Proceedings of the 2023 International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/12/20230294

The application of SLAM technology in indoor navigation to complex indoor environments

Yifan Hu

Collage of Overseas Education, Nanjing Technology University, Nanjing, 210014, China

202021138053@njtech.edu.cn

Abstract. This paper reviews the application progress of Simultaneous localization and mapping (SLAM) in complex indoor environments. SLAM is a technology used to manufacture mobile robots and autonomous vehicle. It can achieve autonomous localization and mapping in unknown environments. In dealing with complex indoor environments, SLAM technology faces many challenges, such as missing local perspectives, detecting moving objects, constructing dense maps, and real-time requirements in large-scale environments. However, with the development of technology, more and more SLAM technologies are being applied to complex indoor environments, among which the fusion of multimodal vision and deep learning technology, enhanced camera positioning technology, and navigation algorithms based on intelligent platforms are currently relatively advanced technologies. This article will briefly introduce the main development process of SLAM technology in dealing with complex indoor environments and the application of deep learning technology. By utilizing this technology, map features can be extracted more accurately, objects can be recognized, and obstacle avoidance can be achieved. This provides a good development direction for SLAM technology.

Keywords: simultaneous localization and mapping, SLAM technology, indoor environment, multi-sensor fusion, deep learning.

1. Introduction

In recent years, the demand for autonomous mobile robot and autonomous vehicle systems is increasing, and SLAM technology is one of the key technologies to realize autonomous localization and mapping of these systems. However, in complex indoor environments, autonomous mobile robots and autonomous vehicles face many challenges, such as occlusion, shading, relief of terrain, non-rigid objects, etc., which bring great challenges to SLAM technology. Therefore, how to deal with these challenges and improve the robustness of SLAM technology in indoor environment has become one of the research hotspots of SLAM technology.

Over the past few decades, SLAM has undergone several important developments that have moved it from a laboratory environment to a practical application. Traditional SLAM technology is mainly based on Lidar or camera data algorithm, which is described in detail by Wei and Ebel [1, 2]. However, with the emergence of a variety of sensor technologies, such as inertial navigation, GPS, sonar, etc., multi-sensor fusion technology continues to develop, further improving the performance and accuracy

DOI: 10.54254/2755-2721/12/20230294

of SLAM system. At the same time, the emergence of deep learning technology also brings new ideas to SLAM system, and how to apply deep learning technology to SLAM system has become a research hotspot.

In this article, we will focus on the application of SLAM technology in complex indoor environments and introduce some new SLAM research. These studies propose some optimization and improvement methods to address the challenges faced in complex environments, aiming to improve the accuracy and robustness of SLAM localization and mapping in unknown environments.

2. Multi-sensor fusion technology

SLAM multi-sensor fusion technology is a technology that fuses data from multiple sensors to improve the performance, accuracy, and robustness of the SLAM system. Traditional SLAM technology mainly relies on data from a single sensor (such as lidar and camera), but these sensor data may have shortcomings, such as occlusion, missing data, noise, and other problems, which will affect the performance and accuracy of SLAM system.

The emergence of multi-sensor fusion technology aims to integrate multi-sensor data and make up for the limitation of a single sensor. At present, SLAM multi-sensor fusion technology usually includes camera, lidar, IMU, GPS and other sensors. These sensors can complement each other to provide more comprehensive and accurate sensor data.

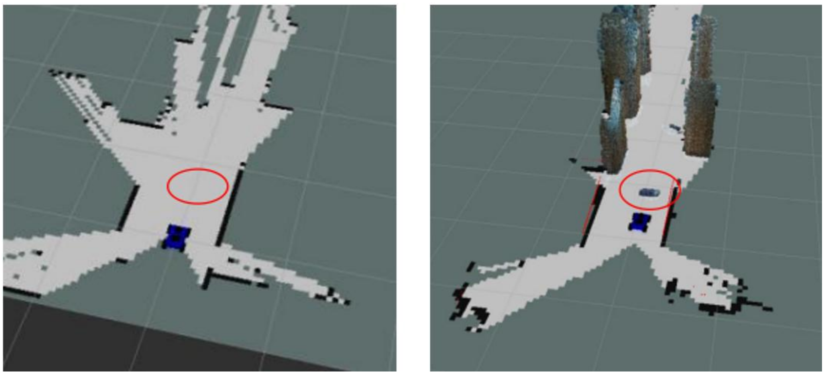
In multi-sensor fusion technology, data fusion is usually divided into the following steps:

1. Data alignment: The data of multiple sensors are aligned in time and space to ensure that subsequent data fusion can be performed correctly.

