, p. 613-639.

[2] Yu.N. Chelnokov, “Quaternion regularization of the equations of the perturbed spatial restricted three-body problem: II,” Mechanics of Solids, 2018, vol. 53, no. 6, p. 633-650.

[3] Yu.N. Chelnokov, “Regular equations of the three-dimensional two-body problem,” Mechanics of Solids, 1984, vol. 19, no. 1, p. 1-7.

[4] Yu.N. Chelnokov, M.Yu. Loginov, “New quaternion models of spaceflight regular mechanics and their applications in the problems of motion prediction for cosmic bodies and in inertial navigation in space,” 28th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2021, 4 p.

[5] Yu.N. Chelnokov, Ya.G. Sapunkov, M.Yu. Loginov, A.F. Schekutev, “Forecast and correction of the orbital motion of the space vehicle using regular quaternion equations and their solutions in the Kustaanheimo–Stiefels variables and isochronic derivatives, Prikladnaya matematika i mekhanika, 2023, vol. 87, no. 2, p. 124-156.

[6] V.K. Abalakin, E.P. Aksenov, E.A. Grebenikov, V.G. Demin, Yu.A. Ryabov, “Spravochnoe rukovodstvo po nebesnoi mekhanike i astrodinamike (A reference guide to celestial mechanics and astrodynamics.)”, Nauka, Moscow, 1976.

[7] N.N. Moiseev, “Chislennye metody v teorii optimal'nyh sistem (Numerical methods in the theory of optimal systems),” Nauka, Moscow, 1971.

[8] V.V. Ivashkin, V.G. Petukhov. “Traektorii pereleta s maloi tiagoi mezhdu orbitami sputnikov Zemli i Luny pri ispol'zovanii orbity zakhvata Lunoi”, Preprint № 81, IPM im. M.V. Keldysha RAN. Moscow, 2008, 32 p.

[9] I.A. Pankratov, Ya.G. Sapunkov, Yu.N. Chelnokov, “Quaternion models and algorithms for solving the general problem of optimal reorientation of spacecraft orbit,” Izvestiya of Saratov University. Mathematics. Mechanics. Informatics, 2020, vol. 20, iss. 1, pp. 93-104.

The Kinematic Models of SINS and its Errors on the $\mathbb{SE}_2(3)$ Group: Analysis and Comparison

Ke Fang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
fangke@seu.edu.cn

Tijing Cai
School of Instrument Science and Engineering
Southeast University
Nanjing, China
caitij@seu.edu.cn

***Abstract*—This paper derives the left- and right-invariant error models on the $\mathbb{SE}_2(3)$ group for SINS under Earth-Centered Inertial (ECI, i-frame), Earth-Centered Earth-Fixed (ECEF, e-frame), and local geographic (LGEO, n-frame) coordinate systems. These models are analyzed and compared from both theoretical derivation and turntable experiment to reveal the relationships between left- and right-invariant errors, reference coordinate system selection, and the definition of coordinate system inconsistency errors, as well as the pros and cons of different $\mathbb{SE}_2(3)$ models. Results show the impact of navigation frame selection and error model formulation on attitude estimation performance, demonstrating that the ECI-$\mathbb{SE}_2(3)$-L and ECEF-$\mathbb{SE}_2(3)$-L models generally exhibit better robustness and convergence properties under large initial misalignment and high-dynamic conditions.**

***Index Terms*—Attitude determination, Coordinate system inconsistency errors, Integrated navigation systems, Lie group and Lie algebra, SINS kinematics.**

I. Introduction

Strapdown Inertial Navigation System (SINS) provides a passive, covert, and reliable navigation solution but suffers from error divergence over time. Thus, it is commonly integrated with other navigation systems such as Global Navigation Satellite System (GNSS) and Doppler Velocity Log (DVL). Lacking direct attitude observations, precise attitude estimation remains challenging, particularly in large misalignment angles and high dynamics.

In conventional SINS kinematics, attitude and its errors are represented in the $\mathbb{SO}(3)$ space, whereas velocity, position, and their errors lie in the vector space, resulting in spatiotemporal inconsistency between angular motion and linear motion [1], [2]. Moreover, as the n-frame is commonly selected as the navigation frame, the coupling between angular and linear motions induced by Coriolis effect renders the SINS error kinematics inherently non-globally independent [3], [4]. Specifically, the propagation of SINS errors is intrinsically dependent on global navigation parameters, while the feedback of these errors, in turn, influences the global parameters. This interdependence leads to severe degradation in navigation accuracy and filtering model precision when significant errors exist. Barrau and Bonnabel [5], [6] introduced the Invariant Extended Kalman Filter (IEKF) to solve the problem, and Zhu et al. [7] derived an $\mathbb{SE}(3)$-based SINS error model with quaternion attitude representation. Chang et al. [8], [9] developed an $\mathbb{SE}_2(3)$-based SINS error model with e-frame selected as the navigation frame, and then extended it to a globally independent kinematic model in n-frame. Furthermore, Fang et al. [10] incorporated the error kinematic models of the b-frame and the n-frame within the $\mathbb{SE}(3)$ group structure relative to the i-frame. These models effectively mitigate inconsistency errors in error propagation. However, under different navigation coordinate systems, the performance of SINS error kinematics on the matrix Lie groups varies, and distinctions also exist between the left- and right-invariant error models in the group spaces. This paper derives the SINS error kinematic models on the $\mathbb{SE}_2(3)$ group, for different navigation coordinate systems, and develops the corresponding SINS/GNSS integrated navigation measurement models. Theoretical derivation and experimental results explore the relationships among navigation frame selection, the coordinate system inconsistency definition, and the left- and right-invariant error models on the $\mathbb{SE}_2(3)$ group.

II. The Kinematic Models of SINS and its Errors on the $\mathbb{SE}_2(3)$ Group

A. The model with the i-frame as the navigation frame (ECI-$\mathbb{SE}_2(3)$ model)

When the i-frame is selected as the navigation frame, the $\mathbb{SE}_2(3)$ group element representing the global navigation state of the SINS is expressed as

$$\boldsymbol{T}_b^i = \begin{bmatrix} \boldsymbol{C}_b^i & \boldsymbol{v}_{ib}^i & \boldsymbol{p}_{ib}^i \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \in \mathbb{SE}_2(3) \tag{1}$$

where, $\boldsymbol{C}_b^i$ is the attitude matrix of the b-frame relative to the i-frame; $\boldsymbol{v}_{ib}^i$ denotes the velocity vector of the b-frame with respect to the i-frame, projected onto the i-frame; and $\boldsymbol{p}_{ib}^i$ represents the position vector of the b-frame with respect to the i-frame, projected onto the i-frame.

The kinematic model of $\boldsymbol{T}_b^i$ is

$$\begin{cases} \dot{\boldsymbol{C}}_b^i = \boldsymbol{C}_b^i \left(\boldsymbol{\omega}_{ib}^b \times\right) \\ \dot{\boldsymbol{v}}_{ib}^i = \boldsymbol{f}^i + \boldsymbol{G}^i \\ \dot{\boldsymbol{p}}_{ib}^i = \boldsymbol{v}_{ib}^i \end{cases} \tag{2}$$

where, $\boldsymbol{\omega}_{ib}^b$ is the angular velocity of the b-frame relative to the i-frame, projected onto the b-frame, where the symbol "$\times$" denotes its skew-symmetric matrix; $\boldsymbol{f}^i$ represents the specific force vector projected onto the i-frame; $\boldsymbol{G}^i$ denotes the gravitational vector projected onto the i-frame.

1) Left-invariant error model (ECI-$\mathbb{SE}_2(3)$-L model)

Since the i-frame is a quasi-static reference, the left-invariant error of the SINS is expressed as

$$\begin{aligned} \boldsymbol{\delta}_{ib}^l = \left(\hat{\boldsymbol{T}}_b^i\right)^{-1} \boldsymbol{T}_b^i &= \begin{bmatrix} \boldsymbol{C}_b^{b'} & \delta\boldsymbol{v}_{ib}^l & \delta\boldsymbol{p}_{ib}^l \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{C}_b^{b'} & \boldsymbol{C}_i^{b'}\left(\boldsymbol{v}_{ib}^i - \hat{\boldsymbol{v}}_{ib}^i\right) & \boldsymbol{C}_i^{b'}\left(\boldsymbol{p}_{ib}^i - \hat{\boldsymbol{p}}_{ib}^i\right) \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \end{aligned} \tag{3}$$

where, b' represents the calculated body frame, and $\boldsymbol{C}_i^{b'}$ denotes the calculated attitude matrix of the b-frame relative to the i-frame; the parameters marked with a "$\hat{}$" indicates the calculated values; $\boldsymbol{C}_b^{b'}$ represents the attitude error of the ECI-$\mathbb{SE}_2(3)$-L model; $\delta\boldsymbol{v}_{ib}^l$ is the invariant velocity error of this model; $\delta\boldsymbol{p}_{ib}^l$ is the invariant position error of this model.

