An overview study of multi-sensor fusion for map building and localization

Abstract

This paper reviews the application of multi-sensor fusion in simultaneous localisation and map construction (SLAM) technology. With the development of robotics, autonomous driving and virtual reality, there is an increasing demand for precise localisation and map construction.SLAM technology has emerged to solve the problem of autonomous robot localisation and map construction in unknown environments. However, single-sensor SLAM systems have limitations, such as limited sensing capability and susceptibility to noise interference. Multi-sensor fusion SLAM significantly improves the performance and robustness of the system by integrating the advantages of multiple sensors.The multi-sensor fusion SLAM system includes key components such as sensor data reading, front-end visual odometry, back-end optimisation, loopback detection and map building. The sensor data is first optimised to reduce noise and then further processed according to the task requirements. Sensors are categorised into internal sensors (e.g. IMUs and wheeled odometers) and external sensors (e.g. cameras, LIDAR, UWB sensors, etc.). Common data fusion methods include filter-based fusion (e.g., Extended Kalman Filter, Particle Filter), optimisation-based fusion (e.g., Graph Optimisation, Nonlinear Least Squares) and deep learning-based fusion (e.g., Convolutional Neural Network, Recurrent Neural Network). These methods are able to handle different types of sensor data and improve the performance of SLAM systems.

Multi-sensor fusion SLAM technology has a wide range of applications in fields such as robot navigation, autonomous driving, virtual reality, geographic information systems and mapping. In the future, the technology will pay more attention to the combination of multimodal fusion and deep learning, optimising the computational efficiency of the algorithms, improving the real-time and robustness of the system, as well as the fusion of cloud SLAM and edge computing, to promote the development and advancement of the related fields.

Keyword: lidar；SLAM；Sensor Fusion；state estimate

1 Introduction

In the context of the rapid development of modern technology, robotics, autonomous driving, virtual reality and other fields have put forward higher requirements for accurate positioning and map construction technology. Simultaneous Localisation and Map Building (SLAM) technology has emerged as a key technology to address this need.The core of SLAM technology lies in solving the problem of how to accurately know where a robot is during movement (localisation) and how to describe its surroundings in detail (map building).SLAM technology allows a robot or device to autonomously localise itself in unknown environments and to build an environment map. maps, providing the basis for subsequent tasks such as navigation and path planning. However, single-sensor SLAM systems often have limitations, such as limited ability to perceive the environment and susceptibility to noise interference. To overcome these shortcomings, fused SLAM technology has emerged, which significantly improves the performance and robustness of SLAM systems by integrating the advantages of multiple sensors. Aiming at the problems of low efficiency of existing gas leakage detectors and the inability to locate the leakage source, Chen Dongyi proposed a double-vehicle linkage cycle inspection scheme based on STM32, and at the same time successfully achieved the accurate positioning of the leakage source location by using the method of multisensor fusion. In order to improve the positioning accuracy of robot outdoor long-duration positioning, Xia Linlin proposed a graph optimisation-based Global Navigation Satellite System (GNSS)/Binocular Vision/Inertial Simultaneous Localisation and Map Building (SLAM) system development and application. Line features in space are integrated into the front-end feature extraction and back-end position optimisation thread as a supplement to the geometric constraints to improve the position solving accuracy. Meanwhile, the graph structure for joint optimisation is constructed with a factor graph and a global observation error model is derived. Although single-sensor systems have their conclusions and applicable scenarios, multi-sensor fusion SLAM systems face the problem of data complexity due to the increase in the number of sensors, which makes it difficult to achieve effective integration between different data sources. To address this challenge, the data must be pre-optimised to enhance the robustness of the system. By doing so, the overall performance and reliability of the multi-sensor fusion SLAM system can be improved.

2 Research on multi-sensor fusion slam method

2.1 Overview of multi-sensor fusion slam

Multi-sensor fusion SLAM is a popular technology framework that improves the accuracy of localisation and map building by integrating data from different sensors. Such frameworks typically contain key components such as sensor data reading, front-end visual odometry, back-end optimisation, loopback detection, and map building. Compared to traditional SLAM approaches, multi-sensor fusion SLAM is particularly suitable for mobile robot applications in dynamic, low-texture, and complex environments, which often lead to image drift and excessive interference points in conventional SLAM systems.