2. Data fusion: Fusion of data from different sensors, such as lidar and camera data to identify feature points and fuse these into a complete map, or IMU data integrated into lidar and camera data to provide location information for robot trajectory estimation.

3. Data filtering: In order to avoid noise and instability caused by multi-sensor data, the filtering technology is often used to smooth the data after fusion.

Guan introduced the basic principle of multi-sensor fusion technology in detail in the paper, including the steps of data alignment, fusion, and filtering [3]. Aiming at the common problems of autonomous mobile robot localization and mapping in complex indoor environment, a multi-sensor fusion localization mapping scheme is proposed. This scheme uses Lidar, camera, IMU, GPS and other sensors for data fusion, which can improve the performance and accuracy of the system, and can independently establish accurate indoor maps (as shown in Figure 1), but there are still real-time and precision problems.



(a) **(b)**

Figure 1. Compare the results of experimental mapping [3]. (a) Lidar SLAM; (b) multi-sensor fusion SLAM.

Deng introduced the existing SLAM technology and multi-sensor fusion technology, and then introduced the technical route and implementation steps of this method in detail, including data preprocessing, data association and map construction [4]. Among them, data preprocessing mainly includes correction, filtering, and noise reduction steps; Data association is mainly realized by deep neural network. Map construction is realized by probability graph model. The experimental results show that the proposed method can realize the robot localization and map construction quickly, and achieve good accuracy and robustness. At the same time, the author also points out that there are still some problems in this method, such as high requirements for data synchronization of sensor and difficulty of camera data processing, which need to be further improved and optimized.

In terms of coping with indoor dynamic environment, literature proposes a real-time VSLAM method named DIR-SLAM, which is used to remove interference in dynamic environment and improve the robustness and reliability of the system [5]. This method uses the idea of event graph to convert visual data into event signals, which are used together with filters to remove dynamic interference. In addition, a new evolutionary algorithm is proposed to optimize the system performance. It can provide direction for the solution of indoor dynamic environment.

3. Deep Learning Techniques

SLAM Deep learning technology is a new SLAM technology developed in recent years. It introduces deep learning technology into SLAM to process sensor data such as vision and Lidar, to realize more accurate and robust indoor positioning and map construction. The current popular SLAM deep learning techniques mainly include the following aspects:

3.1. Visual SLAM deep learning techniques

Camera data is processed by deep learning, including image feature extraction, attitude estimation and scene reconstruction, to realize indoor scene positioning and map construction.

3.2. Semantic SLAM technology based on deep learning

Combined with deep learning method, multi-sensor data such as cameras and lidar can be used to effectively deal with indoor dynamic environment and complex scenes, and realize indoor object recognition, tracking and modelling.

3.3. Exploration techniques based on reinforcement learning

This paper mainly introduces the development of semantic SLAM technology based on deep learning in dealing with complex indoor environment.

Literature proposes a new semantic optimization method. In this method, deep neural networks are used for semantic segmentation of objects in the scene, which is combined with the feature matching process in SLAM [6]. Finally, an experimental verification was carried out based on KITTI data set. The experimental results show that feature SLAM using this method has better robustness and accuracy, and can achieve high-quality and accurate track and map reconstruction in large scale indoor and outdoor environments. It proposes a new feature SLAM method, which combines semantic information with feature matching process, improves the accuracy and robustness of SLAM in indoor and outdoor scenes through adaptive updating of the optimizer, and provides a new idea and method for the development of semantic SLAM in the future. But there are still some limitations when dealing with dynamic environments.

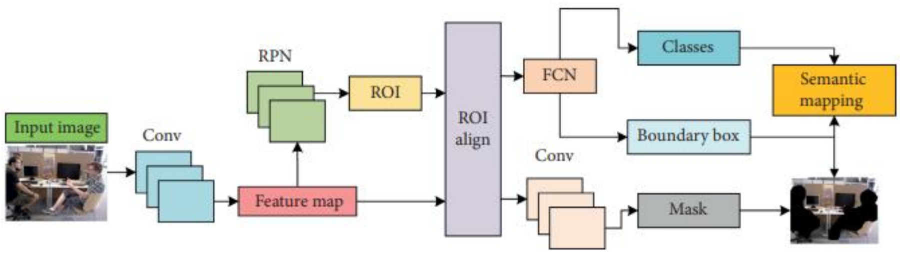


Figure 2. Pipeline of dynamic object detection using MASK-RCNN [5].