The error kinematics of the ECI-$\mathbb{SE}_2(3)$-L model is

$$\dot{\boldsymbol{\delta}}_{ib}^l = \boldsymbol{\delta}_{ib}^l \boldsymbol{\Xi}_{ib}^l = \begin{bmatrix} \boldsymbol{C}_b^{b'} & \delta\boldsymbol{v}_{ib}^l & \delta\boldsymbol{p}_{ib}^l \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \begin{bmatrix} \left(\boldsymbol{\omega}_{b'b}^b \times\right) & \Delta\boldsymbol{v}_{ib}^l & \Delta\boldsymbol{p}_{ib}^l \\ \mathbf{0} & 0 & 0 \\ \mathbf{0} & 0 & 0 \end{bmatrix} \tag{4}$$

$$\boldsymbol{\omega}_{b'b}^b = \left(\boldsymbol{I} - \boldsymbol{C}_{b'}^b\right) \hat{\boldsymbol{\omega}}_{ib}^b - \delta\boldsymbol{\omega}_{ib}^b \tag{5}$$

$$\Delta\boldsymbol{v}_{ib}^l = -\boldsymbol{C}_{b'}^b \left(\hat{\boldsymbol{\omega}}_{ib}^b \times\right) \delta\boldsymbol{v}_{ib}^l + \left(\boldsymbol{I} - \boldsymbol{C}_{b'}^b\right) \hat{\boldsymbol{f}}^b - \delta\boldsymbol{f}^b - \boldsymbol{C}_{b'}^b \boldsymbol{C}_i^{b'} \delta\boldsymbol{G}^i \tag{6}$$

$$\Delta\boldsymbol{p}_{ib}^l = -\boldsymbol{C}_{b'}^b \left(\hat{\boldsymbol{\omega}}_{ib}^b \times\right) \delta\boldsymbol{p}_{ib}^l + \boldsymbol{C}_{b'}^b \delta\boldsymbol{v}_{ib}^l \tag{7}$$

where, $\boldsymbol{\Xi}_{ib}^l$ represents the Lie algebra of $\boldsymbol{\delta}_{ib}^l$ on the local tangent plane of the $\mathbb{SE}_2(3)$ group space; $\hat{\boldsymbol{\omega}}_{ib}^b$ denotes the angular velocity of the b-frame relative to the i-frame in the b-frame, which is measured directly by the gyroscope, with $\delta\boldsymbol{\omega}_{ib}^b$ indicating its error; $\hat{\boldsymbol{f}}^b$ is the specific force of the b-frame relative to the i-frame in the b-frame, measured directly by the accelerometer, with $\delta\boldsymbol{f}^b$ representing its error; $\delta\boldsymbol{G}^i$ stands for the gravitational acceleration error.

2) Right-invariant error model (ECI-$\mathbb{SE}_2(3)$-R model)

Similarly, the right-invariant error is defined as

$$\begin{aligned} \boldsymbol{\delta}_{ib}^r = \boldsymbol{T}_i^b \left(\hat{\boldsymbol{T}}_i^b\right)^{-1} &= \begin{bmatrix} \boldsymbol{C}_{b'}^b & \delta\boldsymbol{v}_{ib}^r & \delta\boldsymbol{p}_{ib}^r \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{C}_{b'}^b & \boldsymbol{v}_{ib}^b - \boldsymbol{C}_{b'}^b \hat{\boldsymbol{v}}_{ib}^b & \boldsymbol{p}_{ib}^b - \boldsymbol{C}_{b'}^b \hat{\boldsymbol{p}}_{ib}^b \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \end{aligned} \tag{8}$$

where, $\boldsymbol{C}_{b'}^b$ is the attitude error of the ECI-$\mathbb{SE}_2(3)$-R model; $\delta\boldsymbol{v}_{ib}^r$ denotes the invariant velocity error of this model; $\delta\boldsymbol{p}_{ib}^r$ represents the invariant position error of the model.

This work is supported by National Key R&D Program of China under Grant No.2023YFC2907004.

The error kinematics of the ECI-$\mathbb{SE}_2(3)$-R model is

$$\dot{\boldsymbol{\delta}}^r_{ib}=\boldsymbol{\delta}^r_{ib}\boldsymbol{\Xi}^r_{ib}=\begin{bmatrix}\boldsymbol{C}^b_{b'} & \delta\boldsymbol{v}^r_{ib} & \delta\boldsymbol{p}^r_{ib}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\begin{bmatrix}\left(\boldsymbol{\omega}^{b'}_{bb'}\times\right) & \Delta\boldsymbol{v}^r_{ib} & \Delta\boldsymbol{p}^r_{ib}\\ \boldsymbol{0} & 0 & 0\\ \boldsymbol{0} & 0 & 0\end{bmatrix} \tag{9}$$

$$\boldsymbol{\omega}^{b'}_{bb'}=\left(\boldsymbol{I}-\boldsymbol{C}^{b'}_{b}\right)\hat{\boldsymbol{\omega}}^b_{ib}+\boldsymbol{C}^{b'}_{b}\delta\boldsymbol{\omega}^b_{ib} \tag{10}$$

$$\Delta\boldsymbol{v}^r_{ib}=-\boldsymbol{C}^{b'}_{b}\left(\hat{\boldsymbol{\omega}}^b_{ib}\times\right)\delta\boldsymbol{v}^r_{ib}+\left(\boldsymbol{I}-\boldsymbol{C}^{b'}_{b}\right)\hat{\boldsymbol{f}}^b-\boldsymbol{C}^{b'}_{b}\delta\boldsymbol{f}^b+\boldsymbol{C}^{b'}_{i}\delta\boldsymbol{G}^i \tag{11}$$

$$\Delta\boldsymbol{p}^r_{ib}=-\boldsymbol{C}^{b'}_{b}\left(\hat{\boldsymbol{\omega}}^b_{ib}\times\right)\delta\boldsymbol{p}^r_{ib}+\boldsymbol{C}^{b'}_{b}\delta\boldsymbol{v}^r_{ib} \tag{12}$$

where, $\boldsymbol{\Xi}^r_{ib}$ represents the Lie algebra of $\boldsymbol{\delta}^r_{ib}$ on the local tangent plane of the $\mathbb{SE}_2(3)$ group space.

Remark 1: A comparative analysis reveals that the ECI-$\mathbb{SE}_2(3)$-L and ECI-$\mathbb{SE}_2(3)$-R models exhibit strict structural symmetry, stemming from their common adoption of the quasi-inertial and precisely defined i-frame as the navigation frame. The inconsistency errors are characterized by the discrepancies between the b-frame and b'-frame. The primary distinction between the two models lies in the reference frame for velocity and position errors. In the ECI-$\mathbb{SE}_2(3)$-L model, linear motion errors are expressed relative to the b'-frame obtained through navigation computation, whereas in the ECI-$\mathbb{SE}_2(3)$-R model, these errors are formulated with respect to the ideal b-frame, which has been compensated for estimated inconsistency errors. Subsequent analysis will investigate the impact of this distinction on integrated navigation filtering performance under large misalignment angles.

B. The model with the e-frame as the navigation frame (ECEF-$\mathbb{SE}_2(3)$ model)

Relative to the e-frame, the global motion state of the SINS can be formulated as

$$\boldsymbol{T}^e_b=\begin{bmatrix}\boldsymbol{C}^e_b & \boldsymbol{v}^e_{eb} & \boldsymbol{p}^e_{eb}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\in\mathbb{SE}_2(3) \tag{13}$$

where, $\boldsymbol{C}^e_b$ is the attitude matrix of the b-frame relative to the e-frame; $\boldsymbol{v}^e_{eb}$ is the velocity vector of the b-frame with respect to the e-frame (also referred to as ground velocity or relative velocity), projected onto the e-frame; $\boldsymbol{p}^e_{eb}$ is the position vector of the b-frame with respect to the e-frame projected onto the e-frame.

Its kinematic model is

$$\begin{cases}\dot{\boldsymbol{C}}^e_b=\boldsymbol{C}^e_b\left(\boldsymbol{\omega}^b_{ib}\times\right)-\left(\boldsymbol{\omega}^e_{ie}\times\right)\boldsymbol{C}^e_b\\ \dot{\boldsymbol{v}}^e_{eb}=\boldsymbol{C}^e_b\boldsymbol{f}^b-2\boldsymbol{\omega}^e_{ie}\times\boldsymbol{v}^e_{eb}+\boldsymbol{g}^e\\ \dot{\boldsymbol{p}}^e_{eb}=\boldsymbol{v}^e_{eb}\end{cases} \tag{14}$$

where, $\boldsymbol{\omega}^e_{ie}$ denotes the angular velocity of Earth's rotation, expressed in the e-frame; $\boldsymbol{g}^e$ refers to the gravity vector projected onto the e-frame, which not only contains gravitational acceleration, but also centrifugal acceleration.

1) Left-invariant error model (ECEF-$\mathbb{SE}_2(3)$-L model)

The left-invariant error of $\boldsymbol{T}^e_b$ is defined as

$$\begin{aligned}\boldsymbol{\delta}^l_{eb}&=\left(\hat{\boldsymbol{T}}^e_b\right)^{-1}\boldsymbol{T}^e_b=\begin{bmatrix}\boldsymbol{C}^{b'}_b & \delta\boldsymbol{v}^l_{eb} & \delta\boldsymbol{p}^l_{eb}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\\&=\begin{bmatrix}\boldsymbol{C}^{b'}_b & \boldsymbol{C}^{b'}_e\left(\boldsymbol{v}^e_{eb}-\hat{\boldsymbol{v}}^e_{eb}\right) & \boldsymbol{C}^{b'}_e\left(\boldsymbol{p}^e_{eb}-\hat{\boldsymbol{p}}^e_{eb}\right)\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\end{aligned} \tag{15}$$

where, $\boldsymbol{C}^{b'}_e$ is the attitude matrix of the computational b-frame relative to the e-frame; $\boldsymbol{C}^{b'}_b$ is the attitude error of the ECEF-$\mathbb{SE}_2(3)$-L model; $\delta\boldsymbol{v}^l_{eb}$ denotes the velocity error of the model; $\delta\boldsymbol{p}^l_{eb}$ is the position error of the model.