In multi-sensor fusion SLAM, the sensor data is first optimally processed to eliminate or reduce noise. This optimised data is then further processed according to the requirements of the particular task. There are various types of sensors, which can be classified into two main categories: internal and external sensors. Internal sensors, such as IMUs and wheel odometers, are mainly used to capture the internal states of the robot, such as velocity and acceleration. Whereas external sensors, including cameras, LiDAR (Lidar), Ultra Wide Band (UWB) sensors, magnetic field meters, and manometers, are responsible for collecting information about the robot's external environment, such as position and distance.

The core advantage of multi-sensor fusion SLAM is the ability to combine the strengths of individual sensors to complement each other in order to overcome the limitations of a single sensor. Through this fusion, the system is able to sense the environment more accurately and improve robustness and reliability in complex scenarios. The application of this technology not only improves the autonomous navigation ability of robots, but also brings new possibilities for the development of the SLAM field.

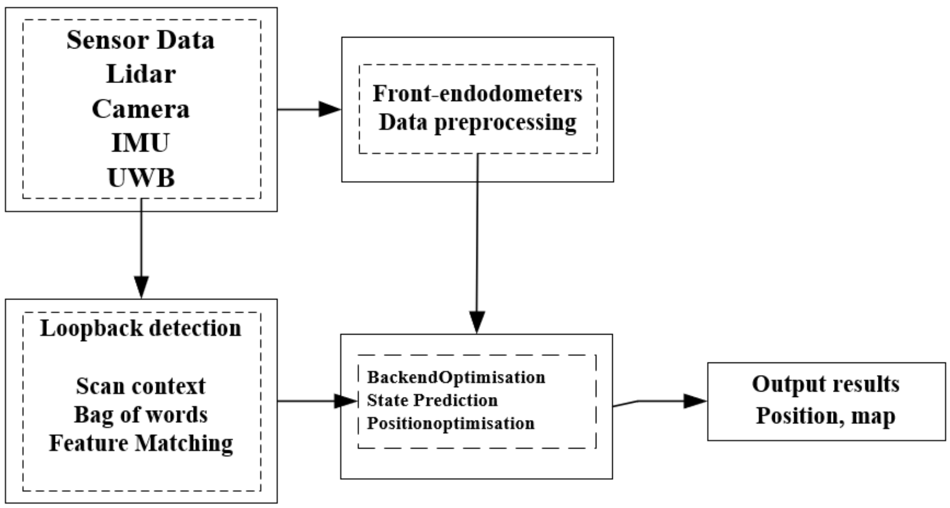

Fig. 1 Multi-sensor fusion slam architecture

2.2 Multi-sensor fusion methods

Vision-based SLAM systems are prone to problems such as degradation of positioning accuracy and failure of map construction under poor lighting conditions and lack of texture features in the scene; while LIDAR-based SLAM systems, although they perform better in structured environments, are limited in dynamic environments or complex terrains. In addition, data from a single sensor is susceptible to noise interference, leading to insufficient robustness of the SLAM system. Therefore, fusing the advantages of multiple sensors to achieve the complementary and fusion of multi-source information becomes a key way to improve the performance of SLAM systems. Multi-sensor data fusion is the core of fused SLAM technology. It involves how to effectively integrate data from different sensors to obtain more accurate and reliable positioning and map construction results. Common data fusion methods include filter-based fusion, optimisation-based fusion and deep learning-based fusion.

Filter-based fusion: e.g. Extended Kalman Filter (EKF), Particle Filter, etc. By establishing the state model and observation model of the system, real-time updating and estimation of the sensor data is carried out to achieve the optimal estimation of the system state. This method is suitable for linear or nearly linear systems, and can better deal with the noise and uncertainty of sensor data[1].

Optimisation-based fusion: e.g. graph optimisation, nonlinear least squares, etc., solves the optimal system state by constructing an optimisation problem with sensor data as constraints. This method can make full use of the correlation information between sensor data and is suitable for dealing with large-scale and complex datasets.

Deep learning-based fusion: deep learning models, such as convolutional neural networks (CNN) and recurrent neural networks (RNN), are used to extract and fuse features from sensor data to achieve perception and understanding of the environment. This approach is able to learn deeper feature representations from the data and is suitable for processing high dimensional and complex data[2].

For indoor environments that lack obvious location markers, to achieve precise positioning of UAVs, Jun Zhang proposed a multi-sensor fusion algorithm. The algorithm integrates median filtering, threshold segmentation, projection transformation, and least squares in order to achieve accurate estimation of UAV position and attitude[3].