DOI: 10.54254/2755-2721/12/20230294

Zhao introduces a visual semantic SLAM method called OFM-SLAM, which is used to deal with dynamic indoor environments [7]. In this method, Faster R-CNN network is used for target detection, feature pyramid network (FPN) is used for semantic segmentation, and ORB-SLAM algorithm is used for visual SLAM. It also describes how the OFM-SLAM approach deals with dynamic environments, including eliminating the effects of dynamic objects through a dynamic object removal algorithm, and improving the accuracy of localization by integrating localization information from other sources. Experimental results in a dynamic indoor environment show that the OFM-SLAM method can accurately track robot movements and efficiently detect various objects, as shown in Figure 2.

Zhu and Wang both proposed a visual semantic navigation method for indoor mobile robots based on deep learning [8, 9]. Zhu introduced a visually-based semantically untraced FastSLAM method. The Unscented Kalman Filter is adopted, and RGB-D cameras for visual SLAM and semantic classification at the same time [8]. At the same time, ORB-SLAM algorithm is used for feature extraction and tracking, and untraced Kalman filter is used for location estimation. At the same time, the objects around the robot are classified and mapped by classifying and labeling semantic information. Wang's method consists of three main steps: semantic map generation, state representation and path generation [9]. Semantic map generation uses full convolutional neural networks (FCN) to achieve semantic segmentation and convert RGB images into semantic maps. State representation the convolutional neural network (CNN) is used to convert the image into state representation, and path generation is realized by using deep reinforcement learning algorithm. To sum up, indoor mobile robot navigation has a wider application potential for robot navigation and environment perception.

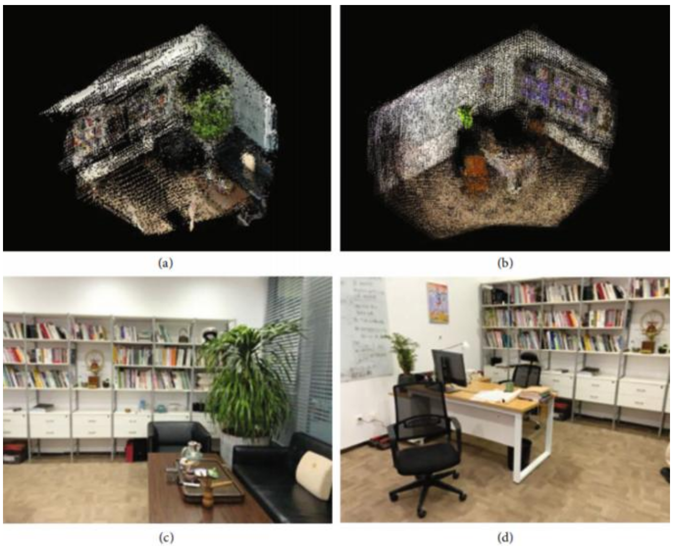


Figure 3. The semantic 3D point cloud maps of real environments. The figures in (a, b) are the semantic 3D point cloud maps. The figures in (c, d) are the color images of experimental scenes. [10]

At the same time, deep learning technology and multi-sensor fusion technology can be combined. You et al[10] introduces a multi-modal semantic SLAM method, which can realize real-time localization and map construction in a dynamic environment. MISD-SLAM method uses both visual sensors and inertial measurement units as auxiliary inputs. This method introduces semantic information into the visual-inertial SLAM framework, uses deep neural networks to conduct semantic segmentation of RGB images, and combines semantic information with positioning and mapping processes to improve the reliability of the system. In addition, the inertial meter information is used to estimate the robot's speed and attitude, thus further improving the positioning accuracy. In summary, MISD-SLAM is an innovative multi-modal semantic SLAM method with excellent performance in dynamic environments. This method not only considers the visual and inertial information, but also introduces the semantic information to expand the application scenarios of the robot in the complex indoor environment, as shown in Figure 3.