The corresponding error kinematics is

$$\dot{\boldsymbol{\delta}}^l_{eb}=\boldsymbol{\delta}^l_{eb}\boldsymbol{\Xi}^l_{eb}=\begin{bmatrix}\boldsymbol{C}^{b'}_b & \delta\boldsymbol{v}^l_{eb} & \delta\boldsymbol{p}^l_{eb}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\begin{bmatrix}\left(\boldsymbol{\omega}^b_{b'b}\times\right) & \Delta\boldsymbol{v}^l_{eb} & \Delta\boldsymbol{p}^l_{eb}\\ \boldsymbol{0} & 0 & 0\\ \boldsymbol{0} & 0 & 0\end{bmatrix} \tag{16}$$

$$\boldsymbol{\omega}^{b'}_{bb'}=\left(\boldsymbol{I}-\boldsymbol{C}^{b'}_{b}\right)\hat{\boldsymbol{\omega}}^b_{ib}+\boldsymbol{C}^{b'}_{b}\delta\boldsymbol{\omega}^b_{ib} \tag{17}$$

$$\begin{aligned}\Delta\boldsymbol{v}^l_{eb}=&-\boldsymbol{C}^b_{b'}\left[\left(\hat{\boldsymbol{\omega}}^b_{ib}+\boldsymbol{C}^{b'}_{e}\boldsymbol{\omega}^e_{ie}\right)\times\right]\delta\boldsymbol{v}^l_{eb}-\delta\boldsymbol{f}^b\\&+\left(\boldsymbol{I}-\boldsymbol{C}^b_{b'}\right)\hat{\boldsymbol{f}}^b-\boldsymbol{C}^b_{b'}\boldsymbol{C}^{b'}_{e}\delta\boldsymbol{g}^e\end{aligned} \tag{18}$$

$$\Delta\boldsymbol{p}^l_{eb}=\boldsymbol{C}^b_{b'}\left[\left(\boldsymbol{C}^{b'}_{e}\boldsymbol{\omega}^e_{ie}-\hat{\boldsymbol{\omega}}^b_{ib}\right)\times\right]\delta\boldsymbol{p}^l_{eb}+\boldsymbol{C}^b_{b'}\delta\boldsymbol{v}^l_{eb} \tag{19}$$

where, $\boldsymbol{\Xi}^l_{eb}$ is the Lie algebra of $\boldsymbol{\delta}^l_{eb}$ on the local tangent plane of the $\mathbb{SE}_2(3)$ group manifold.

Remark 2: In the above equations, $\delta\boldsymbol{g}^e$ denotes the projection of the discrepancy between the gravity value utilized in navigation calculations and the true gravity at the carrier's location in the e-frame. This discrepancy arises from three primary sources of error: computational errors induced by position inaccuracies, centrifugal acceleration errors, and gravity disturbances caused by the heterogeneous mass distribution in the local area of the Earth.

2) Right-invariant error model (ECEF-$\mathbb{SE}_2(3)$-R model)

The right-invariant error of the ECEF-$\mathbb{SE}_2(3)$ model is

$$\begin{aligned}\boldsymbol{\delta}^r_{eb}&=\boldsymbol{T}^e_b\left(\hat{\boldsymbol{T}}^e_b\right)^{-1}=\begin{bmatrix}\boldsymbol{C}^e_{e'} & \delta\boldsymbol{v}^r_{eb} & \delta\boldsymbol{p}^r_{eb}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\\&=\begin{bmatrix}\boldsymbol{C}^e_{e'} & \boldsymbol{v}^e_{eb}-\boldsymbol{C}^e_{e'}\hat{\boldsymbol{v}}^e_{eb} & \boldsymbol{p}^e_{eb}-\boldsymbol{C}^e_{e'}\hat{\boldsymbol{p}}^e_{eb}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\end{aligned} \tag{20}$$

where, $\boldsymbol{C}^e_{e'}$ represents the attitude error of the ECEF-$\mathbb{SE}_2(3)$-R model; $\delta\boldsymbol{v}^r_{eb}$ denotes its velocity error; $\delta\boldsymbol{p}^r_{eb}$ is the position error of the model.

The error kinematics is

$$\dot{\boldsymbol{\delta}}^r_{eb}=\boldsymbol{\delta}^r_{eb}\boldsymbol{\Xi}^r_{eb}=\begin{bmatrix}\boldsymbol{C}^e_{e'} & \boldsymbol{\delta}\boldsymbol{v}^r_{eb} & \boldsymbol{\delta}\boldsymbol{p}^r_{eb}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\begin{bmatrix}\left(\boldsymbol{\omega}^{e'}_{ee'}\times\right) & \Delta\boldsymbol{v}^r_{eb} & \Delta\boldsymbol{p}^r_{eb}\\ \boldsymbol{0} & 0 & 0\\ \boldsymbol{0} & 0 & 0\end{bmatrix} \tag{21}$$

$$\boldsymbol{\omega}^{e'}_{ee'}=\left(\boldsymbol{I}-\boldsymbol{C}^{e'}_{e}\right)\boldsymbol{\omega}^e_{ie}-\boldsymbol{C}^{e'}_{b}\delta\boldsymbol{\omega}^b_{ib} \tag{22}$$

$$\begin{aligned}\Delta\boldsymbol{v}^r_{eb}=&\left\{\left[\left(\boldsymbol{I}-\boldsymbol{C}^{e'}_{e}\right)\boldsymbol{\omega}^e_{ie}+\boldsymbol{C}^{e'}_{b}\delta\boldsymbol{\omega}^b_{ib}\right]\times\right\}\hat{\boldsymbol{v}}^e_{eb}-2\boldsymbol{C}^{e'}_{e}\left(\boldsymbol{\omega}^e_{ie}\times\right)\delta\boldsymbol{v}^r_{eb}\\&-\boldsymbol{C}^{e'}_{b}\delta\boldsymbol{f}^b-\left(\boldsymbol{I}-\boldsymbol{C}^{e'}_{e}\right)\hat{\boldsymbol{g}}^e-\boldsymbol{C}^{e'}_{e}\delta\boldsymbol{g}^e\end{aligned} \tag{23}$$

$$\Delta\boldsymbol{p}^r_{eb}=\boldsymbol{C}^{e'}_{e}\delta\boldsymbol{v}^r_{eb}-\left\{\left[\left(\boldsymbol{I}-\boldsymbol{C}^{e'}_{e}\right)\boldsymbol{\omega}^e_{ie}-\boldsymbol{C}^{e'}_{b}\delta\boldsymbol{\omega}^b_{ib}\right]\times\right\}\hat{\boldsymbol{p}}^e_{eb} \tag{24}$$

where, $\boldsymbol{\Xi}^r_{eb}$ is the Lie algebra of $\boldsymbol{\delta}^r_{eb}$ on the local tangent plane of the $\mathbb{SE}_2(3)$ group manifold.

Remark 3: In contrast to the previously discussed ECI-$\mathbb{SE}_2(3)$ models, there are significant differences in the errors and their propagation equations between the ECEF-$\mathbb{SE}_2(3)$-L model and the ECEF-$\mathbb{SE}_2(3)$-R model. In the ECEF-$\mathbb{SE}_2(3)$-L model, all inconsistency errors originate from the differences between the b-frame and the computational b'-frame, and the propagation of the attitude error is driven by $\boldsymbol{\omega}^b_{ib}$. In the ECEF-$\mathbb{SE}_2(3)$-R model, however, the inconsistency errors stem from the differences between the e-frame and the e'-frame, so the error propagation is driven by $\boldsymbol{\omega}^e_{ie}$.

C. The model with the n-frame as the navigation frame (LGEO-$\mathbb{SE}_2(3)$ model)

Adopting the n-frame as the navigation frame introduces greater complexity, which results from both the e-frame and n-frame being non-inertial, increasing Coriolis-driven coupling effects. Under this circumstance, the $\mathbb{SE}_2(3)$ group element describing the global motion state is

$$\boldsymbol{T}^n_b=\begin{bmatrix}\boldsymbol{C}^n_b & \boldsymbol{v}^n_{en} & \boldsymbol{p}^n_{en}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\in\mathbb{SE}_2(3) \tag{25}$$

where, $\boldsymbol{C}^n_b$ is the attitude matrix of the b-frame relative to the n-frame; $\boldsymbol{v}^n_{en}$ denotes the velocity vector of the n-frame with respect to the e-frame, projected onto the n-frame; $\boldsymbol{p}^n_{en}$ signifies the position vector of the n-frame relative to the e-frame, expressed in the n-frame.

Remark 4: The velocity vector here represents the ground-relative velocity. Since the origins of the n-frame and b-frame coincide at the center of the SINS, the velocity and position vectors used here differ solely in symbolic notation.

The kinematics of $\boldsymbol{T}^n_b$ is

$$\begin{cases}\dot{\boldsymbol{C}}^n_b=\boldsymbol{C}^n_b\left(\boldsymbol{\omega}^b_{ib}\times\right)-\left(\boldsymbol{\omega}^n_{in}\times\right)\boldsymbol{C}^n_b\\ \dot{\boldsymbol{v}}^n_{en}=\boldsymbol{C}^n_b\boldsymbol{f}^b-\left(\boldsymbol{\omega}^n_{ie}+\boldsymbol{\omega}^n_{in}\right)\times\boldsymbol{v}^n_{en}+\boldsymbol{g}^n\\ \dot{\boldsymbol{p}}^n_{en}=-\boldsymbol{\omega}^n_{en}\times\boldsymbol{p}^n_{en}+\boldsymbol{v}^n_{en}\end{cases} \tag{26}$$

where, $\boldsymbol{\omega}^n_{in}$ denotes the angular velocity of the n-frame relative to the i-frame, expressed in the n-frame; $\boldsymbol{\omega}^n_{ie}$ refers to the angular velocity of Earth's rotation, projected onto the n-frame; $\boldsymbol{\omega}^n_{en}$ represents the angular velocity of the n-frame relative to the e-frame, expressed in the n-frame; $\boldsymbol{g}^n$ is the gravity vector projected onto the n-frame.

