IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 January 2023, accepted 5 February 2023, date of publication 9 February 2023, date of current version 14 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3243633

RESEARCH ARTICLE

Research on Multi-Sensor Fusion SLAM Algorithm Based on Improved Gmapping

CHENGJUN TIAN, HAOBO LIU, ZHE LIU, HONGYANG LI, AND YUYU WANG
School of Electronic Information Engineering, Changchun University of Science and Technology, Changchun 130022, China

Corresponding author: Chengjun Tian (tianchengjun@cust.edu.cn)

ABSTRACT Simultaneous Localization and Mapping (SLAM) is the core technology of the intelligent robot system, and it is also the basis for its autonomous movement. In recent years, it has been found that SLAM using a single sensor has certain limitations, such as Inertial Measurement Unit (IMU) noise and serious drift, and 2D radar can only detect environmental information on the same horizontal plane. In this regard, this paper constructs a multi-sensor back-end fusion SLAM algorithm that combines vision, laser, encoder and IMU information. Experiments have proved that compared with using a single sensor, the application of a multi-sensor fusion system makes the edges of the constructed map clearer and the noise reduced. Aiming at the problem of increased calculation caused by particle degradation and too many particles, this paper improves the Gmappping algorithm, and uses the combination of selective resampling and Kullback-Leibler Distance (KLD) sampling to complete resampling. It has been proved by experiments that compared with the original algorithm of Gmapping, the application of the improved algorithm increases the particle convergence speed by 39.85% in the process of indoor mapping. Aiming at the problems that the traditional loop detection algorithm is easily affected by environmental factors, resulting in low detection accuracy, and the loop detection algorithm based on deep convolutional neural network has a large amount of calculation and takes a long time to detect. The main research of this paper is to apply a deep learning-based loop detection algorithm on the multi-sensor fusion framework, and use the combination of high-dimensional and low-dimensional features of the image for loop detection. This paper uses different algorithms to conduct comparative experiments on the dataset CityCentre. The experimental results show that compared with the traditional algorithms Bag of Words (BoW), AlexNet algorithm, VGG19 algorithm, and ResNet32 algorithm, the accuracy of the algorithm proposed in this paper has increased by 31.26%, 14.21%, 3.05%, and 1.56%, respectively. In addition, the comparison experiment results of SLAM mapping with the original Real-Time Appearance-Based Mapping (RTAB-MAP) algorithm prove that the loop closure detection algorithm based on deep learning proposed in this paper can enable the system to better build a globally consistent map, including more environmental information.

INDEX TERMS Laser SLAM, Gmapping algorithm, multi-sensor fusion SLAM system, loop closure detection.

I. INTRODUCTION

In recent years, intelligent robots have played an important role in contactless services such as material distribution, regional disinfection, and public area security patrols. Simultaneous Localization and Mapping (SLAM) as the basic technology of intelligent robots has significantly benefited from the development of artificial intelligence. There are three main types of SLAM [1] systems: laser-based SLAM systems, multi-sensor-based SLAM systems, and vision-based SLAM systems.

The classic SLAM system includes five modules: sensor data reading, front-end visual odometry, loop closure detection, nonlinear back-end optimization, and building maps. Among them, loop closure detection is an essential link in the visual SLAM system, and it is also an effective method

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

to eliminate accumulated errors. Traditional loop closure detection algorithms usually use artificially designed features, and the problems mainly concentrate on unsatisfactory accuracy and a large amount of calculation. In recent years, the research of loop closure detection algorithms based on deep learning has gradually emerged and has shown excellent performance. In summary, this paper considers integrating the deep learning-based loop closure detection algorithm into the multi-sensor SLAM system, which could improve the system's overall performance. Therefore, this paper introduces a deep learning-based loop closure detection algorithm into the multi-sensor fusion SLAM system. We used this loop closure detection method to replace the traditional loopback detection algorithm provided by Real-Time Appearance-Based Mapping (RTAB-MAP) [2] and respectively conducted SLAM mapping experiments.

