**Review Article**

**A Review of Navigation and SLAM Technologies in Orchard Environments**

**Abstract**

This paper reviews the research status of orchard environment navigation technology and Simultaneous Localization and Mapping (SLAM) technology. In the field of orchard navigation, researchers primarily utilize LiDAR and visual sensors to achieve autonomous navigation, enhancing the operational efficiency of robots through map construction and path planning. LiDAR, with its all-weather operational capability, shows broad application prospects in orchard environments, while visual sensors perform poorly under limited lighting conditions. As a core technology for robot navigation, SLAM has evolved from traditional methods to modern optimization algorithms. Currently, laser SLAM and visual SLAM each have their advantages in different scenarios. Laser SLAM demonstrates higher robustness in complex environments, while visual SLAM is more cost-effective and better at capturing detailed environmental information. Future research will focus on multi-sensor fusion and algorithm optimization to further improve the navigation capabilities of robots in complex environments.

**Keywords**

Orchard robots, mapping, navigation

Introduction

With the rapid development of intelligent robot technology, the field of agricultural intelligence has attracted significant attention from researchers both domestically and internationally. To meet the specific needs of different crops, researchers have developed various intelligent agricultural devices, such as fruit-picking robots, automated spray robots, and intelligent weeding robots. These devices are typically based on mobile robot platforms and are equipped with manipulators, spraying devices, and other actuators, enabling them to efficiently perform specific agricultural tasks. This paper reviews the current state of research on agricultural mobile robots from the perspective of autonomous navigation technology.

Xu Yuyang [1] explored map construction and path planning methods for mobile robots in orchard environments. The study used 2D LiDAR to construct maps of orchard environments and systematically evaluated the accuracy of the algorithms employed. Additionally, to address path planning challenges from start to end points, the author proposed a strategy of inserting intermediate navigation points at turning points, breaking down the overall path planning task into multiple local path planning tasks between navigation points, thereby improving the efficiency and accuracy of path planning.

Hu Guangrui [2] and others designed an orchard navigation system based on solid-state LiDAR, as shown in Figure 1 (a). This system collects point cloud data of the orchard environment using solid-state LiDAR and extracts initial values of inter-row navigation lines through the RANSAC algorithm. Subsequently, an artificial potential field method is used to optimize the initial path, generating a navigation path that effectively avoids tree canopies and other obstacles, enabling autonomous robot navigation.

|  |  |
| --- | --- |
| （a） | (b) |

Figure-1 Mobile robot for orchards

Liu Weihong [3] and others proposed a 3D LiDAR-based inter-row navigation method for orchards, as shown in Figure 1 (b). This method uses 3D LiDAR to collect real-time point cloud data of the orchard environment and extracts the position information of tree trunks through point cloud segmentation technology. Based on the spatial distribution characteristics of tree rows on both sides, the research team combined the RANSAC algorithm with the least squares method (LSM) to fit the tree rows and generate a centerline as the navigation reference. Additionally, this method improved the pure tracking algorithm, enabling differential-drive robots to track tree rows more accurately, thereby enhancing the stability and accuracy of navigation.

Opiyo S [4] and others designed an autonomous navigation and tracking system for orchard robots based on Gabor filters and K-means algorithms. This system segments RGB images of the orchard environment by combining these two algorithms, extracting the central axis of the inter-row ground as the navigation reference. Furthermore, the research team developed a fuzzy controller to track the central axis in real-time, achieving efficient navigation tasks in orchard environments.

Astolfi P [5] and others integrated LiDAR and vehicle odometry on a commercial chassis to conduct experimental studies on localization and navigation in vineyard environments. This experimental platform, developed under the ROS (Robot Operating System) framework, used the Gmapping algorithm to construct a 2D grid map of the orchard environment and achieved precise localization through the AMCL algorithm. Although the study primarily relied on simulation environments due to insufficient experimental data, it provided valuable insights into developing orchard robot navigation systems based on open-source platforms, as shown in Figure 2a.

|  |  |
| --- | --- |
| （a） | （b） |

Figure 2 Orchard operation robot platform

Caballero F [6] and others proposed a localization technology based on 3D LiDAR and point cloud maps, successfully applied to unmanned aerial vehicle platforms. The core of this method lies in achieving high-precision registration between point clouds and maps, minimizing the distance error between point clouds and maps through nonlinear optimization algorithms. Given an initial pose, combined with odometry-based pose prediction, the method effectively tracks the real-time pose of the robot and achieves significant localization results.