1) Left-invariant error model (LGEO-$\mathbb{SE}_2(3)$-L model)

The left-invariant error of $\boldsymbol{T}^n_b$ is

$$\begin{aligned}\boldsymbol{\delta}^l_{nb}&=\left(\hat{\boldsymbol{T}}^n_b\right)^{-1}\boldsymbol{T}^n_b=\begin{bmatrix}\boldsymbol{C}^{b'}_b & \delta\boldsymbol{v}^l_{nb} & \delta\boldsymbol{p}^l_{nb}\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\\&=\begin{bmatrix}\boldsymbol{C}^{b'}_b & \boldsymbol{C}^{b'}_n\left(\boldsymbol{v}^n_{en}-\hat{\boldsymbol{v}}^n_{en}\right) & \boldsymbol{C}^{b'}_n\left(\boldsymbol{p}^n_{en}-\hat{\boldsymbol{p}}^n_{en}\right)\\ \boldsymbol{0} & 1 & 0\\ \boldsymbol{0} & 0 & 1\end{bmatrix}\end{aligned} \tag{27}$$

where, $\boldsymbol{C}_n^{b'}$ is the attitude matrix of the computational b'-frame relative to the n-frame; $\boldsymbol{C}_b^{b'}$ represents the attitude error of the LGEO-$\mathbb{SE}_2(3)$-L model; $\delta\boldsymbol{v}_{nb}^l$ denotes the velocity error of the model; $\delta\boldsymbol{p}_{nb}^l$ signifies the position error of the model.

The corresponding error kinematics is

$$\dot{\boldsymbol{\delta}}_{nb}^l = \boldsymbol{\delta}_{nb}^l \boldsymbol{\Xi}_{nb}^l = \begin{bmatrix} \boldsymbol{C}_b^{b'} & \delta\boldsymbol{v}_{nb}^l & \delta\boldsymbol{p}_{nb}^l \\ \boldsymbol{0} & 1 & 0 \\ \boldsymbol{0} & 0 & 1 \end{bmatrix} \begin{bmatrix} \left(\boldsymbol{\omega}_{b'b}^b\times\right) & \Delta\boldsymbol{v}_{nb}^l & \Delta\boldsymbol{p}_{nb}^l \\ \boldsymbol{0} & 0 & 0 \\ \boldsymbol{0} & 0 & 0 \end{bmatrix} \tag{28}$$

$$\boldsymbol{\omega}_{b'b}^b = \left(\boldsymbol{I} - \boldsymbol{C}_{b'}^b\right)\hat{\boldsymbol{\omega}}_{ib}^b + \boldsymbol{C}_{b'}^b\boldsymbol{C}_n^{b'}\delta\boldsymbol{\omega}_{in}^n - \delta\boldsymbol{\omega}_{ib}^b \tag{29}$$

$$\begin{aligned}\Delta\boldsymbol{v}_{nb}^l = &-\boldsymbol{C}_{b'}^b\left[\left(\boldsymbol{C}_n^{b'}\hat{\boldsymbol{\omega}}_{ie}^n + \hat{\boldsymbol{\omega}}_{ib}^b\right)\times\right]\delta\boldsymbol{v}_{nb}^l + \boldsymbol{C}_{b'}^b\boldsymbol{C}_n^{b'}\left[\left(2\delta\boldsymbol{\omega}_{ie}^n + \delta\boldsymbol{\omega}_{en}^n\right)\times\right]\hat{\boldsymbol{v}}_{en}^n \\ &+ \left(\boldsymbol{I} - \boldsymbol{C}_{b'}^b\right)\hat{\boldsymbol{f}}^b - \delta\boldsymbol{f}^b - \boldsymbol{C}_{b'}^b\boldsymbol{C}_n^{b'}\delta\boldsymbol{g}^n\end{aligned} \tag{30}$$

$$\begin{aligned}\Delta\boldsymbol{p}_{nb}^l = &\boldsymbol{C}_{b'}^b\left[\left(\boldsymbol{C}_n^{b'}\hat{\boldsymbol{\omega}}_{ie}^n - \hat{\boldsymbol{\omega}}_{ib}^b\right)\times\right]\delta\boldsymbol{p}_{nb}^l \\ &+ \boldsymbol{C}_{b'}^b\boldsymbol{C}_n^{b'}\left(\delta\boldsymbol{\omega}_{en}^n\times\right)\hat{\boldsymbol{p}}_{en}^n + \boldsymbol{C}_{b'}^b\delta\boldsymbol{v}_{nb}^l\end{aligned} \tag{31}$$

where, $\boldsymbol{\Xi}_{nb}^l$ represents the Lie algebra of $\boldsymbol{\delta}_{nb}^l$ on the local tangent plane of the $\mathbb{SE}_2(3)$ group space; $\delta\boldsymbol{\omega}_{in}^n$ is the error of $\boldsymbol{\omega}_{in}^n$; $\delta\boldsymbol{\omega}_{en}^n$ is the error of $\boldsymbol{\omega}_{en}^n$; $\delta\boldsymbol{g}^n$ signifies the gravity error in the n-frame.

Remark 5: In the LGEO-$\mathbb{SE}_2(3)$-L model, although the attitude error is represented by $\boldsymbol{C}_{b'}^b$, its error propagation is not congruent with that of the ECI-$\mathbb{SE}_2(3)$-L and ECEF-$\mathbb{SE}_2(3)$-L models derived previously. Since the n-frame is a non-inertial frame, dynamic errors are introduced as it moves with the SINS, and thus $\boldsymbol{\omega}_{b'b}^b$ incorporates the influence of $\delta\boldsymbol{\omega}_{in}^n$. Analogously, the ECEF-$\mathbb{SE}_2(3)$-L model should theoretically account for the perturbation in the Earth's rotational angular velocity, $\delta\boldsymbol{\omega}_{ie}^e$. However, because Earth's rotation is typically regarded as both precise and uniform, this term is conventionally neglected, thereby reducing the error propagation in the ECEF-$\mathbb{SE}_2(3)$-L model to degenerate to the same form as that in the ECI-$\mathbb{SE}_2(3)$-L model.

2) Right-invariant error model (LGEO-$\mathbb{SE}_2(3)$-R model)

The right-invariant error of $\boldsymbol{T}_b^n$ is defined as

$$\begin{aligned}\boldsymbol{\delta}_{nb}^r &= \left(\hat{\boldsymbol{T}}_b^n\right)^{-1}\boldsymbol{T}_b^n = \begin{bmatrix} \boldsymbol{C}_{n'}^n & \delta\boldsymbol{v}_{nb}^r & \delta\boldsymbol{p}_{nb}^r \\ \boldsymbol{0} & 1 & 0 \\ \boldsymbol{0} & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{C}_{n'}^n & \boldsymbol{v}_{en}^n - \boldsymbol{C}_{n'}^n\hat{\boldsymbol{v}}_{en}^n & \boldsymbol{p}_{en}^n - \boldsymbol{C}_{n'}^n\hat{\boldsymbol{p}}_{en}^n \\ \boldsymbol{0} & 1 & 0 \\ \boldsymbol{0} & 0 & 1 \end{bmatrix}\end{aligned} \tag{32}$$

where, $\boldsymbol{C}_{n'}^n$ represents the attitude error of the LGEO-$\mathbb{SE}_2(3)$-R model; $\delta\boldsymbol{v}_{nb}^r$ denotes its velocity error; $\delta\boldsymbol{p}_{nb}^r$ represents its position error.

The error propagation equation is given by

$$\dot{\boldsymbol{\delta}}_{nb}^r = \boldsymbol{\delta}_{nb}^r\boldsymbol{\Xi}_{nb}^r = \begin{bmatrix} \boldsymbol{C}_{n'}^n & \delta\boldsymbol{v}_{nb}^r & \delta\boldsymbol{p}_{nb}^r \\ \boldsymbol{0} & 1 & 0 \\ \boldsymbol{0} & 0 & 1 \end{bmatrix} \begin{bmatrix} \left(\boldsymbol{\omega}_{nn'}^{n'}\times\right) & \Delta\boldsymbol{v}_{nb}^r & \Delta\boldsymbol{p}_{nb}^r \\ \boldsymbol{0} & 0 & 0 \\ \boldsymbol{0} & 0 & 0 \end{bmatrix} \tag{33}$$

$$\boldsymbol{\omega}_{nn'}^{n'} = \left(\boldsymbol{I} - \boldsymbol{C}_n^{n'}\right)\hat{\boldsymbol{\omega}}_{in}^n + \boldsymbol{C}_n^{n'}\delta\boldsymbol{\omega}_{in}^n - \boldsymbol{C}_b^{n'}\delta\boldsymbol{\omega}_{ib}^b \tag{34}$$