Although the lidar used in laser SLAM has the advantages of high measurement accuracy, fast speed, and a simple measurement error model, it has obvious shortcomings. 2D radar [3] can only detect environmental information on the same level, and the cost of 3D lidar is too high. For other sensors, although the price of IMU is low and the frequency is high, it has high noise and severe drift. The wheel encoder has a significant error when the wheel slips or runs on uneven ground; When the camera cannot obtain depth information, there is a significant deviation in the estimation. Only using a single sensor can only meet the needs of a specific scene and cannot adequately deal with complex environments. Therefore, multi-sensor SLAM technology [4] has become a hot spot in the field of SLAM research. This technology inputs the information obtained by various sensors into the SLAM system, and uses the information fusion algorithm to use the information fusion results as the pose estimation, trajectory estimation and mapping of the mobile robot [5], so that the mobile robot has a multi-dimensional environment. Information acquisition capability, able to implement a robust SLAM system in complex environments.

The classic laser SLAM [6], [7] algorithms include the Gmapping algorithm, Cartographer algorithm, and HectorSLAM algorithm. Among them, the Gmapping algorithm based on particle filter has the following advantages:

- Ability to build indoor maps in real-time;
- When building a small scene map, the amount of computation required is small, and the accuracy is high;
- Effective use of odometer information, Etc.

However, the disadvantages of the Gmapping Algorithm will lead to particle degradation, the use of excessive particles, and an increase in the amount of calculation. The disadvantages are as follows:

- Need information from odometer;
- Not suitable for drones and rough terrain;
- No loop closure detection module;
- When building a large scene, there are too many particles and a massive amount of calculation.

Therefore, this paper improves the Gmapping algorithm and changes the original sampling method to a method that alternatively performs selective resampling and Kullback-Leibler Distance (KLD) sampling.In summary, this paper studies three multi-sensor fusion SLAM frameworks of laser and encoder; laser, encoder, and Inertial Measurement Unit (IMU) [8]; laser, encoder, IMU and vision, and conducts related experiments.

This paper contributes to practical applications. On the one hand, the Gmapping [9] algorithm can build indoor maps in real time, and requires less calculation and higher precision to build small scene maps. However, as the number of particles required increases as the scene grows, because each particle carries a map, the amount of memory and calculation required to construct a large map will increase, and the convergence speed will become slower and slower. Through the improved method in this paper, the particle convergence speed is increased, and the shortcomings of the Gmapping algorithm in large scenes are effectively overcome. And for indoor planar motion robots (such as sweepers, food delivery robots, etc.), the effect will be better, and the particle convergence speed will be faster, which can improve the cleaning speed of the sweeper and the food delivery rate. Bring better user experience to users. On the other hand, single sensors have certain limitations in application. Single-sensor SLAM systems based on lidar sensors often fail in unstructured scenarios such as long corridors or flat open spaces. The performance of vision-based single-sensor systems is sensitive to initialization, light intensity, and changes. In this paper, the fusion of lidar, camera and IMU can overcome such problems, and effectively improve the robustness, accuracy and reliability of the system. More and more multi-sensor fusion frameworks are applied to the field of robotics and autonomous driving.

In summary, the main contributions of this paper are as follows:

- The particle filter-based Gmapping algorithm has two main problems: particle degradation and excessive number of particles leading to an increase in the amount of calculation. This paper makes the following improvements to the Gmapping algorithm for the above two problems. Selective resampling and KLD sampling are alternately carried out. resampling stage. The experiment proves that the particle convergence speed of the improved Gmapping algorithm in this paper has increased by 39.85%, and a better improvement result has been obtained.
- Aiming at the limitations of a single sensor or a small number of sensor fusions, this paper studies three types of multi-sensor fusion SLAM frameworks. The experiments have fully demonstrated that the overall performance of the SLAM system with the fusion of vision, laser, encoder and IMU [10] is superior.
- As an important part of the visual SLAM system, loop detection is especially important for the SLAM system of multi-sensor fusion. In this paper, instead of the original traditional loop detection algorithm, a loop detection algorithm based on deep learning is introduced into the

multi-sensor fusion SLAM system. Experiments have proved that the improved system can better construct a globally consistent map.

The rest of this article is arranged as follows: the second part mainly introduces the related work, including the classic algorithm of laser SLAM, the current situation of multi-sensor fusion, Etc.; the third part mainly introduces the improved Gmapping method, multi-sensor fusion theory and the loop closure applied in this paper Detection method; The fourth part mainly introduces the relevant experiments and the analysis of the experimental results in this paper, and the fifth part is the conclusion.

