International Journal of Multidisciplinary Research and Publications
ISSN (Online): 2581-6187

A Review on Instant Localization and Map Building for Mobile Robots with Multi-Sensor Fusion

YiMeng Zheng
School of Management Science and Engineering, Anhui University of Finance and Economics
Email address: 3188563957@qq.com

Abstract**—*The main objective of this review is to systematically sort out and analyze the existing multi-sensor fusion SLAM techniques and explore the impact of different sensor combinations and fusion strategies on SLAM performance. This paper will assess the limitations of the current techniques and suggest possible directions for improvement, provide a comprehensive overview of the techniques, analyze key algorithms and methods, as well as provide guidance and recommendations for future research.*

Keywords**—*Multi-sensor fusion SLAM data fusion.*

I. INTRODUCTION

With the rapid development of automation and intelligent technologies, mobile robots have shown their unique potential for applications in numerous fields, such as self-driving cars, service robots, industrial automation, and exploration tasks. In these applications, Simultaneous Localization and Mapping (SLAM) technology plays a crucial role. SLAM allows robots to navigate autonomously in unknown environments, which is achieved by constructing a map of the environment in real time and localizing their position.[1]

SLAM is an algorithm that integrates the two key tasks of Localization and Mapping. It enables a robot moving through an unknown environment to simultaneously map the surroundings and figure out where it is in the map. The goal of SLAM is to achieve two main functions: one is to provide an accurate model of the environment, and the other is to provide the robot's precise position in that environment. This is critical for autonomous robot navigation and decision making. However, a single sensor is often limited by its own characteristics; viewing angle, resolution, or environmental adaptation can affect the final result, which limits the performance and reliability of SLAM systems.

Multi-sensor fusion methods can combine the strengths of different sensors and compensate for their respective shortcomings, thus improving the robustness, accuracy and coverage of the system. For example, laser radar (LiDAR) provides accurate distance measurements, while cameras are able to capture rich color and texture information. By fusing this data, SLAM systems are able to more accurately understand and adapt to the environment in which they operate.

II. SENSORS

A. *Types of sensors*

Multi-sensor fusion SLAM systems are capable of integrating multiple types of sensors to improve the accuracy and robustness of localization and map building. These sensors include laser radar (LiDAR), which excels in indoor and outdoor especially structured environments by providing accurate distance and angle measurements. Cameras, on the other hand, capture rich visual information, including color and texture, and are well suited for feature recognition and scene understanding. Inertial Measurement Units (IMUs) provide information about the robot's acceleration and angular velocity, which helps to estimate the robot's motion in the absence of external visual information. Global Positioning System (GPS) provides global positioning information in outdoor environments, although it can be affected by occlusion and signal fading. Ultrasonic sensors are less expensive and suitable for short-range measurements, but may be limited in complex environments. In addition, other types of sensors, such as radar and tactile sensors, may be selected depending on the needs of a particular application. By fusing data from these sensors, the SLAM system is able to perceive the environment more comprehensively and improve its navigation capabilities.[2]

B. *Integration Strategy*

Multi-sensor fusion strategy plays a crucial role in SLAM systems, and its core objective is to enhance the system's localization accuracy and map-building robustness by integrating data from different sensors. The fusion strategy is divide into three main levels:

The first is data-level fusion, a strategy that directly integrates raw data from different sensors at the lowest level, e.g., combining images captured by a vision sensor with distance measurements from a LiDAR. This requires precise synchronization of the timestamps and spatial coordinates of each sensor to ensure temporal and spatial consistency of the data. Data-level fusion has the advantage of being able to directly utilize the raw information from the sensors, but differences in data format and resolution must also be addressed.[3]

Feature-level fusion sits at an intermediate level and involves converting data from different sensors into a uniform representation of features, such as point clouds, edges, corner points, or lines. This conversion allows the features to have a consistent format and resolution for easy integration and processing. Feature-level fusion allows the system to capitalize on the strengths of each sensor and extract critical environmental information while reducing data redundancy and noise.