$$\begin{aligned}\Delta\boldsymbol{v}_{nb}^r = &\left\{\left[\boldsymbol{C}_n^{n'}\delta\boldsymbol{\omega}_{ie}^n + \left(\boldsymbol{I} - \boldsymbol{C}_n^{n'}\right)\hat{\boldsymbol{\omega}}_{ie}^n + \boldsymbol{C}_b^{n'}\delta\boldsymbol{\omega}_{ib}^b\right]\times\right\}\hat{\boldsymbol{v}}_{en}^n \\ &- \boldsymbol{C}_n^{n'}\left[\left(\hat{\boldsymbol{\omega}}_{ie}^n + \hat{\boldsymbol{\omega}}_{in}^n\right)\times\right]\delta\boldsymbol{v}_{nb}^r - \boldsymbol{C}_b^{n'}\delta\boldsymbol{f}^b \\ &- \left(\boldsymbol{I} - \boldsymbol{C}_n^{n'}\right)\hat{\boldsymbol{g}}^n - \boldsymbol{C}_n^{n'}\delta\boldsymbol{g}^n\end{aligned} \tag{35}$$

$$\begin{aligned}\Delta\boldsymbol{p}_{nb}^r = &\left\{\left[-\left(\boldsymbol{I} - \boldsymbol{C}_n^{n'}\right)\hat{\boldsymbol{\omega}}_{ie}^n + \boldsymbol{C}_n^{n'}\delta\boldsymbol{\omega}_{en}^n + \boldsymbol{C}_b^{n'}\delta\boldsymbol{\omega}_{ib}^b\right]\times\right\}\hat{\boldsymbol{p}}_{en}^n \\ &- \boldsymbol{C}_n^{n'}\left(\hat{\boldsymbol{\omega}}_{en}^n\times\right)\delta\boldsymbol{p}_{nb}^r + \boldsymbol{C}_n^{n'}\delta\boldsymbol{v}_{nb}^r\end{aligned} \tag{36}$$

where, $\boldsymbol{\Xi}_{nb}^r$ is the Lie algebra of $\boldsymbol{\delta}_{nb}^r$ on the local tangent plane of the $\mathbb{SE}_2(3)$ group manifold.

Remark 6: There are also differences in the errors and their propagation equations between the LGEO-$\mathbb{SE}_2(3)$-L and LGEO-$\mathbb{SE}_2(3)$-R models. In the LGEO-$\mathbb{SE}_2(3)$-L model, all inconsistency errors arise from the disparity between the b-frame and the b'-frame, whereas in the LGEO-$\mathbb{SE}_2(3)$-R model, the errors come from the differences between the n-frame and the n'-frame. Additionally, compared to the ECI-$\mathbb{SE}_2(3)$ and ECEF-$\mathbb{SE}_2(3)$ models, the LGEO-$\mathbb{SE}_2(3)$ models incorporates more coupling terms induced by the Coriolis effect.

D. The measurement models for SINS/GNSS integration

Generally, GNSS provides auxiliary linear motion observations, and the corresponding measurement models for the abovementioned $\mathbb{SE}_2(3)$-based SINS error kinematics are as follows.

$$\begin{aligned}
&\text{ECI–}\mathbb{SE}_2(3)\text{–L:} \begin{bmatrix}\boldsymbol{y}_v \\ \boldsymbol{y}_p\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{v}}_{ib}^i - \boldsymbol{v}_{ib,\mathrm{GNSS}}^i \\ \hat{\boldsymbol{p}}_{ib}^i - \boldsymbol{p}_{ib,\mathrm{GNSS}}^i\end{bmatrix} = -\boldsymbol{C}_{b'}^i\begin{bmatrix}\delta\boldsymbol{v}_{ib}^l \\ \delta\boldsymbol{p}_{ib}^l\end{bmatrix} \\
&\text{ECI–}\mathbb{SE}_2(3)\text{–R:} \begin{bmatrix}\boldsymbol{y}_v \\ \boldsymbol{y}_p\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{v}}_{ib}^i - \boldsymbol{v}_{ib,\mathrm{GNSS}}^i \\ \hat{\boldsymbol{p}}_{ib}^i - \boldsymbol{p}_{ib,\mathrm{GNSS}}^i\end{bmatrix} = -\boldsymbol{C}_{b'}^i\boldsymbol{C}_b^{b'}\begin{bmatrix}\delta\boldsymbol{v}_{ib}^r \\ \delta\boldsymbol{p}_{ib}^r\end{bmatrix} \\
&\text{ECEF–}\mathbb{SE}_2(3)\text{–L:} \begin{bmatrix}\boldsymbol{y}_v \\ \boldsymbol{y}_p\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{v}}_{eb}^e - \boldsymbol{v}_{eb,\mathrm{GNSS}}^e \\ \hat{\boldsymbol{p}}_{eb}^e - \boldsymbol{p}_{eb,\mathrm{GNSS}}^e\end{bmatrix} = -\boldsymbol{C}_{b'}^e\begin{bmatrix}\delta\boldsymbol{v}_{eb}^l \\ \delta\boldsymbol{p}_{eb}^l\end{bmatrix} \\
&\text{ECEF–}\mathbb{SE}_2(3)\text{–R:} \begin{bmatrix}\boldsymbol{y}_v \\ \boldsymbol{y}_p\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{v}}_{eb}^e - \boldsymbol{v}_{eb,\mathrm{GNSS}}^e \\ \hat{\boldsymbol{p}}_{eb}^e - \boldsymbol{p}_{eb,\mathrm{GNSS}}^e\end{bmatrix} = \begin{bmatrix}\left(\boldsymbol{I} - \boldsymbol{C}_{e'}^e\right)\hat{\boldsymbol{v}}_{eb}^e - \delta\boldsymbol{v}_{eb}^r \\ \left(\boldsymbol{I} - \boldsymbol{C}_{e'}^e\right)\hat{\boldsymbol{p}}_{eb}^e - \delta\boldsymbol{p}_{eb}^r\end{bmatrix} \\
&\text{LGEO–}\mathbb{SE}_2(3)\text{–L:} \begin{bmatrix}\boldsymbol{y}_v \\ \boldsymbol{y}_p\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{v}}_{en}^n - \boldsymbol{v}_{en,\mathrm{GNSS}}^n \\ \hat{\boldsymbol{p}}_{en}^e - \boldsymbol{p}_{en,\mathrm{GNSS}}^e\end{bmatrix} = \begin{bmatrix}-\boldsymbol{C}_{b'}^n\delta\boldsymbol{v}_{nb}^l \\ -\boldsymbol{C}_n^e\boldsymbol{C}_{b'}^n\delta\boldsymbol{p}_{nb}^l\end{bmatrix} \\
&\text{LGEO–}\mathbb{SE}_2(3)\text{–R:} \begin{bmatrix}\boldsymbol{y}_v \\ \boldsymbol{y}_p\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{v}}_{en}^n - \boldsymbol{v}_{en,\mathrm{GNSS}}^n \\ \hat{\boldsymbol{p}}_{en}^e - \boldsymbol{p}_{en,\mathrm{GNSS}}^e\end{bmatrix} = \begin{bmatrix}\left(\boldsymbol{I} - \boldsymbol{C}_{n'}^n\right)\hat{\boldsymbol{v}}_{en}^n - \delta\boldsymbol{v}_{nb}^r \\ -\boldsymbol{C}_n^e\boldsymbol{C}_n^{n'}\delta\boldsymbol{p}_{nb}^r\end{bmatrix}
\end{aligned} \tag{37}$$

Remark 7: It is evident that all right-invariant measurement models introduce inconsistency errors in the prior estimation of the Kalman filter. In contrast, all left-invariant models feature the error model-measurement model independence property. This distinction renders the left-invariant models structurally superior to the right-invariant models for the same navigation frame.

III. Turntable Experiment, Results and Discussion

A. Experiment Setup

A turntable experiment was conducted to assess and compare the efficacy of the proposed $\mathbb{SE}_2(3)$ group-based models for attitude determination in SINS/GNSS integrated navigation. The experiment employed a high-precision three-axis turntable, with the rotational accuracy of its inner, outer, and middle frames all surpassing $3''$. A high-precision SINS was mounted at the center of the turntable, and its specifications are listed in Table I, with a data frequency of 200 Hz. The rotation center remained fixed, with its position precisely determined. To emulate GNSS measurements, velocity and position errors with standard deviations of 0.005 m/s and 0.1 m, respectively, were introduced, with simulated GNSS data generated at 1 Hz. Initially, the turntable was aligned in the "East-North-Up" direction, with the initial pitch angle (θ) set to 0°, initial roll angle (γ) set to 0°, and initial yaw angle (ϕ) set to 0° under the "3-1-2 Euler angle" definition. Following a brief stationary phase, the outer, middle, and inner frames oscillated periodically at 0.1 Hz, with amplitudes of 6°, 3°, and 3°, respectively. After approximately 1860 s of oscillation, the turntable stabilized, realigning with the "East-North-Up" frame.

TABLE I
Main parameters of the SINS used in the experiment

Parameters		Values
Accelerometers	Biases	1 mg
	Random Noise	25 $\mu g/\sqrt{\mathrm{Hz}}$
Gyroscopes	Drifts	$0.005^\circ/\mathrm{h}$
	Random Noise	$0.002^\circ/\sqrt{\mathrm{h}}$

B. Results and Discussion

Initially, the models was tested with an initial attitude of $[0^\circ, 0^\circ, 0^\circ]$, and the corresponding error curves are depicted in Fig. 1. When the initial attitude were sufficiently accurate, all models exhibited high attitude estimation accuracy during the early stages of filtering. However, in the steady-state phase after the turntable ceased rotation, the LGEO-$\mathbb{SE}_2(3)$-R model demonstrated significantly inferior accuracy compared to the other models, with the yaw angle error reaching $170''$. The ECEF-$\mathbb{SE}_2(3)$-R model exhibited slightly superior horizontal attitude estimation accuracy; however, its yaw angle error remained considerable, reaching approximately $25''$. The remaining models achieved comparable attitude accuracy, with steady-state horizontal attitude errors ranging from $4''$ to $6''$ and yaw errors around $2''$. Furthermore, the error curves of the two ECI-$\mathbb{SE}_2(3)$ models exhibited a high degree of overlap, corroborating the previously discussed structural symmetry of these two models.