II. RELATED WORK

A SLAM system that uses only lidar or a fusion of other sensors with lidar is called a laser SLAM system. Classic laser SLAM algorithms include: Gmapping algorithm, Cartographer algorithm and Hector SLAM. A visual SLAM system uses a camera as the main sensor. The visual SLAM scheme mainly studied in this paper is the RTAB-MAP algorithm.

A. CLASSIC ALGORITHM OF LASER SLAM

1) GMAPPING ALGORITHM

As we all know, the basic problem that SLAM needs to solve is to complete the robot's positioning and build a map of the surrounding environment simultaneously. This problem used the joint probability distribution model in probability theory to describe:

$$P\left(x_{1:t}, m|z_{1:t}, u_{1:t-1}\right) \tag{1}$$

where $x_{1:t}$ represents the pose sequence of the robot at time 1:t; m represents the map of the robot's surrounding environment; $z_{1:t}$ represents the sensor measurement data of the robot at time 1:t; $u_{1:t-1}$ represents the control data of the robot at time 1:t-1.

That is, on the premise of the current robot sensor measurement value $z_{1:t}$ and robot control data $u_{1:t-1}$, the robot pose state $x_{1:t}$ and the robot's surrounding environment map m can be expressed by the above formula get. From the relevant knowledge of probability theory, deduce the above formula as follows:

$$P\left(x_{1:t}, m|z_{1:t}, u_{1:t-1}\right) = P\left(m|x_{1:t}, z_{1:t}\right) * p\left(x_{1:t}|z_{1:t}, u_{1:t-1}\right) \tag{2}$$

Using Rao-Blackwellized Particle Filtering (RBPF) [11] based on particle filtering, we decomposed the above problem into localization and mapping. Among them, the core idea of particle filters is to use particle set characterization probability. However, the RBPF algorithm has two defects. The first point is that the algorithm uses many particles, which will cause a large amount of calculation and memory consumption. The second point is that performing resampling frequently can cause particle degradation.

Aiming at Two Defects of the RBPF Algorithm, the RBPF-based Gmapping algorithm has made two significant improvements to the RBPF algorithm. On the one hand, an improved proposal distribution is adopted to reduce the number of particles used during the run. On the other hand, selective resampling is used to reduce the resampling frequency.

- Improved proposal distribution
  Sampling can be done from a distribution to obtain the pose estimation of the robot at the next moment, which is the proposed distribution. The proposed distribution is just a surrogate for the target distribution. The difference between the two is the weight values of the particles. In order to solve the problem of too many particles and particle degradation, the Gmapping algorithm proposes to use the latest lidar observation data to improve the proposal distribution, then the proposal distribution becomes:

$$\begin{aligned} P&\left(x_t|m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1}\right) \\ &= \frac{P\left(z_t|m_{t-1}^{(i)}, x_t\right) P(x_t|x_{t-1}^{(i)}, u_{t-1})}{P(z_t|m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1})} \end{aligned} \tag{3}$$

- Selective resampling
  Resampling also determines the performance of particle filtering. The Gmapping algorithm uses an iterative calculation to make the particles approach the target distribution. The resampling stage determines whether particles with high-weight values can effectively replace particles with low-weight values. In general, the number of effective particles is used as the standard to measure the degradation degree of particle weight value, and its calculation formula is as follows:

$$N_{eff} = \frac{1}{\sum_{i=1}^{N}\left(\tilde{\omega}^{(i)}\right)^2} \tag{4}$$

Among them, $\tilde{\omega}^{(i)}$ represents the normalized weight of particle i, that is, the ratio of the target distribution and the proposed distribution. By adopting this constraint method, the number of resampling can be effectively reduced, and the particle degradation process can be slowed down.

2) CARTOGRAPHER ALGORITHM

Cartographer [12] is a classical laser SLAM algorithm based on graph optimization developed by Google. The difference between this algorithm and the particle filter-based Gmapping algorithm is that the Cartographer algorithm not only estimates the mobile robot's current pose state x_t but also the trajectory of the entire environment map construction process $x_{0:t}$. The algorithm framework uses two parts of Local SLAM and Global SLAM.