To meet the autonomous navigation requirements of orchards, a research team from Huazhong University of Science and Technology [7-8] developed an air-ground collaborative perception framework. This framework includes the following key steps: First, aerial imagery data of the entire orchard is acquired using a drone platform. Tree targets are detected using image parsing algorithms based on threshold segmentation and edge detection, generating a structured feature map with tree distribution information. To improve the precision of the navigation reference, geometric feature modeling of tree rows is performed, and classic Hough transform combined with the RANSAC algorithm is used to efficiently extract linear features. During navigation execution, real-time LiDAR point cloud data is registered with the pre-built feature map to solve optimal observation pose parameters. The system ultimately designs an Extended Kalman Filter (EKF) to dynamically fuse motion prediction and observation data, ensuring centimeter-level localization accuracy for the mobile platform in complex orchard scenarios, as shown in Figure 2b.

Summarizing the above content, robots in orchard environments typically rely on LiDAR or visual sensors to achieve autonomous navigation through real-time or pre-set navigation paths. Based on visual or LiDAR technology, robots can extract navigation lines to complete navigation tasks without relying on GPS, achieving good results. Although image processing technology can accurately extract key parameters such as navigation lines, visual sensors are highly sensitive to lighting conditions and struggle to operate normally in nighttime or low-light environments, limiting the operational efficiency of robots. In contrast, LiDAR measures distances by emitting and receiving laser beams, enabling all-weather operation and showing broader application prospects in orchards and outdoor environments. In recent years, using 3D point cloud maps as prior information and determining robot positions through constraints between real-time point clouds and prior maps has gradually become a reliable localization method, providing important technical references for precise localization of orchard robots.

1.2.2 **SLAM Research Status**

The development of SLAM technology can be traced back nearly 40 years, with its evolution roughly divided into three stages, as shown in Figure 3. The first stage is the traditional era, during which researchers attempted various methods to solve the SLAM problem, such as particle filtering, Extended Kalman Filtering (EKF), and maximum likelihood estimation, and conducted in-depth analysis and verification of their convergence. The second stage focused on algorithmic analysis, with research emphasizing fundamental SLAM characteristics, including map consistency and algorithmic convergence. The third stage is the prediction-robustness period, during which researchers optimized existing algorithms and further enhanced environmental perception, localization, and mapping capabilities to meet the demands of more complex scenarios [9]. Mainstream SLAM methods are typically divided into visual SLAM and laser SLAM, each with its advantages in different scenarios, as described below.

（1）**Laser SLAM**

Laser SLAM technology is primarily based on LiDAR and can be categorized into 2D SLAM and 3D SLAM depending on the dimensionality of the LiDAR. 2D SLAM is widely used in indoor environments, such as household robotic vacuum cleaners, while 3D SLAM is suitable for various complex indoor and outdoor environments, such as high-precision map construction and autonomous driving. The appearance of 2D and 3D LiDAR is shown in Figure 3.

|  |  |
| --- | --- |
| LiDAR - RoboSense - LiDAR for Autonomous Driving, Robots, V2X | 【PUCK LITE】VLP-16-LW 三维激光雷达-北京华微中测科技有限公司 |

Figure 3 Two Types of LiDAR

In 2008, Grisetti [23] proposed an improved Rao-Blackwellized filtering SLAM algorithm called Gmapping. This algorithm optimized the proposal distribution of the odometry model by reducing the number of particles, thereby alleviating memory explosion issues caused by significant discrepancies between the proposal distribution and the actual distribution. However, since Gmapping is fundamentally based on particle filtering, its computational efficiency issues remain unresolved, limiting its application in large-scale map scenarios.