4. Conclusion

In summary, SLAM technology has made significant progress in solving problems. Although the development of this technology still has limitations, through the continuous enhancement of sensors, improvement of algorithms, and integration of intelligent technologies, this article introduces the fusion of multimodal vision and deep learning technology. Currently, there are more advanced technologies such as enhanced camera positioning technology. With the continuous development and improvement of large-scale SfM and 3D reconstruction technology, enhanced camera positioning technology has been applied in many indoor environments, and achieved good results. This technology combines SLAM technology and visual odometer technology. By overlaying multiple maps, incremental learning, and other means, more accurate and dense maps can be constructed, achieving more efficient indoor environment navigation. Navigation algorithms based on intelligent platforms: With the gradual maturity of intelligent platforms, the artificial intelligence algorithm libraries, machine learning tools, deep neural networks and other resources provided by them provide more possibilities for the research and application of SLAM. The Jizhi Intelligent Platform provides high-performance GPU and CPU support, which can accelerate the training and optimization of various machine learning algorithms, providing strong support for robot navigation and path planning in complex indoor environments. Therefore, the future development prospects of SLAM technology are still very broad and will continue to provide effective solutions for various application scenarios in the process of sustainable development. In indoor environments, SLAM technology can assist robots in tasks such as map construction, positioning, obstacle avoidance, and path planning, providing important technical support for robot navigation. At the same time, it also has broad application prospects in fields such as human-computer interaction, logistics transmission, warehouse management, and security monitoring. In addition, SLAM technology has other special applications, such as automatic guided robots in hospitals, control systems in smart homes, and a wider range of fields such as intelligent manufacturing and intelligent warehousing. With the continuous progress of technology and the expansion of application scenarios, it is believed that SLAM technology will be increasingly valued and recognized in the field of indoor navigation and modeling.

References

[1] S. Wei, F. Pang, Z. Liu et al. ‘A Review of Simultaneous Localization and Mapping Method Based on Laser Radar’, Computer application research, vol. 2020, pp. 327-332.
[2] Esidine Ebel. ‘Research Based on Laser Radar and Monocular Camera Heterogeneous Sensor Data Fusion Technology [D]’. Tianjin university, vol. 2021, pp. 1-75.
[3] C. Guan. ‘Built Figure Positioning and Navigation Based on Multi-sensor Fusion [D]’, Harbin Institute of Technology, vol. 2021. pp. 1-81
[4] S. Deng, C. Guo. ‘Research on Real-time Location and Map Construction Method Based on Multi-sensor Fusion [C]’, Academic Exchange Center of China Satellite Navigation System Administration Office. Proceedings of the 11th Annual China Satellite Navigation Conference -- S13 Autonomous Navigation. vol. 2020:6. pp. 1-6.
[5] X. Ma et al., ‘DIR-SLAM: Dynamic Interference Removal for Real-Time VSLAM in Dynamic Environments’, Mobile Information Systems, vol. 2023, pp. 1-12, Feb. 2023, doi: 10.1155/2023/1145346.
[6] P. Li, L. Yin, J. Gao, and Y. Sun, ‘Semantic Optimization of Feature-Based SLAM’, Mathematical Problems in Engineering, vol. 2021, pp. 1-10, Apr. 2021, doi: 10.1155/2021/5581788.
[7] X. Zhao, T. Zuo, and X. Hu, ‘OFM-SLAM: A Visual Semantic SLAM for Dynamic Indoor Environments’, Mathematical Problems in Engineering, vol. 2021, pp. 1-16, Apr. 2021, doi: 10.1155/2021/5538840.

DOI: 10.54254/2755-2721/12/20230294

[8] X. Zhu, F. Deng, Y. Ou, L. Liu, and E. Wang, 'Vision-Based Semantic Unscented Fast SLAM for Indoor Service Robot', Mathematical Problems in Engineering, vol. 2015, pp. 1-10, 2015, doi: 10.1155/2015/149206.
[9] L. Wang et al., 'Visual Semantic Navigation Based on Deep Learning for Indoor Mobile Robots', Complexity, vol. 2018, pp. 1-12, 2018, doi: 10.1155/2018/1627185.
[10] Y. You, P. Wei, J. Cai, W. Huang, R. Kang, and H. Liu, 'MISD-SLAM: Multimodal Semantic SLAM for Dynamic Environments', Wireless Communications and Mobile Computing, vol. 2022, pp. 1-13, Apr. 2022, doi: 10.1155/2022/7600669.