The content of Local SLAM includes: ① use the data of odometer and IMU to calculate the trajectory, and give the estimated value of the pose of the robot; ② use the estimated value of the pose of the robot as the initial value, match the data of the lidar, and update The value of the pose estimator; ③ Each frame of lidar data is superimposed after motion

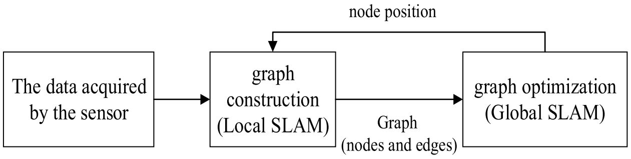

node position
The data acquired by the sensor
graph construction (Local SLAM)
Graph (nodes and edges)
graph optimization (Global SLAM)

FIGURE 1. Cartographer algorithm framework.

filtering, and then forms a submap (Submap). The content of Global SLAM includes: ① loop detection; ② back-end optimization, using all sub-graphs to form a complete and usable map.

3) HECTORSLAM ALGORITHM

The HectorSLAM algorithm [13] is very straightforward: to "align" the laser points with the existing map, that is, Scan matching. Scan matching refers to constructing an error function between the current frame and the acquired map data and using the Gauss-Newton method to obtain the optimal solution and deviation. The main work of the algorithm is to convert the laser points to the raster map. That is to say, for the matching to be successful, all laser points need to be transformed into the raster map.

B. MULTI-SENSOR SLAM

Multi-sensor SLAM usually refers to inputting the information obtained by various sensors into the SLAM system and using the results obtained by the information fusion algorithm as the pose estimation, trajectory estimation, and mapping of the mobile robot.

Hong Kong University of Science and Technology Shen Shaojie et al. [14] proposed an algorithm based on the fusion of monocular vision and IMU Visual-Inertial System Monocular (IMU - VINS-Mono). This method can calibrate the external parameter information between the camera sensor and the IMU online, and use the IMU information as the prior information of the visual odometry, and use the nonlinear sliding window optimization method to solve the pose information. The front-end odometer uses the Kanade-Lucas-Tomasi (KLT) optical flow method to track visual features, which is one of the classic algorithms for mapping and positioning algorithms based on visual-inertial navigation. Wan et al. [15] used the Error State Kalman Filter (ESKF) to fuse the IMU information from the strapdown inertial navigation solution, the lidar odometer information and Global Navigation Satellite System (GNSS) information based on the prior map. Xue et al. [16] used the fused data of IMU and wheel speedometer as the prior value of motion estimation, and then used Extended Kalman Filter (EKF) to fuse the data of laser odometer and IMU. Wisth et al. [17] proposed an efficient odometry model that fuses vision, lidar, and IMU sensors. This method proposes a new method for extracting laser features. By extracting line features and surface features in laser point clouds Meta-features, this method solves the problem that the traditional point cloud matching between frames can only obtain sub-optimal pose information, and uses factor graph optimization to jointly optimize the residual information of lidar, vision and IMU, so as to obtain pose trajectory information and create a laser point cloud map. Meng et al. [18] proposed a tight coupling of monocular vision methods and lidar odometry to extract 3D features from lidar and vision information. In this system, the monocular camera and 3D LIDAR measurement are close together for joint optimization, which can provide accurate data for Six-Degrees of Freedom (6-DOF) pose estimation preprocessing, and use the Iterative Closest Point (ICP) method to construct closed-loop constraints, and perform global pose optimization to obtain high-frequency and high-precision pose estimation.

C. DEEP LEARNING-BASED LOOP CLOSURE DETECTION METHOD

Loop closure detection is an essential link in the visual SLAM system and is an effective method to eliminate accumulated errors. Traditional loop closure detection algorithms usually use artificially designed features with poor accuracy and a large amount of computation. Research on loop closure detection algorithms based on deep learning has gradually emerged and performed well in recent years.