When the initial misalignment angles were set to $[10^\circ, 10^\circ, 50^\circ]$ (see Fig. 2), the ECEF-$\mathbb{SE}_2(3)$-R and LGEO-$\mathbb{SE}_2(3)$-R models failed entirely to correct for the large misalignment. Although the ECI-$\mathbb{SE}_2(3)$-R model was capable of gradually aligning the estimated attitude with the true values, its convergence rate was noticeably slower than that of the other models. After the turntable ceased rotation, all models maintained satisfactory attitude accuracy under

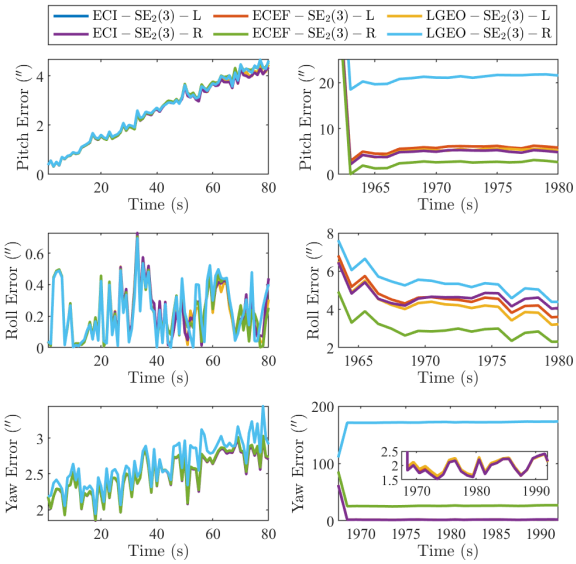


Fig. 1. The attitude error curves under the misalignment set to $[0°, 0°, 0°]$

large dynamic maneuvers, with a pitch angle stability of approximately $2.8'$. Regarding roll and yaw angle stability, the ECI-$\mathbb{SE}_2(3)$ models exhibited a slight advantage over the other two models.

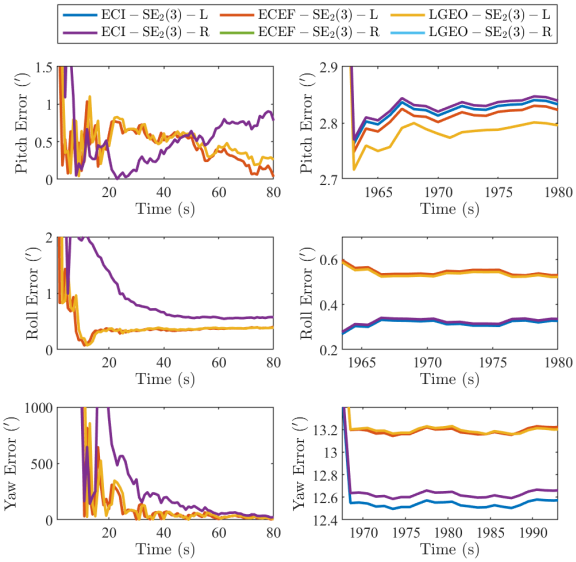


Fig. 2. The attitude error curves under the misalignment set to $[10°, 10°, 50°]$

Next, an extreme misalignment scenario was considered, where the initial misalignment angles were set to $[30°, 30°, 150°]$, with the corresponding results presented in Fig. 3. The ECEF-$\mathbb{SE}_2(3)$-R and LGEO-$\mathbb{SE}_2(3)$-R models remained incapable of handling navigation under large misalignment conditions, despite the availability of accurate velocity and position data from GNSS. The ECI-$\mathbb{SE}_2(3)$-R model also exhibited filter divergence, requiring precise tuning of filtering parameters to correct the misalignment at a relatively slow rate. During the initial filtering phase, the LGEO-$\mathbb{SE}_2(3)$-L model also demonstrated slower correction and higher error fluctuations compared to the other two models. Under this extreme condition, the ECI-$\mathbb{SE}_2(3)$-L and ECEF-$\mathbb{SE}_2(3)$-L models successfully corrected the horizontal attitude angles to near the true values within approximately 15 s and the yaw angles within around 40 s. After the turntable stopped rotating, the stabilized attitude accuracy remained comparable to that observed in the previous scenario.

The three left-invariant models and the ECI-$\mathbb{SE}_2(3)$-R model are all capable of handling large initial misalignments. In these four models, the inconsistency errors are represented through the discrepancies between the b- and b'-frames and are driven by $\boldsymbol{\omega}_{ib}^{b}$. However, in the ECEF-$\mathbb{SE}_2(3)$-R model, this inconsistency error is expressed as $\boldsymbol{C}_{e}^{e'}$ and is governed by $\boldsymbol{\omega}_{ie}^{e}$, whereas in the LGEO-$\mathbb{SE}_2(3)$-R model, it is represented as $\boldsymbol{C}_{n}^{n'}$, driven by $\boldsymbol{\omega}_{in}^{n}$. Evidently, under dynamic motion conditions, the intricate angular dynamics exacerbate attitude errors, making the error kinematics driven by $\boldsymbol{\omega}_{ib}^{b}$ more theoretically justified approach.

Regarding the ECI-$\mathbb{SE}_2(3)$ models, the right-invariant one exhibits significantly higher sensitivity to filtering parameters than the left-

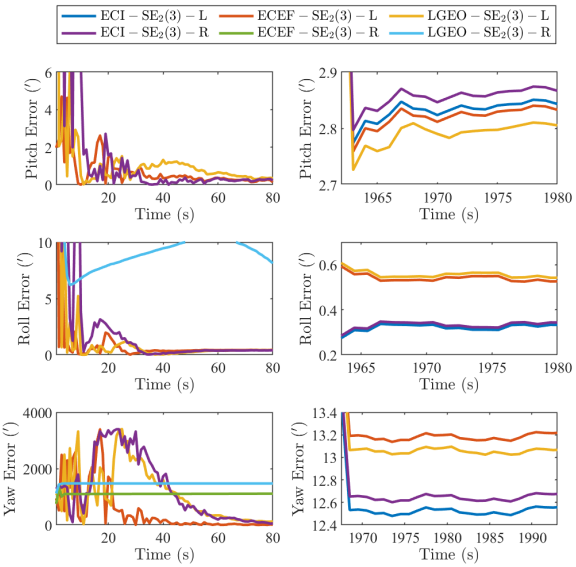


Fig. 3. The attitude error curves under the misalignment set to $[30°, 30°, 150°]$

invariant counterpart. This heightened sensitivity arises from the presence of inconsistency error terms in the right-invariant measurement model, which, in the presence of large misalignments and high-dynamic maneuvers, may degrade the accuracy of the measurement model during the initial filtering stage due to erroneous prior estimations of these inconsistency errors.

From the perspective of global navigation parameter independence, the ECI-$\mathbb{SE}_2(3)$ models maintain a globally independent property, ensuring that their error propagation remains unaffected by inaccuracies in global parameter estimations. In contrast, the LGEO-$\mathbb{SE}_2(3)$-L model exhibits complex error coupling relationships, rendering its error propagation more susceptible to inaccuracies in global parameter estimation. Consequently, its performance is slightly inferior to that of the ECI-$\mathbb{SE}_2(3)$-L and ECEF-$\mathbb{SE}_2(3)$-L models.

IV. Conclusion

This paper systematically derives SINS kinematic and error models on the $\mathbb{SE}_2(3)$ group under different navigation frames. Both theoretical derivation and experimental results and discussion are presented to examine the similarities, distinctions, and relative merits of these models. The study reveals the potential interrelationships between the selection of navigation coordinate system, the left- and right-invariant error models, and the definition of coordinate system inconsistency errors in SINS-based navigation systems. The experimental results demonstrate that the ECI-$\mathbb{SE}_2(3)$-L and ECEF-$\mathbb{SE}_2(3)$-L models exhibit superior robustness and accuracy in attitude estimation, particularly under large initial misalignment and high-dynamic conditions, successfully converging to true values with minimal filtering sensitivity. In contrast, the ECEF-$\mathbb{SE}_2(3)$-R and LGEO-$\mathbb{SE}_2(3)$-R models fail to correct large misalignments, while the ECI-$\mathbb{SE}_2(3)$-R model, though capable of alignment, suffers from slower convergence and higher sensitivity to filtering parameters.

References

[1] M. S. Andrle and J. L. Crassidis, "Attitude estimation employing common frame error representations," *Journal of Guidance, Control, and Dynamics*, vol. 38, no. 9, pp. 1614–1624, 2015.

[2] M. Brossard, S. Bonnabel, and J.-P. Condomines, "Unscented kalman filtering on lie groups," in *2017 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS)*. Vancouver, Canada: IEEE, 2017, pp. 2485–2491.

[3] M. Brossard, A. Barrau, and S. Bonnabel, "Exploiting symmetries to design ekfs with consistency properties for navigation and slam," *IEEE Sensors Journal*, vol. 19, no. 4, pp. 1572–1579, 2019.

[4] ——, "Ai-imu dead-reckoning," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 4, pp. 585–595, 2020.

[5] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *Transactions on Automatic Control*, vol. 62, no. 4, pp. 1797–1812, Apr 2017.