Aiming at the problems of short-term loss of single-sensor data, low positioning accuracy, and asynchronous sensor frequency of mobile robots, Li Hang uses LiDAR, IMU, and wheeled odometer to obtain positioning information, and proposes a combined data fusion method based on Extended Kalman Filtering and Complementary Fusion. First, the initial positioning data are preprocessed by S-G filtering algorithm, and the extended Kalman filter fusion algorithm is used to achieve the fusion of IMU and wheel odometer sensors to obtain the fusion data 1, and then the complementary fusion algorithm is used to fuse the fusion data 1 and lidar to obtain the fusion positioning data 2, in which the fusion data 1 is used to make up for the real-time correction of the lidar to solve the frequency asynchrony of displacement deviation, thus significantly improving the positioning accuracy. The fusion data 1 is used to correct the LiDAR in real time to solve the frequency asynchronous displacement deviation, thus improving the positioning accuracy[4].

2.3 Multi-sensor fusion data spatio-temporal alignment

In a fused SLAM system, data from different sensors are often acquired asynchronously and there are spatial deviations between sensors. Therefore, achieving time synchronisation and spatial alignment of multiple sensors is the key to ensure the accuracy of data fusion.

Time synchronisation: Ensure that the data from different sensors are collected at the same or similar time points by means of hardware synchronisation or software synchronisation. Hardware synchronisation usually relies on devices such as synchronous signal generators, while software synchronisation is corrected by algorithms that estimate the time deviation of sensor data.

Spatial alignment: Establishing a transformation between different sensor coordinate systems to unify sensor data under the same coordinate system. This usually involves the calibration process of the sensors, through which the external references (e.g., rotation matrices and translation vectors) between the sensors are obtained to achieve spatial alignment of the data.

In order to eliminate the systematic and temporal errors of ground radar, J. Zhao proposed a spatio-temporal alignment method for 3D spatial ground radar networks based on Unscented Kalman Filter (UKF). The method combines the systematic error, temporal error and target motion model in a dynamic model and estimates them by the UKF method.Monte-Carlo simulation shows that the method can estimate the systematic error and temporal error of the radar effectively at the same time, and obtain the target trajectory at the same time.

Based on multi-sensor integration and spatio-temporal alignment technology, Wang

Zhilourealised the three-dimensional parameter acquisition of forest trees. Using the self-developed multi-sensor integration system for forest parameters, the data acquisition of forest trees was carried out in the forest sample site of Northeast Forestry University, and the spatial and temporal alignment of point cloud data and image data was realised by using the inter-sensor coordinate system conversion relationship and binocular vision algorithm. The RMSE of X, Y and Z directions were 0.074, 0.117 and 0.153 m, respectively, and the average value of RMSE was 0.115 m. The results showed that the integrated system was effective in collecting data and the alignment effect was good[5].

Pingjun Pan proposed a real-time estimation algorithm for temporal and spatial deviation of radar and infrared sensors by giving a temporal and spatial deviation alignment model. The algorithm combines the motion state of the target and the sensor deviation in the same state equation, constructs the system dynamic equation and the measurement equation of the expanded dimensional state, and through the nonlinear analysis of the measurement equation, the joint estimation of the target state and the alignment deviation is carried out by using the method of two-stage filtering of UKF and KF[6].

In the vehicle-mounted 3D measurement system of Shi Bo, in order to obtain the 3D spatial information needed in the digital city in a fast, real-time and complete way, a variety of sensors, such as the global positioning system (GPS) and the laser scanner (LS), are integrated, in which the principle of multi-sensor data fusion is applied to the processing of the data. The spatial and temporal alignment between the sensors is the key to the effective fusion of multi-sensor data information. The problem of spatial and temporal alignment is solved by solving the relationship between the coordinate systems of the sensors and by using the time recording function of the laser scanner and the GPS time marking function, which is successfully applied to the ‘near-view target three-dimensional measurement technology’, and the expected results are given in the experimental results.

3 General overview of the slam algorithm

Simultaneous localisation and mapping (SLAM) technology has been widely used in various autonomous mobile platforms, of which vision SLAM and LiDAR SLAM are the two main SLAM technology solutions. However, vision SLAM systems are susceptible to changes in the visual environment, while LiDAR SLAM systems suffer from accuracy degradation or even failure in environments such as single structure. As the application scenarios of smart mobile platforms continue to expand, higher requirements are placed on the accuracy and robustness of SLAM systems, and the fusion of multiple complementary sensors is an effective way to improve the performance of SLAM systems.