Finally, there is decision-level fusion, which is the highest level of fusion strategy, in which each sensor performs the

SLAM task independently, generating its own localization and map estimates. These independent estimates are then integrated through voting, weighting, or other decision rules to form the final output. This approach allows the system to combine the judgments of different sensors in the final decision-making phase, thereby improving overall accuracy and robustness.

Every fusion strategy has specific benefits and drawbacks. As technology continues to advance, new fusion strategies and algorithms are being developed and optimized to adapt to increasingly complex and dynamic environmental conditions.

C. Fusion Algorithm

Designing multi-sensor fusion algorithms is a complex task that requires consideration of data synchronization, consistency and complementarity to ensure effective integration of sensor information. Common design approaches include:[4]

Kalman filters and their variants, such as EKF and UKF, integrate observations through prediction and update steps. The prediction is based on the last state and system model, and the update step corrects the prediction using the current observation to achieve the optimal state estimation.

Graph optimization methods construct a global cost function, use the sensor measurements as constraints, and integrate the measurements by solving for the minimum using nonlinear optimization techniques such as gradient descent or Newton's method.

Machine learning, especially deep learning, uses neural networks to learn complex relationships between sensor data, learn potential patterns from large amounts of data, and implement advanced fusion strategies.

The design also needs to consider data synchronization, which ensures temporal synchronization; data consistency, which deals with scale, coordinate system, and time-stamp differences; and data complementarity, such as vision sensors providing environmental information and LiDAR providing distance measurements.

Combining these methods and factors, robust, accurate, and efficient multi-sensor fusion SLAM systems can be designed. Technological advances will enable future algorithms to be more intelligent, automated, and adaptable to complex dynamic environments.

III. MULTI-SENSOR FUSION SLAM APPROACH

A. Filter-based methods

Filtering-based methods are a common technique in SLAM, which utilize probabilistic filters to estimate the robot's dynamic state and environment maps. These methods achieve continuous estimation of robot position and environment by considering information from different sensors and integrating them into the filtering process. Following are a few common filter-based SLAM methods:

1. Kalman Filter: this is a linear optimal estimation method for the case where both the system model and the noise are linear and Gaussian distributed. The Kalman filter updates the state estimates by means of resolution, which minimizes the variance of the estimation error at each time step.

2. Extended Kalman Filter (EKF): The classical Kalman filter is no longer useful in nonlinear systems. the EKF extends the scope of the Kalman filter by linearizing the nonlinear function at each time step (usually using the first-order approximation of the Taylor series expansion) and transforming the problem into a linear one.

3. Untraceable Kalman Filter (UKF): The UKF is a more advanced filtering method that uses a specific set of points (called sigma points) to approximate the propagation of a nonlinear function. These points capture the probability distribution of the nonlinear system by weighted averaging, thus avoiding the linearization error in the EKF.The UKF is suitable for more complex nonlinear systems and can often provide more accurate estimates than the EKF.

These filter-based techniques all have benefits and drawbacks. Kalman filters are very efficient when dealing with linear systems, but may not be accurate enough when confronted with nonlinear systems.Both EKF and UKF are designed to solve nonlinear problems, but they differ in terms of computational complexity and difficulty of implementation. In practical applications, the selection of an appropriate filtering method needs to consider the characteristics of the system, the accuracy of the sensors, the limitation of computational resources, and the real-time requirements.

With the development of SLAM technology, these filtering-based methods are constantly being improved and optimized to adapt to more complex and dynamic environments. Meanwhile, emerging technologies such as deep learning and graph optimization are being combined with filter-based methods to further enhance the performance of SLAM systems.

B. A graph-based approach

Graph-based methods in SLAM are an effective way to transform a problem into a graph optimization problem, which usually use the principles of graph theory to express and solve the robot's localization and graph building problems. The following are two common graph-based approaches to SLAM:

1. Factor graph: in this approach, the graph consists of nodes and factors. The nodes represent the states of the robot (e.g., position and orientation), while the factors represent the measurement models and constraints, which connect related nodes. Factor graphs are particularly well suited to represent latent variable and conditional independence in probabilistic graphs, and are able to construct sparse graph structures, which helps to reduce the amount of computation and increase the efficiency of the solution. By optimizing this graph, the best state estimate that satisfies all constraints can be found.