[6] ——, "Invariant kalman filtering," *Annual Review of Control Robotics and Autonomous Systems*, vol. 1, pp. 237–257, 2018.

[7] T. Zhu, A. Li, K. Li, and F. Qin, "The quaternion based error model based on se(3) of the ins," *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13 067–13 077, Jun 2022.

[8] L. Chang, J. Di, and F. Qin, "Inertial-based integration with transformed ins mechanization in earth frame," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 3, pp. 1738–1749, Jun 2022.

[9] L. Chang, Q. Bian, Y. Zuo, and Q. Zhou, "Sins/gnss-integrated navigation based on group affine sins mechanization in local-level frame," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 5, pp. 2471–2482, 2023.

[10] K. Fang, T. Cai, and B. Wang, "The kinematic models of the sins and its errors on the se(3) group in the earth-centered inertial coordinate system," *Sensors*, vol. 24, no. 12, 2024. [Online]. Available: https://www.mdpi.com/1424-8220/24/12/3864

Quadcopter Motion Control System under Disturbances Based on Fuzzy Logic

A.D. Sheyko
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

V.M. Nikiforov
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

A.A. Gusev
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

K.A. Andreev
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

M.L. Simakov
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

A.V. Shevchenko
Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

***Abstract* — The article is devoted to improving trajectory tracking accuracy using fuzzy logic-based controllers. The work presents the development of a height stabilization system and a linear motion system for a quadcopter based on a fuzzy controller and a fuzzy PID controller, as well as simulation results of trajectory tracking under various wind disturbances**

Keywords* — *quadcopter, mathematical modelling, fuzzy logic, fuzzy controller, control system.

I. Introduction

Currently, there is active development and practical application of systems and technical objects that utilize fuzzy logic for regulation and control.

The advantages of fuzzy logic lie in its ability to represent human thinking—that is, human decision-making methods—and to model complex objects using natural language. This makes fuzzy logic particularly useful in tasks where traditional mathematical models prove insufficiently effective or overly complex to implement.

The objective of this study is to improve the quality of quadcopter motion control along reference trajectories under wind disturbances.

The object of the study is a quadcopter represented by a mathematical model in the form of a system of differential equations.

The subjects of the study are:

1. A fuzzy logic-based controller used for quadcopter control under disturbing influences.
2. A fuzzy PID controller used for quadcopter control under disturbing influences/

II. Problem Statment

The tasks of the work can be formulated as follows:

1. For a quadcopter represented by a mathematical model, it is necessary to develop motion control systems based on a fuzzy controller and a fuzzy PID controller.
2. Conduct mathematical modeling of reference trajectory tracking under wind disturbances.
3. Perform a comparative analysis of various control methods to determine the most effective approach.

III. Mathematical Model of the Quadcopter

The mathematical model of the quadcopter is represented as a system of differential equations [1]:

$$\begin{cases} \ddot{x} = \dfrac{\left(\left(F\left(\sin\varphi\sin\psi+\cos\varphi\cos\psi\sin\theta\right)\right)-f_x\right)}{m} \\ \ddot{y} = \dfrac{\left(\left(F\left(\cos\varphi\sin\psi\sin\theta+\cos\psi\sin\varphi\right)\right)-f_y\right)}{m} \\ \ddot{z} = \dfrac{F\cos\varphi\cos\theta - f_z - mg}{m} \\ \dot{p} = \dfrac{\left(I_y - I_z\right)}{I_x}qr + \dfrac{\tau_\varphi}{I_x} \\ \dot{q} = \dfrac{\left(I_z - I_x\right)}{I_y}pr + \dfrac{\tau_\theta}{I_y} \\ \dot{r} = \dfrac{\left(I_x - I_y\right)}{I_z}pq + \dfrac{\tau_\psi}{I_z} \\ \tau_\varphi = bl\left(w_2^2 - w_4^2\right) \\ \tau_\theta = bl\left(w_3^2 - w_1^2\right) \\ \tau_\psi = d\left(w_1^2 + w_3^2 - w_2^2 - w_4^2\right) \\ F = b\left(w_1^2 + w_2^2 + w_3^2 + w_4^2\right) \\ \dot{\varphi} = p + q\sin\varphi\tan\theta + r\cos\varphi\tan\theta \\ \dot{\theta} = q\cos\varphi - r\sin\varphi \\ \dot{\psi} = \dfrac{\left(q\sin\varphi + r\cos\varphi\right)}{\cos\theta} \end{cases}$$

where:

x, y, z – coordinates in the Earth-fixed reference frame;
φ, θ, ψ – roll, pitch, and yaw angles, respectively;
p, q, r – projections of the angular velocity vector onto the body-fixed coordinate system;

I_x, I_y, I_z – moments of inertia of the vehicle about the body-fixed axes;

τ_φ, τ_θ, τ_ψ – roll, pitch, and yaw moments, respectively;

l – distance from the center of mass to the rotor axis;

w_i – angular velocity of the i-th rotor;

F – total thrust;

b – thrust coefficient;

d – aerodynamic drag coefficient;

f_x, f_y, f_z – projections of air resistance forces.

IV. Design of Controllers

The proposed fuzzy controller and fuzzy PID controller used in this work are applied for regulating altitude (z) and horizontal coordinates (x, y). For angular stabilization (φ, θ, ψ), a classical PID controller is employed.

The fuzzy controller consists of the following blocks: fuzzification, rule base, inference engine (logical decision-making), defuzzification.

The block diagram of the control system with a fuzzy controller and feedback is shown in figure 1 [2].

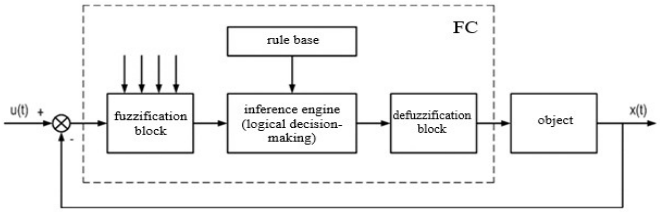


Fig. 1. Control system with a fuzzy regulator

The input and output parameters of the fuzzy controller, that is, linguistic variables are: system error (θ), rate of change of system error (θ̇), control output (m).

Triangular membership functions are used for the linguistic variables, as illustrated in figure 2.

For software implementation of the fuzzy controller, analytical expressions for the control outputs must be derived, assuming identical triangular membership functions with five fuzzy sets: BN (Big Negative), N (Negative), Z (Zero), P (Positive), BP (Big Positive).

The rule base of the fuzzy controller is provided in table 1.

TABLE I. THE RULE BASE OF THE FUZZY CONTROLLER

θ̇ \ θ	BN	N	Z	P	BP
BN	***BN***	***BN***	***N***	***N***	***Z***
N	***BN***	***N***	***N***	***Z***	***P***
Z	***N***	***N***	***Z***	***P***	***P***
P	***N***	***Z***	***P***	***P***	***BP***
BP	***Z***	***P***	***P***	***BP***	***BP***

The initial scaling factors for the fuzzy controller's input and output linguistic variables (A_m, B_m, D_m) are set to: maximum possible error, maximum rate of error change, maximum control output.

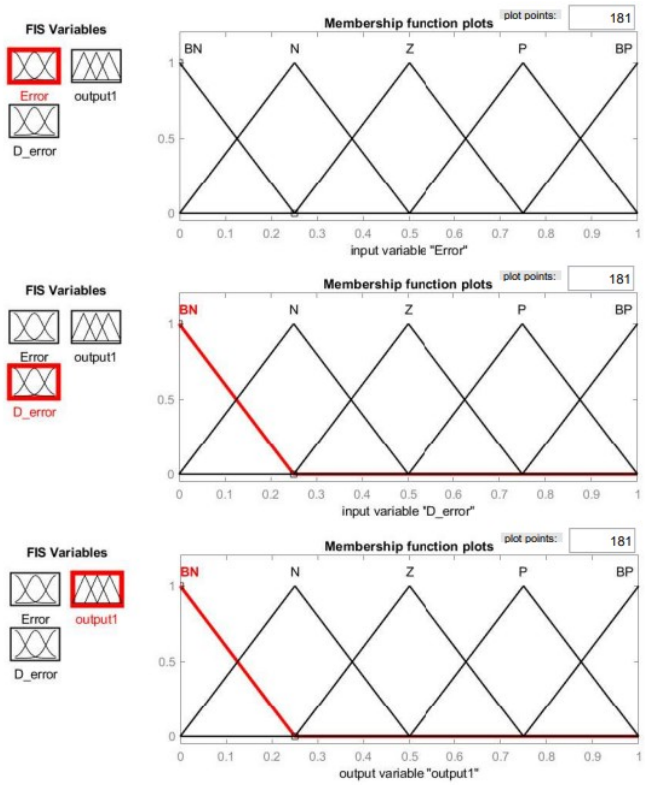


Fig. 2. Membership functions of linguistic variables

The parameter values for the fuzzy controller are listed in table 2.

TABLE II. THE PARAMETER VALUES FOR THE FUZZY CONTROLLER

	A_m	B_m	D_m
x	***3.4883***	***2.6013***	***3564.4***
y	***3.4883***	***2.6013***	***3564.4***
z	***3.6054***	***2.9363***	***291.6074***

The fuzzy PID controller operates by using fuzzy logic to dynamically adjust the PID parameters for optimal performance.

The fuzzy PID controller consists of two layers: conventional PID controller on lower layer, fuzzy inference system on upper layer.