3.1. visual slam technology

Visual SLAM, thanks to its structural simplicity, cost-effectiveness and ability to extract semantic information from the environment, has become a widely used SLAM method. Although vision SLAM may encounter the problem of cumulative errors when operating in dynamic robot environments, especially when confronted with large flat areas or drastic changes in lighting conditions, its performance may suffer. However, the mainstream architectures for vision SLAM can be classified into three types,direct method, feature point method, and semi-direct method.

The direct method directly utilises the pixel intensity information obtained from the camera to estimate the camera motion and the 3D structure of the scene. Its ability to use all pixel information provides dense map building capability, which makes it perform well when dealing with environments with rich textures. Since it does not rely on feature points, the direct method maintains good performance when facing large planes or less textured environments and maintains good robustness when dealing with fast-moving scenes, since it models pixel intensity changes directly. Jakob engel proposed a direct monocular SLAM algorithm that constructs large-scale, consistent maps of the environment, performs highly accurate pose estimation in the presence of scale drift, and reconstructs 3D environments in real time using a large number of pixel-level small-baseline stereo comparison filters. The algorithm employs a direct image alignment method for high-precision pose estimation and reconstructs the 3D environment using an associated keyframe poses-graph and a semi-dense depth map[7].

The feature point method estimates the camera motion by detecting and tracking feature points in the image, which can effectively handle different lighting and occlusion situations and provide relatively high localisation accuracy. By extracting the local features of the object, the feature point method is able to achieve a stable and robust visual feature description, which is not easily affected by lighting changes. Moreover, the bag-of-words model built using feature points can perform effective loopback detection, which is one of the important roles of the feature point method in SLAM systems. However, in scenes with less texture or repetitive objects, the feature point method may encounter matching difficulties, resulting in performance degradation. And it may fail when dealing with fast moving objects or the camera's own motion, which limits its application in dynamic scenes.Raul Mur-Artal proposes the ORB-SLAM system, a robust monocular SLAM system that operates in real time in small-scale environments indoors and outdoors and is capable of dealing with severe motion clutter, supporting extensive baseline loop closure and relocation, and with fully automatic initialisation. The algorithm demonstrates the advantages of real-time, robustness, loop closure and relocation, and automatic initialisation by constructing a feature-based monocular SLAM system and testing the ORB-SLAM system in a variety of environments[8].

The semi-direct method combines the advantages of the direct and feature-point methods to reduce the computational effort by tracking blocks of pixels in an image while maintaining sensitivity to fast motion.Xiang Gao proposed a sparse odometry monocular vision SLAM system (LDSO) with closed-loop detection and bitmap optimisation. As a straightforward technique, DSO can utilise any image pixel with a sufficient intensity gradient, which makes it robust even in featureless regions.LDSO retains this robustness while prioritising corner features in the tracking front-end, thus ensuring repeatability of these points. This repeatability allows for reliable detection of closed-loop candidate points using traditional feature-based bag-of-words (BoW) methods. Closed-loop candidate points are geometrically verified and Sim(3) relative positional constraints are estimated by jointly minimising 2D and 3D geometric error terms. These constraints are fused with relative positional co-visibility maps extracted from the sliding window optimisation of the DSO. Evaluation on a publicly available dataset shows that the modified point selection strategy retains tracking accuracy and robustness, while the integrated bitmap optimisation significantly reduces cumulative rotations, translations and scale drifts, making the overall performance comparable to state-of-the-art feature-based systems[9].

3.2 Laser slam technology

Laser SLAM technology solutions can be mainly divided into two categories: 2D laser SLAM and 3D laser SLAM, each of which has different application scenarios and characteristics:

2D laser SLAM mainly uses single-line LiDAR to detect two-dimensional planar environmental information, suitable for indoor environment, famous for Gmapping, Hector-slam, cartogapher .JingRen Wen proposed a 2D LiDAR SLAM back-end optimisation method that introduces the control network constraints (CNC), in order to improve the mobile mapping accuracy that effectively solves the drift accumulation problem of front-end scan matching. The back-end of the graph-optimised SLAM is optimised by aligning the lidar scan centres with the control vertices of a pre-measured control network to optimise the bit positions of all scans and sub-maps[10][11][12].