In 2010, Konolige [10] developed a graph optimization framework-based SLAM algorithm called Karto SLAM. This method models the constraints between LiDAR poses and landmark points as a graph structure and solves it using optimization algorithms. Although the framework considers sparsity issues, its complex optimization process results in low computational efficiency, making it difficult to meet real-time demands.

In 2011, Kohlbrecher [11] proposed the Hector SLAM method, which uses scan-to-map technology and does not rely on odometry information, making it suitable for flat areas. Notably, Hector SLAM has high frequency requirements for sensors and lacks a loop closure module, making it unsuitable for large-scale map scenarios.

In 2016, Google [12] launched Cartographer, an open-source real-time indoor mapping system. As shown in Figure 5, the system is based on a graph optimization framework. Its core mechanism involves processing each frame of LiDAR data with scan matching algorithms to determine optimal pose estimates and integrate these data into the current sub-map. The scan matching process is optimized only for the current sub-map. After the sub-map is constructed, the system performs local loop closure to ensure spatial consistency between sub-maps using re-localization techniques and pre-built grid data, significantly improving the accuracy of the global map and localization. Finally, the system performs global loop optimization to ensure global consistency of the entire map. As one of the current open-source 2D SLAM algorithms, Cartographer performs excellently.

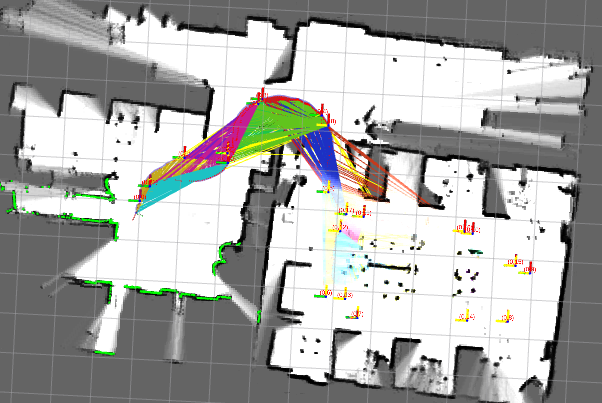


Figure 4 2D Raster Map Example

Currently, 3D laser SLAM technology is primarily divided into two categories: mechanical radar and solid-state radar. Solid-state radar can be further subdivided into three types: phased array, FLASH, and MEMS. Although solid-state radar solutions like loam\_livox have significant technical advantages, their non-repetitive scanning characteristics and limited field of view often lead to degradation issues in point cloud matching due to insufficient feature points. This is particularly true in non-structured environments such as orchards, where scene features are relatively monotonous, limiting the application of solid-state radar. Therefore, this section discusses 3D SLAM solutions based on 3D mechanical LiDAR.

Most 3D mechanical LiDAR solutions are based on the LOAM framework. The core of this framework lies in the front-end processing, which extracts key features (primarily planar and edge features) from the current frame of LiDAR data. The back-end then optimizes the initial values for matching the current frame with the map by fusing information from multiple sensors, playing a crucial role in laser SLAM systems. The LOAM framework, proposed by J. Zhang [26] and others in 2014, is a classic method for laser odometry and mapping. It relies solely on LiDAR data and decomposes the SLAM task into a high-frequency, low-precision front-end module and a low-frequency, high-precision back-end module, achieving real-time LiDAR odometry computation. However, the limitation of LOAM is its reliance on radar odometry, which leads to significant error accumulation in feature-sparse open scenes. The specific process is shown in Figure -5.