2. Keyframe-based approach: this approach constructs maps and trajectories by selectively using keyframes. Key frames are representative frames extracted from the sensor data and they contain important information about the environment. By using only these keyframes, the amount of data to be processed can be reduced, thus reducing computational complexity. At the same time, this approach is able to maintain the consistency and accuracy of the map because it

maintains the global consistency of the map through constraints between keyframes.

The advantage of graph-based SLAM approaches is their ability to naturally handle multi-sensor data fusion, adaptation to dynamic environments, and representation of uncertainty. The graph optimization framework provides a flexible way to integrate different types of constraints and measurements, enabling SLAM systems to be more robust and accurate. However, these approaches also face the challenges of how to design effective optimization algorithms, how to handle large-scale graphs, and how to guarantee real-time performance. With the improvement of computational power and the development of optimization algorithms, the application of graph-based methods in the field of SLAM is promising.

C. *Machine learning based approach*

With the rapid development of deep learning technologies, machine learning-based approaches are becoming increasingly popular in the field of SLAM (Simultaneous Localization and Mapping). These methods utilize the powerful feature extraction and learning capabilities of deep learning models to improve the performance and efficiency of SLAM systems.

1. Deep SLAM: In this approach, Convolutional Neural Networks (CNNs) are used to automatically extract key features in the image, which is more efficient and robust than the traditional hand-designed features.CNNs are able to capture hierarchical information in the image, which provides richer and more accurate feature descriptions for SLAM. In addition, Recurrent Neural Networks (RNN) are also used to process time-series data, such as video streams or sensor data streams, to capture temporal correlations in the data.
2. End-to-end SLAM: The core of this approach lies in training the network directly from raw sensor inputs to the final trajectory and map outputs, omitting many of the intermediate steps in the traditional SLAM process, such as feature extraction, data correlation, and back-end optimization. End-to-end SLAM is typically implemented by training a deep neural network that learns complex mapping relationships from inputs to outputs. The advantages of this approach are simplified system design, reduced manual intervention, and the ability to perform end-to-end optimization with large amounts of data.

The application of deep learning methods in SLAM not only improves the accuracy and robustness of the system, but also provides new possibilities for dealing with large-scale and complex environments. However, these methods also face some challenges, such as the dependence on a large amount of training data, the generalization ability of the model, and the demand for computational resources. Future research will continue to explore how deep learning and SLAM can be combined more effectively to achieve more intelligent and automated navigation and mapping systems.

IV. INTEGRATION STRATEGY

In a multi-sensor fusion SLAM approach, the fusion strategy plays a crucial role in determining how the data from different sensors are integrated to improve the accuracy and robustness of localization and map construction. The fusion strategy is categorized into the following three main levels:

A. *Data-level fusion*

This Strategy involves the direct integration of raw data collected from different sensors. This means that the data from each sensor is combined at the most basic level in order to be able to utilize the strengths of each and to reduce the possible errors or uncertainties of a single sensor.

B. *Feature-level fusion*

At this level, the data from each sensor is first converted into a common representation of features, such as point clouds, edges or corners. These features are then processed in a uniform manner in order to achieve data fusion at the feature level. This approach allows the system to be more flexible in dealing with the characteristics of different sensors and may improve the consistency of feature extraction.

C. *Decision-level fusion*

in this strategy, each sensor performs the SLAM task independently, generating its own localization and map information. These independent results are then integrated through voting, weighting, or other decision rules to form a consolidated output. This approach allows the system to utilize the judgments of different sensors in the final decision-making phase to improve overall accuracy and robustness.

Choosing the appropriate fusion strategy depends on the specific application scenario, the characteristics of the sensors, and the performance requirements of the system. Each strategy has its advantages and limitations, so these factors need to be carefully considered when designing a multi-sensor fusion SLAM system to achieve optimal performance.