3D laser SLAMM uses multi-line LIDAR to acquire 3D data of the environment and performs localisation through feature point matching of the 3D data.3D laser SLAM technology has the advantages of high measurement accuracy, high environmental adaptability, and easy deployment. With the mass production and performance improvement of domestic multi-line LIDAR, 3D LIDAR is gradually moving towards low-cost, low-power and high-reliability applications, and 3D LIDAR-based SLAM algorithms have been developed rapidly.RTAB-Map was initially developed as an appearance-based closed-loop detection method with memory management for dealing with large-scale and long-time online operations. Subsequently, it evolved to enable simultaneous localisation and map building (SLAM) on a wide range of robotic and mobile platforms[13].

Vision SLAM and laser SLAM are undoubtedly integral and important components of sensor fusion SLAM, and their synergy opens up new paths for autonomous robot navigation and environment sensing.

4 Mainstream sensor fusion solutions

Multi-sensor fusion SLAM technology can effectively solve the problems faced by single-sensor SLAM systems, and it improves the performance of SLAM systems by integrating data from different sensors, such as vision, LiDAR (laser radar), and inertial measurement unit (IMU). This fusion approach demonstrates its advantages in the following main areas:

Extended spatial detection range: multi-sensor systems are able to cover a wider space as different sensors can make measurements over a wider range. Improved resolution: by integrating multiple independent but qualitatively identical measurements, multi-sensor fusion SLAM is able to achieve higher resolution beyond the measurement limitations of a single sensor. Enhanced Robustness and Accuracy: Increasing the spatial dimensions of the sensors reduces the amount of interference to the system. Even if one sensor fails, the system can still maintain operation because other sensors can provide the necessary data support.

4.1 Visual inertia slam

The visual inertial slam programme combines data from visual sensors and an inertial measurement unit (IMU). The vision sensor provides rich feature information about the environment, while the IMU provides dynamic information about the robot's motion. Combining these two sensors improves the robustness and accuracy of the system, especially in dynamic and light-changing environments. Fusion of inertial and vision data is widely used to improve attitude estimation of objects. However, this type of fusion is rarely used to estimate further unknown quantities in a vision framework.GabrielNutzi proposed two different approaches to estimate unknown scale parameters in a monocular SLAM framework, which are directly related to scale by estimating the absolute velocity and position of an object in 3D space. The first method is a spline fitting task from Jung and Taylor, and the second is an extended Kalman filter, both of which have been used to analyse the behaviour of arbitrary camera paths and the quality of the resulting scale estimates by simulating them offline. The online multirate extended Kalman filter is then embedded into Klein and Murray's Parallel Tracking and Mapping (PTAM) algorithm with an inertial sensor. In this inertial/monocular SLAM framework, real-time, robust and rapidly converging scale estimation results are demonstrated. A camera and a low-cost Inertial Measurement Unit (IMU) form the monocular Vision-Inertial System (VINS), the smallest sensor suite (in terms of size, weight and power consumption) for metric six-degrees-of-freedom (DOF) state estimation.Tong Qin proposes the VINS-Mono: a robust and versatile monocular Vision-Inertial state estimator. The method starts with a robust process of estimator initialisation, and uses a tightly coupled, nonlinear optimisation-based approach to obtain a highly accurate visual-inertial odometry by fusing pre-integrated IMU measurements and eigenobservations. Combined with the tightly coupled formulation, the loopback detection module is able to achieve repositioning with minimal computational effort. A 4-degree-of-freedom position map optimisation is also performed to ensure global consistency[14][15].

4.2 Laser inertial slam

The laser-inertial slam scheme utilises data from LIDAR and IMU, with the radar providing accurate distance measurements and the IMU providing motion state information.TiXiao Shan proposed a tightly-coupled LIDAR-inertial position measurement framework, called LIO-SAM, implemented by smoothing and mapping, for highly accurate real-time mobile robot trajectory estimation and map construction.LIO- SAM formulates the LiDAR-inertial positioning problem as a factorial map, allowing the introduction of a large number of relative and absolute measurements from different sources as factors, including closed loops. Pre-integrated estimates from the Inertial Measurement Unit (IMU) de-tilt the point cloud and generate an initial guess for the lidar positioning optimisation. The acquired lidar positioning solution is used to estimate the deviation of the IMU. To ensure high performance in real-time, the old LiDAR scans are marginalised for attitude optimisation rather than matching the LiDAR scans to the global map. Matching scans on a local rather than a global scale significantly improves the real-time performance of the system, with the selective introduction of keyframes and an effective sliding window approach to register new keyframes into a fixed-size set of previous 'subkeyframes'[16].