Li et al. [19] proposed a monocular visual odometry system—UnDeepVO. Able to estimate the 6-DoF pose of a monocular camera and its depth of view by using a deep neural network. The system utilizes spatial loss and temporal loss between stereoscopic image sequences for unsupervised training. During testing, the system can perform pose estimation and dense depth map estimation on monocular images, and what differentiates this system from other model-based or learning-based monocular visual odometry (VO) methods is that it restores scale during the training phase. Chen et al. [20] proposed a monocular visual odometry framework based on convolutional long short-term memory network (LSTM) and convolutional neural network (CNN), called LSTM visual odometry (LSTMVO). It uses an unsupervised end-to-end deep learning framework to simultaneously estimate the 6-DoF pose and depth of view of a monocular camera. It effectively solves the problem that the supervised visual odometry method requires high training data and the current unsupervised learning method fails to effectively integrate context information into pose estimation. Future system optimization can consider extending it to the visual SLAM system, and reducing the global pose drift by introducing a wider range of closed-loop constraints. Southeast University Yu Yu et al. [21] aimed at the problem that traditional loop-back detection uses artificially designed features to sharply reduce the detection accuracy in complex environments, and combined convolutional neural networks with local sensitive hashing algorithms to propose a loopback algorithm based on deep learning. Detection method. Zhou et al. [22] of Shandong University proposed a new method based on deep learning-based Local 3D Depth Descriptor (L3D) to solve problems related to loop closure detection in SLAM. L3D is an emerging compact representation of patches extracted from point clouds learned from data using deep learning algorithms. The authors propose a new overlap metric for loop closure detection, a novel approach capable of accurately loop detecting and estimating six degrees of freedom poses in the presence of small overlaps. The future optimization research direction can make a breakthrough by fusing 2D visual cues extracted from images with 3D local descriptors to deal with scenes lacking geometric structures, and focusing on the optimization of descriptor calculations. Daniele Cattaneo et al. [23] propose a novel loop closure detection(LCD)Net architecture for loop closure detection and point cloud registration, which efficiently detects Loop closures in LiDAR point clouds. LCD Net consists of a shared feature extractor based on a Point-Voxel RCNN (PV-RCNN) network, a position recognition head that captures discriminative global descriptors, and a novel differentiable relative pose head based on imbalanced optimal transfer theory, which can operate without any Effectively aligning two point clouds without prior information about their initial misalignment, experiments show that this architecture has good generalization ability.

D. RTAB-MAP

The visual SLAM scheme mainly studied in this paper is based on the RTAB-MAP algorithm [24], which is essentially a time- and scale-independent graph-based optimization algorithm, which focuses on solving the online loop detection problem when running in a large environment. At present, RTAB-MAP supports the use of various sensors to complete 2D or 3D SLAM tasks. It is a relatively complete and effective SLAM open source library.

Its core idea is that when the number of positioning points in the map makes the time to find a positioning match exceed a certain set threshold, the RTAB-MAP algorithm transfers the positioning points in the Working Memory (WM) that are less likely to form a closed loop to the LTM Long - Term Memory (LTM), the transferred positioning point will no longer participate in the next closed-loop detection calculation, and the Short-Term Memory (STM) module in the WM module is used to observe the temporal similarity of consecutive images and update the positioning accordingly Point weight. There are two key points in the loopback detection part, that is, the first key point is to take out the anchor points from the LTM module and put them back into the WM module, and the second key point is to transfer the anchor points in the WM module that meet the transfer criteria to the LTM module. RTAB-MAP uses the EKF-based multi-sensor loosely coupled data fusion method. The overall flowchart of RTAB-MAP is shown in FIG.2.

III. METHOD

This section will elaborate on the improved Gmapping algorithm, the SLAM framework for three sensor fusions, and the deep learning-based loop closure detection algorithm applied in this paper.

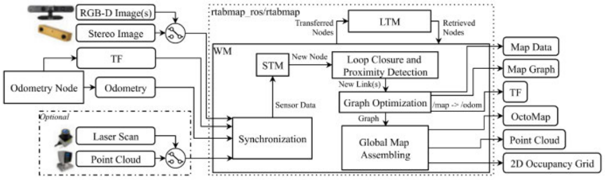


FIGURE 2. Overall flow chart of RTAB-MAP.

A. IMPROVED GMAPPING ALGORITHM

This paper chooses to improve the Gmapping algorithm as follows:

Particle filtering is an iterative calculation. Each particle carries a map, so corresponding computing power is required to support it. When building a map in a small indoor environment, the quality of the map can be guaranteed by optimizing the resampling stage, and the particles' convergence speed can be accelerated simultaneously. Particle filtering uses particles to approximate the posterior distribution; that is, the more the number of particles in the area, the more likely the actual state of the robot will fall in this area. In order to better reflect the actual situation of the robot, a large number of particles is required. So, how many particles are needed in an actual scene to be considered to reflect reality? The following calculation formula can be used to calculate the difference between the estimated probability distribution $p()$ and the actual probability distribution $q()$, namely KL-distance, and the formula is specifically:

$$K\left(p,q\right)=\sum\nolimits_{x}p\left(x\right)\log\frac{p\left(x\right)}{q\left(x\right)} \tag{5}$$

Usually, the target distribution is unavailable, but the χ^2 distribution can approximate the distance between the estimated distribution and the true target distribution. Aiming at the two problems of particle degradation and the increase in calculation amount caused by too many particles, this paper proposes to complete the resampling stage by alternatively performing selective resampling and KLD sampling. The improved Gmapping algorithm is described in Table 1 below.