The use of fuzzy PID controller allows you to improve the quality of regulation and reduce overshoot. In addition, it can adapt to changes in the operating conditions of the system, which makes it more flexible and efficient. Figure 3 shows the structural diagram of a fuzzy PID controller.

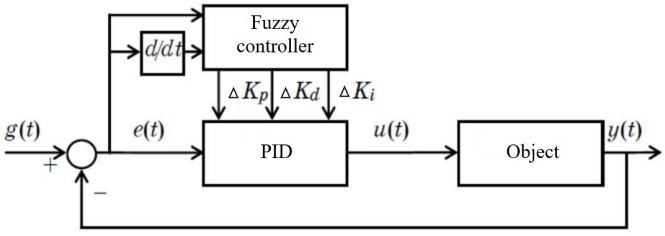


Fig. 3. Block diagram of the fuzzy PID controller

The rule base of the fuzzy PID controller is identical to that in table 1. The membership functions are the same as those used in the fuzzy controller as illustrated in figure 2.

The parameter values for the fuzzy logic component of the proposed controller are provided in table 3.

TABLE III. PARAMETER VALUES OF THE FUZZY PID CONTROLLER

	A_m	B_m	D_{mK_p}	D_{mK_i}	D_{mK_d}
x	*3.4883*	*2.6013*	*4*	*0*	*14*
y	*3.4883*	*2.6013*	*4*	*0*	*14*
z	*3.6054*	*2.9363*	*4*	*1.5*	*14*

V. REFERENCE TRAJECTORY TRACKING PERFORMANCE

Simulations were conducted for two trajectories (vertical circle and helical path) under the following wind disturbances according to the Beauport scale: Moderate wind (5.6 m/s), Strong wind (11.2 m/s), Very strong wind (18 m/s).

The mean-square error (MSE) was used as a quantitative measure of trajectory tracking accuracy.

A comparison was made with a PID-based control system from prior research [1].

Table 4 presents the MSE values for each control method and trajectory under varying wind conditions.

TABLE IV. MEAN-SQUARE ERROR

Axis	Fuzzy Controller	Fuzzy PID Controller	PID Controller
Vertical Circle – Moderate Wind (5.6 m/s)			
x, м	0	0.0033	0.0038
y, м	0.0056	0.0014	0.0029
z, м	0.0089	0.0325	0.0337
Vertical Circle – Strong Wind (11.2 m/s)			
x, м	0.0030	0.1128	0.0721
y, м	0.0129	0.3889	0.0340
z, м	0.0096	0.0549	0.0469
Vertical Circle – Very Strong Wind (18 m/s)			
x, м	0.0700	0.8575	26.4553
y, м	1.3806	0.5355	17.7129
z, м	0.0176	0.1465	0.1656
Helical Path – Moderate Wind (5.6 m/s)			
x, м	0.0016	0.0031	0.0033
y, м	0.0542	0.0467	0.0465
z, м	0.0063	0.0122	0.0123
Helical Path – Strong Wind (11.2 m/s)			
x, м	0.0021	0.0026	0.0025
y, м	0.0606	0.0513	0.0504
z, м	0.0072	0.0118	0.0119
Helical Path – Very Strong Wind (18 m/s)			
x, м	0.6986	13.5325	52.5325
y, м	0.8632	1.5545	12.4323
z, м	0.0245	0.0595	0.0735

Moderate Wind (5.6 m/s): the fuzzy PID controller and PID controller exhibit similar accuracy for both trajectories. The fuzzy controller outperforms them by an order of magnitude in z-axis precision.

Strong Wind (11.2 m/s): the PID and fuzzy PID controllers shows comparable performance. The fuzzy controller maintains significantly higher accuracy (order of magnitude better).

Very Strong Wind (18 m/s): the PID controller fails to track the reference trajectories. The fuzzy PID controller succeeds only for the vertical circle, but with reduced accuracy. The fuzzy controller remains operational for both trajectories, demonstrating superior robustness.

The trajectory tracking performance of the proposed controllers under 18 m/s wind is illustrated in figures 4 and 5.

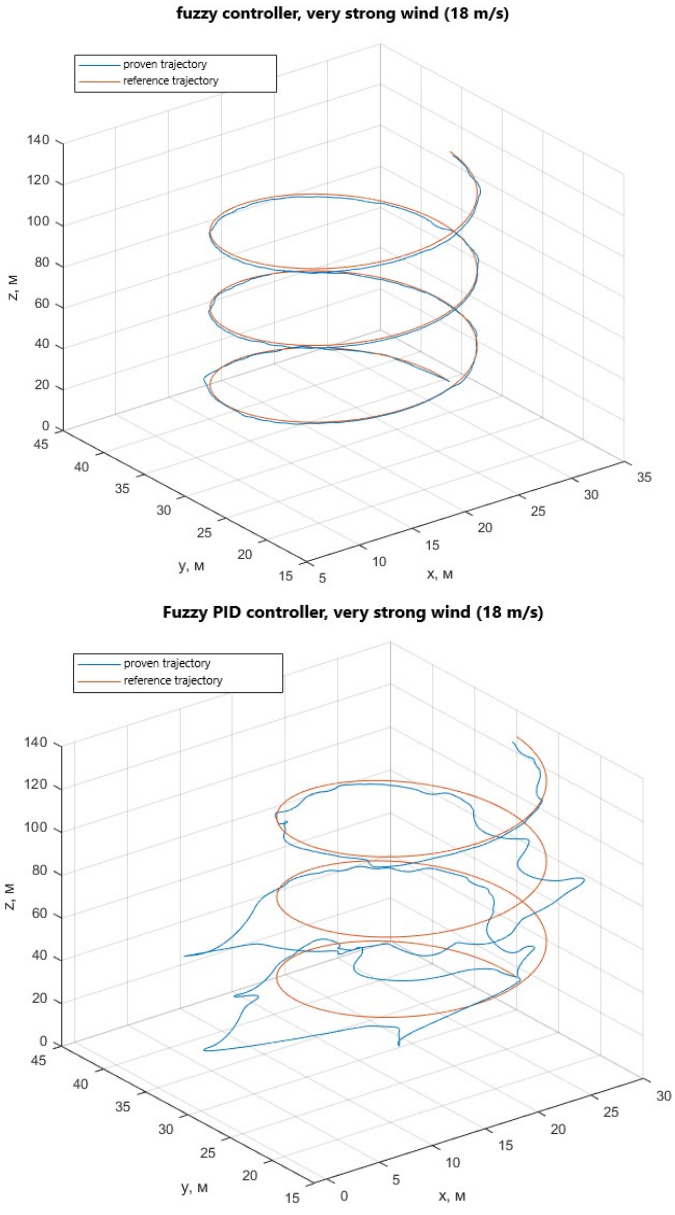


Fig. 4. Helical path

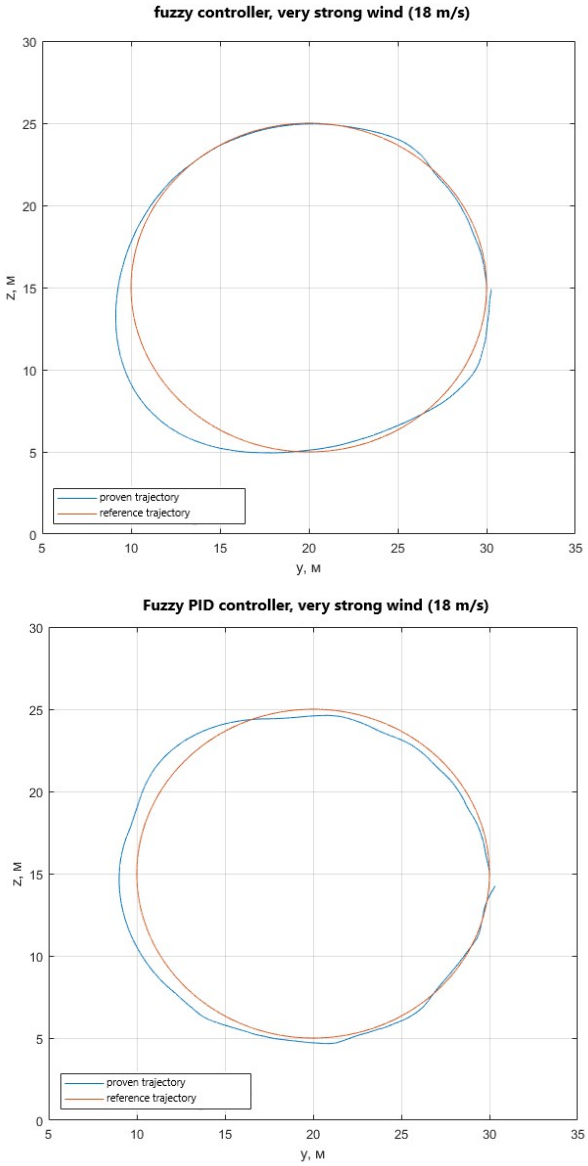


Fig. 5. Vertical circle

VI. CONCLUSION

The study demonstrated that the fuzzy logic-based control system provides higher accuracy and stability under strong wind disturbances compared to both the fuzzy PID controller and the conventional PID controller. The fuzzy controller successfully tracks trajectories even under very strong wind conditions (18 m/s), where other methods fail.

REFERENCES

[1] A.D. Sheyko Trajectory Tracking Control of a Quadcopter Based on Inverse Dynamics Solution, Trudy FGUP "NPTsAP". Systemy i pribory upravleniia, 2023, no. 4, pp. 64-76.

[2] V.I. Gostev, Design of Fuzzy Controllers for Automatic Control Systems, St. Petersburg: BHV-Petersburg, 2011, pp. 238-306