4.3 Laser vision inertial slam

Laser vision-inertial slam enables real-time localisation of the robot and simultaneous construction of environment maps by processing data from vision, LiDAR and IMU in real-time.Jiarong Lin proposed a novel LiDAR-inertial-visual sensor fusion framework called R3LIVE, which utilises measurements from LiDAR, inertial and visual sensors to achieve robust and accurate state estimation.R3LIVE consists of two subsystems, LiDAR-Inertial Odometry (LIO) and Visual-Inertial Odometry (VIO).The LIO subsystem constructs the geometry of the map using measurements from LiDAR and inertial sensors.The VIO subsystem renders the texture of the map using visual-inertial sensor data (. Specifically, the VIO subsystem directly and efficiently fuses visual data by minimising frame-to-map photometric errors[17].

4.4 Partial fusion slam method

Table 1 Summary of multi-sensor mapping methodology options

Program	Time	Sensor	source or not	Map Type
R3LIVE	2021	monocular camera+Lidar+Imu	Yes	Point cloud
OpenRealm	2020	Vision+Lidar+Imu+Gps	Yes	Point cloud
Lvi-Sam	2020	monocular camera+lidar+imu	Yes	Point cloud
Fast-Lio	2021	Lidar+imu	Yes	Point cloud
Vins-Mono	2020	monocular camera+imu	Yes	Point cloud
Gao-S	2022	monocular camera+imu+Gps	No	Point cloud
Li-k	2021	Lidar+imu	No	Point cloud
Camvox	2021	Lidar+imu+Rgb-d	Yes	Point cloud

5 Summary

5.1 Integration challenges

Data synchronisation and fusion: multi-sensor data synchronisation is a major challenge in multi-sensor fusion SLAM. The data acquisition frequencies and timestamps of different sensors may not be consistent, which requires an accurate synchronisation mechanism to ensure data consistency. In addition, the data formats and characteristics of different sensors vary widely, and it is a technical challenge to effectively fuse these data to improve the accuracy and robustness of SLAM.

Demand for computational resources: multi-sensor fusion SLAM needs to process a large amount of data, including high-resolution images, point cloud data and IMU data. This requires the system to have high computational power for real-time processing and decision making. On resource-constrained devices, how to optimise the algorithms to reduce the consumption of computational resources is an urgent problem to be solved.

Real-time: In dynamic environments, such as self-driving cars, real-time is a key requirement for SLAM systems. Multi-sensor fusion SLAM needs to process a large amount of data in a short period of time and respond quickly, which puts high demands on the processing speed of the system.

Environmental adaptability: the adaptability of multi-sensor fusion SLAM in different environments is also a challenge. For example, in environments with low light or lack of texture, the performance of vision sensors may degrade, while the performance of LIDAR and IMU may be more stable. How to adapt the sensor fusion strategy to different environments to maintain the stability and accuracy of the system is an issue that requires in-depth research.

Sensor errors and noise: all sensors are affected by certain errors and noise, which may be amplified in the data fusion process. How to correct and compensate these errors effectively is the key to improve the performance of SLAM system[18][19].

5.2 Directions for Development

Algorithm optimisation: future research will focus on the development of more efficient algorithms to reduce the consumption of computational resources and increase processing speed. This may include improving existing filtering and optimisation algorithms, as well as developing new machine learning and deep learning techniques to process sensor data.

Hardware development: as sensor technology advances, higher performance and lower cost sensors may be available in the future. This will help improve the performance of multi-sensor fusion SLAM systems while reducing costs.

Artificial Intelligence and Machine Learning: the use of artificial intelligence and machine learning techniques to improve the adaptive and learning capabilities of SLAM systems is an important research direction. By training models to recognise and predict environmental changes, SLAM systems can better adapt to different environments and tasks.

Multimodal data fusion: future research may explore the fusion of more types of sensor data such as fused vision, LIDAR, IMU, radar and ultrasonic sensors. This will help to improve the perception of the system in complex environments.Real-time and robustness: future research will aim to improve the real-time and robustness of the system, especially in dynamic and uncertain environments[20].