B. MULTI-SENSOR FUSION FRAMEWORK

The multi-sensor fusion frameworks studied in this paper are as follows:

1) LASER AND ENCODER FUSION SLAM FRAMEWORK

The SLAM framework based on the fusion of laser and encoder is traditional. The following figure shows the specific framework. The principles of the Gmapping algorithm, Cartographer algorithm and HectorSLAM algorithm have been briefly studied in the previous section [25]. Among them, only the Cartographer algorithm includes loop closure detection. However, it only performs a graph optimization of all Submaps at intervals of a certain number of scans.

TABLE 1. Improved gmapping algorithm.

Improved Gmapping algorithm:	
***Step1*:**	Approach the motion state of the mobile robot through a large number of particles so that the particles are evenly distributed in space;
***Step2*:**	According to the state transition equation, each particle is substituted into it for calculation, and then the corresponding predicted particle is obtained;
***Step3*:**	Evaluate the predicted particles, which is calculated as the weight of the particle, and assign a larger weight to the particle that is closer to the real state;
***Step4*:**	Evaluate the predicted particles, which is calculated as the weight of the particle, and assign a larger weight to the particle that is closer to the real state;
***Step5*:**	Input the resampled predicted particles into the state transition equation to obtain new predicted particles. That is, repeat ***Step2***.

It uses the errors estimator as a constraint to optimize the estimator, which is challenging to elute with its estimation to optimize its suspicion. Neither the Gmapping algorithm nor HectorSLAM has a loop closure detection module [26]. Therefore, the loop closure detection and the subsequent calculation constraints in FIG.3 are represented by dashed line graphs.

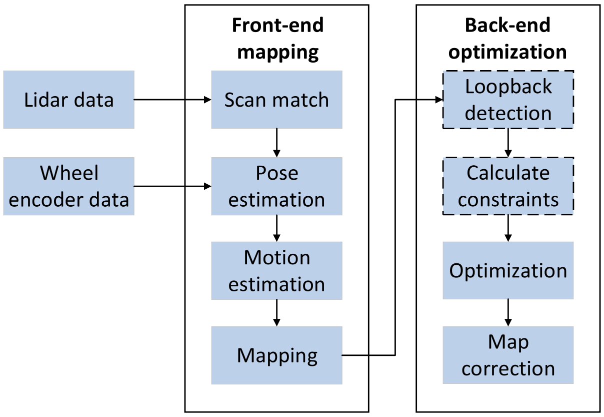


FIGURE 3. **SLAM framework based on the fusion of laser and encoder.**

The framework mainly uses the data provided by the lidar scan frame and the mileage data provided by the wheel encoder for positioning. Then it estimates the optimal pose of the mobile robot after fusion with the extended Kalman filter. Then, use this information to complete the subsequent mapping and optimization work, in which the drawn map is a raster map [27].

2) LASER, ENCODER AND IMU FUSION SLAM FRAMEWORK

The SLAM framework based on laser, encoder, and IMU adds the IMU to the system to provide acceleration and angular velocity information based on the previous section [13], as shown in FIG.3 below for a specific framework diagram. Among the three algorithms studied above, the Cartographer algorithm and HectorSLAM do not require odometry information to complete the SLAM function. In contrast, the particle filter-based Gmapping algorithm relies heavily on the information provided by the odometer. In the subsequent experiments, we use all the sensor information to complete the mapping experiment.

The framework uses the displacement, acceleration and angular velocity information provided by the wheel encoder and the IMU. It fuses the data in a loosely coupled manner through the extended Kalman filter algorithm and then derives the robot's pose information through the track. Then, the pose estimated from the data provided by the lidar scanning frame is also fused by the extended Kalman filter to estimate the current optimal pose of the robot. Finally, use this information to complete subsequent mapping and optimization work, and the drawn map is a raster map.