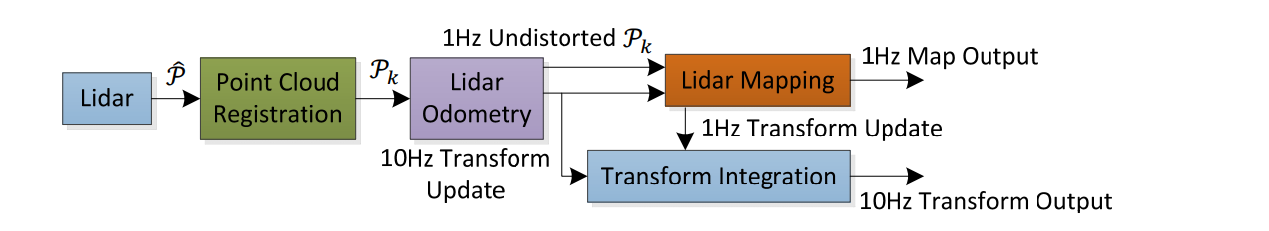


Figure 5 LOAM algorithm flow

In 2018, Shan and Englot [13] proposed the LeGO-LOAM algorithm based on LOAM. This algorithm separates ground points from non-ground points by calculating the Gaussian curvature of each point in the point cloud, thereby reducing interference from non-ground points on laser odometry. However, in geometrically similar scenes, due to the limited detection range of LiDAR, the algorithm may lack sufficient constraints, leading to significant performance degradation or even failure of LOAM-based algorithms. To further enhance performance, SC-LeGO-LOAM introduces a loop closure detection mechanism based on Scan context, significantly improving the speed and accuracy of loop closure detection, although it also increases computational resource consumption. Figure 6 shows an example of a 3D point cloud map generated by 3D SLAM.

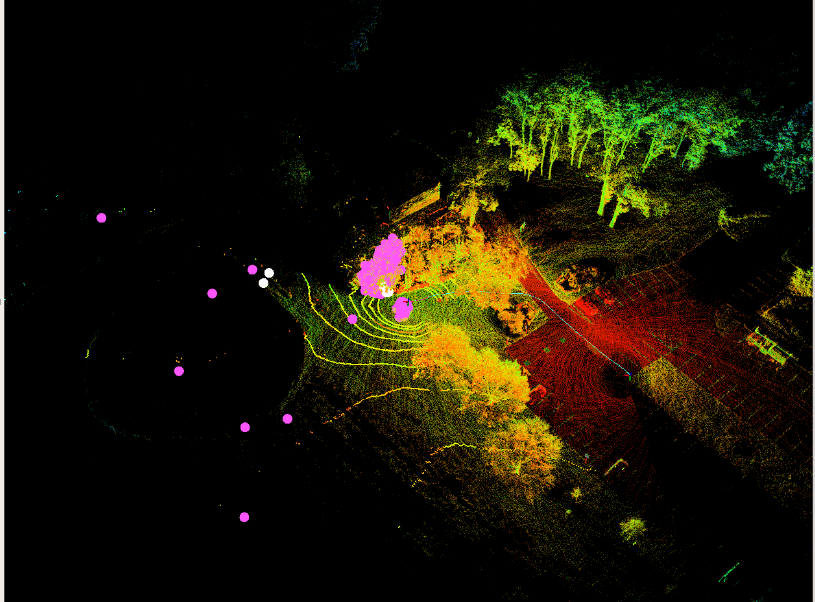


Figure 6 Example of 3D Point Cloud Map

（2）**Visual SLAM**

Compared to laser SLAM, visual SLAM has significant advantages in cost control and technical maturity. Laser SLAM has certain limitations in re-localization capabilities, particularly in scenes with geometrically similar features or dramatic dynamic changes, where its adaptability and robustness are inferior. In contrast, visual sensors (such as monocular, binocular, and multi-camera systems) can capture rich environmental detail information, including object appearance and shape features, which are important for environmental perception and difficult to obtain through other sensors. In recent years, thanks to the rapid development of digital image processing technology and significant improvements in computer hardware performance, visual SLAM technology has become increasingly mature. Despite this, visual SLAM systems still face many challenges, such as susceptibility to weather and lighting conditions, and issues with stability and reliability in dynamic scenes, unevenly distributed feature points (too many or too few), and occlusion scenarios.

Visual odometry is a technology based on image data processing, used to estimate the position and orientation of a camera carrier. Its implementation methods primarily include feature point-based and direct methods.

**Feature Point-Based SLAM Research**

The feature point method is an important approach in visual odometry. Its core idea is to perform feature association and matching on each frame of point cloud data, and then use spatial geometric relationships to calculate camera pose and environmental landmarks. This method is stable, insensitive to lighting changes and dynamic object interference, and has become a mature and widely used solution. Due to its stability and insensitivity to lighting changes, the feature point method has become a mature approach in the field of visual odometry.