5.3 Application Scenarios of Converged SLAM Technology

In the field of robotics, fused SLAM technology is widely used in robot navigation and autonomous exploration tasks. By fusing data from multiple sensors, robots are able to accurately locate and build maps in unknown environments, so that they can autonomously plan paths, avoid obstacles, and complete various tasks. For example, in service robots, fused SLAM technology can help robots navigate in complex environments such as shopping malls and hospitals and provide guide services for customers; in industrial robots, fused SLAM technology can achieve real-time perception and modelling of factory environments, improving the productivity and flexibility of robots.Autonomous driving is another important application area of fused SLAM technology. In the automatic driving system, fused SLAM technology can provide vehicles with high-precision positioning and environment sensing capabilities, so that vehicles can drive safely in complex road environments. By fusing data from on-board cameras, LIDAR, IMU and other sensors, self-driving vehicles can acquire real-time information about the surrounding environment, including road signs, traffic signals, pedestrians, vehicles, etc., so as to realise precise control of the vehicle's movement state and path planning. In the field of virtual reality (VR) and augmented reality (AR), the integration of SLAM technology can achieve real-time perception and modelling of the user's surroundings, providing the user with an immersive interactive experience. For example, in AR applications, converged SLAM technology can accurately superimpose virtual objects into the real world, enabling users to interact with virtual objects in a natural way; in VR applications, converged SLAM technology can realise accurate tracking of users' movements, providing a more realistic and smooth virtual environment experience. In the field of geographic information system (GIS) and surveying and mapping, fused SLAM technology can be used to rapidly acquire geospatial data and improve the efficiency and accuracy of surveying and mapping. By fusing data from LIDAR, IMU, GPS and other sensors, high-precision 3D modelling and topographic mapping can be carried out in complex terrains or hard-to-reach areas, providing important basic data support for urban planning, environmental monitoring and disaster assessment. In the future, fusion SLAM technology will pay more attention to the combination of multimodal fusion and deep learning. By deeply fusing sensor data from different modalities and using deep learning models to extract and analyse the features of the data, the sensing ability and intelligence level of the SLAM system can be further improved. For example, by combining multiple sensor data such as vision, LIDAR, sonar, etc., a deep learning model is used to learn the multidimensional feature representation of the environment to achieve comprehensive perception and understanding of the complex environment. With the continuous expansion of application scenarios, the real-time and robustness of the fusion SLAM system has put forward higher requirements. Future research will be devoted to optimising the computational efficiency of the algorithm and reducing the time delay of data processing and fusion to meet the real-time requirements. At the same time, the robustness of the SLAM system in various complex environments is improved so that it can better cope with challenges such as lighting changes, dynamic objects, and noise interference[21].

The convergence of cloud SLAM and edge computing will be an important development direction in the future. By combining the SLAM system with the cloud computing platform, the powerful computing and storage capabilities of the cloud platform can be used to achieve the processing and analysis of large-scale data and improve the performance and scalability of the SLAM system. At the same time, combined with edge computing technology, data preprocessing and preliminary fusion can be carried out near the data source, reducing the delay of data transmission and improving the response speed of the system.

6 Conclusion

As an important cutting-edge technology, fused SLAM technology has shown great application potential and value in many fields. Through the fusion of multi-sensor data, the fused SLAM system can achieve high-precision perception and modelling of the environment, which provides a solid technical foundation for applications such as robot navigation, autonomous driving and virtual reality. In the future, with the continuous progress and innovation of technology, the fusion SLAM technology will make greater breakthroughs in multimodal fusion, deep learning, real-time and robustness, cloud SLAM and edge computing, human-computer interaction and collaborative SLAM, etc., and promote the development and progress of related fields.

reference

[1] Chen Dongyi, Xie Xuzhe, Wen Jie, Zhang Min, Kang Zewei. Design of leakage source inspection and localisation system based on multi-sensor fusion[J]. Internet of Things Technology, 2025, 15 (01): 27-31.