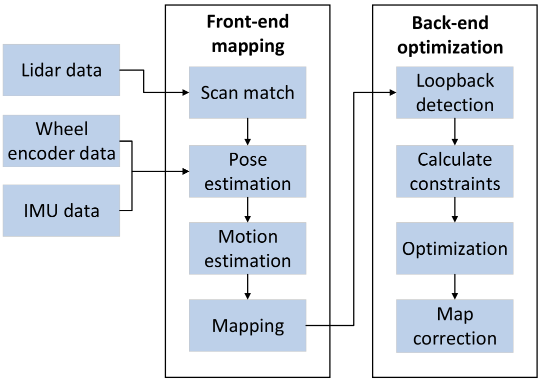


FIGURE 4. **SLAM framework based on the fusion of laser, encoder and IMU.**

3) FOUR TYPES OF SENSOR FUSION SLAM FRAMEWORK

The SLAM framework based on the fusion of vision, laser, encoder and IMU adds an RGB-D camera [28] based on the previous section, as shown in FIG.5 for the specific framework.

The framework is a back-end fusion algorithm, which essentially perceives the multi-dimensional comprehensive data after fusion. The back-end fusion algorithm is loose and is also called a loosely coupled algorithm [29]. All sensors are independent before the results are obtained, and there is no sensor-to-sensor constraint. The data from the RGB-D camera and 2D lidar are subjected to the corresponding perception algorithm to obtain the result and then aggregated and fused with the estimation results of the encoder and IMU [30]. Then the results obtained after the aggregation and fusion are transmitted to the loop closure detection and graph optimization module [31]. Finally, the system framework outputs OctoMap, point cloud, 2D occupancy raster map, map data, Map Graph and TF [32].

C. A DEEP LEARNING-BASED LOOP CLOSURE DETECTION TECHNOLOGY ADOPTED

A deep learning-based loop closure detection algorithm is adopted in this paper, which can improve the accuracy and

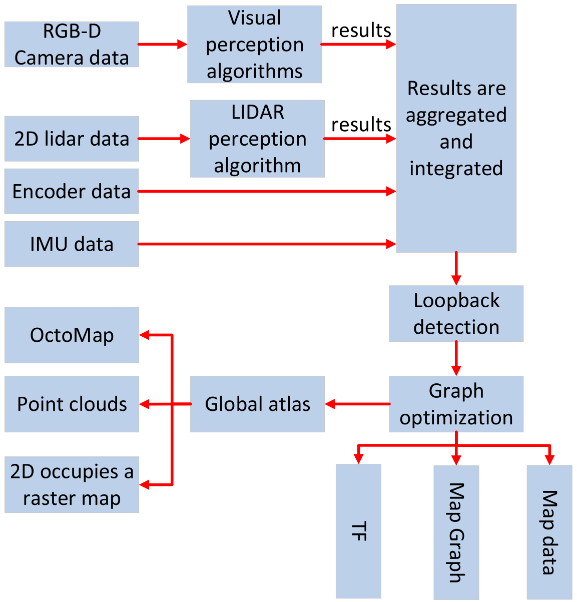


FIGURE 5. SLAM framework base on the fusion of vision, laser, encoder and IMU.

efficiency of loop closure detection, eliminate the accumulated error in the process of robot motion, and enhance the robustness of SLAM in complex environments. The algorithm includes unified image specification, keyframe selection, high-dimensional feature vector library and low-dimensional feature vector library construction, Etc [33].

The schematic diagram of the loop detection algorithm proposed in this paper is shown in Figure 6, and the specific algorithm steps are described in Table 2. Among them, high-dimensional image features are extracted by MobileNetV3-Large, while low-dimensional image features are extracted by AlexNet.

The specific steps of the loop closure detection algorithm are as follows:

- Unify the image specifications of the moving process images obtained by the RGB-D camera to obtain continuous images $p_0 \to p_n$.
- The low-dimensional and high-dimensional feature vectors of $p_0 \to p_n$ are simultaneously extracted in parallel, in which low-dimensional feature vectors are extracted by AlexNet, and high-dimensional feature vectors are extracted by MobileNetV3-Large, and *Data* (L) and *Data* (H) are established respectively.
- Use *Data* (L) to determine the key frame and its frame number used in loop closure detection through its proposed key frame selection algorithm, and record the high-dimensional feature vector corresponding to the number of keyframe frames.
- Compare the high-dimensional feature vector of the current frame p_c and the keyframe. If it is found that the similarity result compared with a keyframe is greater than α, find the previous key frame and the next key frame of this keyframe to determine the range of frames in which possibly exists the loop closure is.
- According to the range of frame numbers that may have loop closures, use the high-dimensional feature vector of p_c to compare with the corresponding feature vector of the range of frame numbers that may have loopbacks in *Data* (H). If there is a frame whose comparison result is more excellent than γ, it is determined to form a loop closure.