In 2017, Raúl [14] and his team from the University of Zaragoza proposed a feature point-based visual SLAM algorithm called ORB-SLAM2. This algorithm uses epipolar geometry and the PnP method to compute inter-frame pose transformations and introduces a loop closure detection mechanism to effectively address error accumulation during system operation.

In 2018, Qin [15] and others from the Hong Kong University of Science and Technology proposed a VINS-Mono algorithm that fuses a monocular camera with an IMU. This algorithm resolves the scale uncertainty issue in monocular SLAM using scale information provided by the IMU, significantly improving performance in scenes with dramatic lighting changes or texture Missing .

In 2022, Weng Jianhong [16] and his team from Guangdong University of Technology conducted an in-depth study on the scale quantization error issue in ORB-SLAM2. They proposed a twin filter algorithm that constructs descriptors in a log-polar coordinate system and generates twin descriptors in a Cartesian coordinate system within an image pyramid, significantly enhancing the scale invariance of feature points. This algorithm not only improves the discreteness of feature points but also increases the number and accuracy of matched points, thereby effectively improving system performance.

Guo Qiang [17] from Nanjing University of Posts and Telecommunications addressed the performance limitations of ORB-SLAM2 in low-texture environments by proposing an improved method. This method enhances the extraction of linear features in the environment by introducing an LSD line feature detection mechanism, thereby improving the robustness of the algorithm in low-texture scenes. Additionally, the author combined point features, line features, and semantic information to construct an error function, further optimizing camera pose estimation. Experimental results showed that, compared to traditional visual SLAM algorithms that rely solely on point features, this improved method significantly enhanced the accuracy of pose estimation.

**Direct Method-Based SLAM Research**

The direct method is a visual odometry approach that does not rely on extracting feature points from images but instead directly uses pixel value information for pose estimation, thereby utilizing image information more comprehensively. The primary advantage of the direct method is its ability to operate even when feature points are insufficient or missing. However, during processing, photometric errors must be considered, and efficient computational methods are required when handling large-scale pixel information.

In 2011, Newcombe and his team from the University of Washington proposed a 3D reconstruction method based on depth cameras called Kinect Fusion. This method directly acquires point cloud data, uses the ICP algorithm to construct a loss function for optimal pose estimation, and introduces a TSDF model to describe target surface information and the distance relationship with the camera. Experimental results showed that the reconstruction speed of this algorithm can keep pace with the frame rate of the depth camera.

In 2014, Kerl and others from the Technical University of Munich proposed a direct method-based visual SLAM algorithm called DVO SLAM. This algorithm introduces an unknown transformation quantity and constructs a residual function for optimization. Although it performed well on public datasets, its reliance on the grayscale invariance assumption limits its practical applicability, as real-world application scenarios often fail to meet this condition.

In 2016, Engel and his team from the Technical University of Munich proposed the DSO algorithm to address the issue of insufficient odometry accuracy. This algorithm projects map points of the current frame into keyframes within a sliding window and constructs a residual function for optimization, thereby obtaining inter-frame transformation results. DSO demonstrated excellent performance in terms of accuracy and robustness but was prone to tracking loss when image quality was poor or initial pose estimation was inaccurate.

In 2018, Wu Yuxiang [20] and her team from the South China University of Technology proposed an innovative semi-direct SLAM technique. This technique selects FAST feature points with significant grayscale gradients in images and optimizes camera localization by minimizing photometric errors. To enhance algorithm stability, the research team used pixel blocks instead of individual pixels to construct the error function. Although this algorithm performed well in real-time processing, its optimization process was prone to falling into local optima.