V. INTEGRATION STRATEGY

A. *Data-level fusion*

SLAM technology has made significant progress, but it still faces a number of limitations and challenges. First, many SLAM systems are highly dependent on specific types of sensors, which limits their ability to be applied in diverse environments. Second, moving objects and changing conditions in dynamic environments pose significant challenges to the stability and accuracy of SLAM systems, requiring systems that can quickly adapt to these changes. In addition, real-time SLAM algorithms need to operate with limited computational resources, which limits the complexity and accuracy of the algorithms and requires algorithm design that is both efficient and accurate.

The complexity of sensor fusion is also an issue; effectively fusing data from different sensors requires solving data synchronization, calibration, and consistency problems, which increases the complexity of the system. Finally, the robustness of existing SLAM systems in the face of sensor failures, extreme environments, or incomplete data still needs to be improved. This means that systems need to be able to better handle uncertainties and anomalies to ensure reliable service under all conditions. Addressing these challenges will be key to driving the further development of SLAM technology.

YiMeng Zheng, “A Review on Instant Localization and Map Building for Mobile Robots with Multi-Sensor Fusion,” *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 6, Issue 12, pp. 177-180, 2024.

A. Future technology trends

The future development of SLAM technology is moving towards several distinct trends. First, multimodal sensing fusion will become a key direction, where system robustness and accuracy can be significantly improved by combining multiple types of sensors, such as vision, LIDAR, radar, and inertial measurement units (IMUs). Second, the application of deep learning and artificial intelligence will revolutionize SLAM, and feature extraction, data fusion, and decision making can be significantly improved using these advanced models. [5]In addition, the exploration of edge computing and distributed SLAM will enable algorithms to run on edge devices, thereby reducing processing latency and improving system responsiveness. Finally, the development of adaptive and self-learning capabilities will drive SLAM systems towards greater intelligence, enabling the system to adapt and optimize itself to changing conditions based on different environments. These trends signal that SLAM technology will become more powerful, flexible and intelligent to meet the demands of complex and changing applications in the future. Real-time and efficiency: optimizing algorithms to achieve faster processing speeds and lower computational costs.

B. Decision-level fusion

In the field of SLAM, we are facing several potential research opportunities, which include the development and integration of novel sensors, such as event cameras or quantum sensors, to provide richer information about the environment, as well as algorithmic innovations that are being explored to address the limitations of existing SLAM systems, such as the development of new approaches based on topology or graph theory. In addition, SLAM technology is being explored for applications in new domains, such as medical robotics, deep sea exploration, or disaster response, and these cross-domain applications present new challenges and opportunities for SLAM technology. Security and privacy protection are also important aspects of research in the SLAM field, and researchers are committed to securing SLAM systems while ensuring user privacy. Finally, the development and maintenance of open source tools and platforms in order to promote collaboration and knowledge sharing among the research community is also an important research direction in the SLAM field.

ACKNOWLEDGEMENT

This work is supported by Anhui University of Finance and Economic Undergraduate Scientific Research Innovation Fund Project under Grant No. XSKY24147.

REFERENCES

[1] LU Chunxiao, ZHONG Huan, LIU Wei, et al. Multi-sensor fusion SLAM technology in complex terrain environment[J/OL]. Robotics:1-11[2024-06-08].https://doi.org/10.13973/j.cnki.robot.230288.

[2] Yao Erliang, Song Haitao, Zhao Jing, et al. Lidar-inertial SLAM algorithm considering point cloud structure and surface information[J/OL]. Robotics:1-14[2024-06-08]. https://doi.org/10.13973/j.cnki.robot.230244.

[3] LIU Dong, YU Tao, CONG Ming, et al. A dynamic environment visual SLAM method based on deep learning image features[J]. Journal of Huazhong University of Science and Technology(Natural Science Edition),2024,52(06):156-163.DOI:10.13245/j.hust.240658.

[4] Aiqiang Ma,Stubborn Yao. Multi-sensor adaptive fusion SLAM method for underground mobile robots in coal mines[J/OL]. Industrial and Mining Automation:1-12[2024-06-08].https://doi.org/10.13272/j.issn.1671-251x.2024050031.

[5] PENG Songzheng. Application of laser SLAM scanning technology in dark pipe exclusion[J]. Value Engineering,2024,43(16):91-93.