[2] Xia Linlin, Song Ziwei, Fang Liang, Sun Wu Hongzhi. Development and application of GNSS/binocular vision/inertial SLAM system based on graph optimisation[J]. Chinese

Journal of Inertial Technology, 2024, 32 (05): 475-483.
[3] Jun Zhang. Research on indoor positioning technology of quadcopter UAV with multi-sensor fusion algorithm[J]. Internet of Things Technology, 2025, 15 (01): 67-71.
[4] MAO Jun, FU Hao, CHU Chaoqun, HE Xiaofeng, CHEN Changhao. A review of inertial/vision/LIDAR SLAM technologies[J]. Navigation, Positioning and Timing, 2022, 9 (04): 17-30.
[5] ZHAO Jie, JIANG Jing, TONG Qian, ZHOU Xianjing. Spatio-temporal alignment method based on UKF ground radar network[J]. Modern Radar, 2008, (06): 17-20.
[6] P.J. Pan, X.X. Feng, P. Sun, J. Liu, Y.K. Liu, F. Li. Radar and infrared alignment algorithm with UKF and KF two-stage filtering[J]. Optical Engineering, 2008, (04): 28-34.
[7] Jakob Engel, Thomas Schöps, Daniel Cremers. Lsd-Slam: Large-Scale Direct Monocular Slam[J], Lecture Notes in Computer Science, 2014, 8690(): 834-849.
[8] Raul Mur-Artal, J. M. M. Montiel, Juan D. Tardos. ORB-SLAM: A Versatile and Accurate Monocular SLAM System[J], Computing Research Repository, 2015, 31(5): 1147-1163.
[9] Xiang G, Rui W, Nikolaus D, Daniel C, et al. LDSO: Direct Sparse Odometry with Loop Closure[C], Proceedings of the ... IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2018
[10] TD Wu, CK Watanabe. GMAP: a Genomic Mapping and Alignment Program for Mrna and EST Sequences[J], Bioinformatics, 2005, 21(9): 1859-1875.
[11] Stefan K, Johannes M, Thorsten G, Karen P, Uwe K, Oskar V S, et al. Hector Open Source Modules For Autonomous Mapping And Navigation With Rescue Robots[C], Robot Soccer World Cup, 2014, 8371(): 624-631.
[12] Abdurahman D, Mohd A A R, Mohd H M A, Hairi Z, Wan M H W A, et al. Cartographer SLAM Method for Optimization with an Adaptive Multi-Distance Scan Scheduler[J], APPLIED SCIENCES-BASEL, 2020, 10(1)
[13] Mathieu Labbe, Francois Michaud. RTAB-Map As an Open-Source Lidar and Visual SLAM Library for Large-Scale and Long-Term Online Operation[J], Advances in Bioinformatics, 2019, 36(2): 416-446.
[14] Gabriel N, Stephan W, Davide S, Roland S, et al. Fusion of IMU and Vision for Absolute Scale Estimation in Monocular SLAM[J], Journal of Intelligent & Robotic Systems, 2010, 61(1-4): 287-299.
[15] Tong Qin, Peiliang Li, Shaojie Shen. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator[J], IEEE Transactions on Robotics, 2018, 34(4): 1004-1020.
[16] Tixiao S, Brendan E, Drew M, Wei W, Carlo R, Daniela R, et al. LVI-SAM: Tightly-coupled Lidar-Visual-Inertial Odometry Via Smoothing and Mapping[C], IEEE/RJS International Conference on Intelligent RObots and Systems, 2020: 5135-5142.
[17] Jiarong Lin, Fu Zhang. R3LIVE: A Robust, Real-time, RGB-colored, LiDAR-Inertial-Visual Tightly-Coupled State Estimation and Mapping Package[C], IEEE International Conference on Robotics and Automation, 2024, 46(12): 11168-11185.
[18] Grace Wahba. A Least Squares Estimate of Satellite Attitude[J], SIAM Review, 1965,

7(3): 409-409.

[19] Been K, Michael K, Luke F, John L, Abraham B, Nicholas R, Seth T, et al. Multiple Relative Pose Graphs for Robust Cooperative Mapping[C], IEEE International Conference on Robotics and Automation, 2010 Been K, Michael K, Luke F, John L, Abraham B, Nicholas R, Seth T, et al. Multiple Relative Pose Graphs for Robust Cooperative Mapping[C], IEEE International Conference on Robotics and Automation, 2010

[20] Philip Cadell, Ben Upcroft. Distributed data acquisition unit with microsecond-accurate wireless clock synchronisation.[C], Intelligent Sensors, Sensor Networks and Information Processing, 2013: 117-122.

[21] S S, E N, JH K, M R, A G, H D, et al. The Anser Project: Data Fusion Across Multiple Uninhabited Air Vehicles[J], The International journal of robotics research, 2003, 22(7-8): 505-539.