**Conclusion**

Research on orchard environment navigation technology indicates that LiDAR and visual sensors are key technologies for achieving autonomous navigation. LiDAR excels in all-weather operation and high-precision localization, making it highly effective in orchard environments, while visual sensors, though cost-effective, have significant limitations under poor lighting conditions. As the core technology for robot navigation, SLAM has made significant progress in algorithm optimization and environmental perception in recent years. Laser SLAM demonstrates higher robustness in complex environments, while visual SLAM offers advantages in capturing detailed environmental information and cost control. Future research should further explore multi-sensor fusion technologies to combine the strengths of LiDAR and visual sensors, enhancing the navigation accuracy and stability of robots in complex orchard environments. Additionally, to address the real-time and robustness challenges of SLAM algorithms, researchers need to develop more efficient optimization algorithms to meet practical application requirements.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:1

1.

2.

3.

Reference

1. Xu Yuyang. Research on Map Construction and Path Planning for Orchard Mobile Robots [D]. Nanjing Forestry University, 2023.
2. Hu Guangrui, Kong Weiyu, Qi Chuang, et al. Path Optimization for Mobile Harvesting Robot in Orchard Environment [J]. *Journal of Agricultural Engineering*, 2021, 37(09): 175-184.
3. Liu Weihong, He Xiongkui, Liu Yajia, et al. 3D LiDAR-Based Navigation Method for Orchard Rows [J]. *Journal of Agricultural Engineering*, 2021, 37(09): 165-174.
4. Opiyo S, Okinda C, Zhou J, et al. Medial Axis-Based Machine-Vision System for Orchard Robot Navigation [J]. *Computers and Electronics in Agriculture*, 2021: 106153.
5. Astolfi P, Gabrielli A, Bascetta L, et al. Autonomous Navigation in Vineyard for the Echord++ Grape Experiment [J]. *IFAC-PapersOnLine*, 2018, 51(11): 704-709.
6. Caballero F, Merino L. DLL: Direct LiDAR Localization. A Map-Based Localization Approach for Aerial Robots [C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 5491-5498.
7. Zhang Zheyuan. Key Technologies for Autonomous Navigation of Orchard Robots [D]. Huazhong University of Science and Technology, 2016.
8. Huang Hao. Development of Orchard Autonomous Cruise Operation Robot [D]. Huazhong University of Science and Technology, 2016.
9. Wu Jianqing, Song Xiuguang. Review of Simultaneous Localization and Mapping Technology [J]. *Journal of Shandong University (Engineering Science)*, 2021, 51(05): 16-31.
10. Grisetti G, Stachniss C, Burgard W. Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters [J]. *IEEE Transactions on Robotics*, 2007, 23: 34-46.
11. Kohlbrecher S, Stryk O, Meyer J, et al. A Flexible and Scalable SLAM System with Full 3D Motion Estimation [C]//Proceedings of the IEEE International Symposium on Safety, Security, and Rescue Robots (SSRR). IEEE, 2011.
12. Hess W, Kohler D, Dapp H, et al. Real-Time Loop Closure in 2D LiDAR SLAM [C]//2016 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2016.
13. Zhang Ji, Singh S. LOAM: Lidar Odometry and Mapping in Real-Time [C]//Robotics: Science and Systems. 2014.
14. Shan T, Englot B. Lego-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain [C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2018: 4758-4765.
15. Mur-Artal R, Tardós J D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras [J]. *IEEE Transactions on Robotics*, 2017, 33(5): 1255-1262.
16. Tong Q, Peiliang L, Shaojie S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator [J]. *IEEE Transactions on Robotics*, 2018, 34(4): 1004-1020.
17. Weng Jianhong, Lei Qun, Tao Jie, et al. Improved Visual Odometry Method Based on ORB [J]. *Control and Decision*, 2023: 1-9.
18. Guo Qiang. Monocular Semantic Visual SLAM Based on Point-Line Fusion [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2022.
19. Newcombe R, Izadi S, Hilliges O, et al. KinectFusion: Real-Time Dense Surface Mapping and Tracking [C]//Proceedings of the IEEE Symposium on Mixed and Augmented Reality. IEEE, 2012.
20. Engel J, Koltun V, Cremers D. Direct Sparse Odometry [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 48(1): 1-11.
21. Wu Yuxiang, Wang Chao, Xian Yingxian. SLAM for Mobile Robots Based on Sparse Direct Method and Graph Optimization [J]. *Journal of Instrumentation*, 2018, 39(4): 257-263.