Description: *Data* (L) is a low-dimensional feature vector database; *Data* (H) is a high-dimensional feature vector database of keyframes; α is a preset value for the possibility of a preliminary judgment of loop closure; γ is a preset reliability parameter threshold.

Unified image specification process: ① Unify image specifications after continuous video information is processed; ② Input 227∗227 and 224∗224 images to AlexNet and MobileNetV3-Large respectively.

Key frame selection algorithm process:

TABLE 2. Key frame selection algorithm.

Directions:	$Data(H)$: High-dimensional feature vector database; β:preset similarity reference value; cosine similarity $f_{cosine}(A, B)$; reference frame P_{Lc}.
Data:	Image features in Data(L) $P_{L0} \to P_{Ln}$
Result:	Keyframe P_{Lk}

Set	P_{Lk}=P_{L0}, $P_{Lc} = P_{L1}$
	when (c <= n)
	  for (j = 1; j <= n; j++)
	  if $f_{cosine}\left(P_{Lc}, P_{L(c+j)}\right) < \beta$
	   $P_{Lk} = P_{L(c+j)}$
	   $P_{Lc} =$
	  end
	end
	return P_{Lk}

This deep learning-based loop closure detection algorithm is introduced into the multi-sensor SLAM system to replace the traditional loop closure detection algorithm provided by RTAB-MAP. The introduced framework is shown in FIG.7. Due to the limitation of the computing power of the mobile robot itself, it is necessary to share part of the computing pressure on the deep learning server mentioned above and use the topology to disperse the computing pressure.

IV. EXPERIMENT AND ANALYSIS

This section is the experiment and analysis part. Above all, the hardware and software platforms required for the experiment are introduced. First, the improved Gmapping algorithm proposed in this paper is compared with the original algorithm, and the experimental results are analyzed.

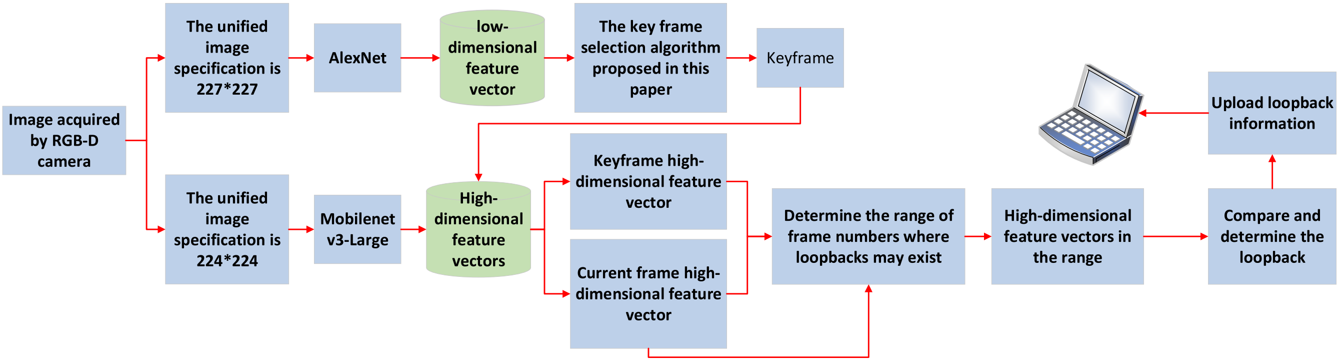


Image acquired by RGB-D camera
The unified image specification is 227*227
AlexNet
low-dimensional feature vector
The key frame selection algorithm proposed in this paper
Keyframe
The unified image specification is 224*224
Mobilenet v3-Large
High-dimensional feature vectors
Keyframe high-dimensional feature vector
Current frame high-dimensional feature vector
Determine the range of frame numbers where loopbacks may exist
High-dimensional feature vectors in the range
Compare and determine the loopback
Upload loopback information

FIGURE 6. Schematic diagram of the deep learning loop detection algorithm proposed in this paper.

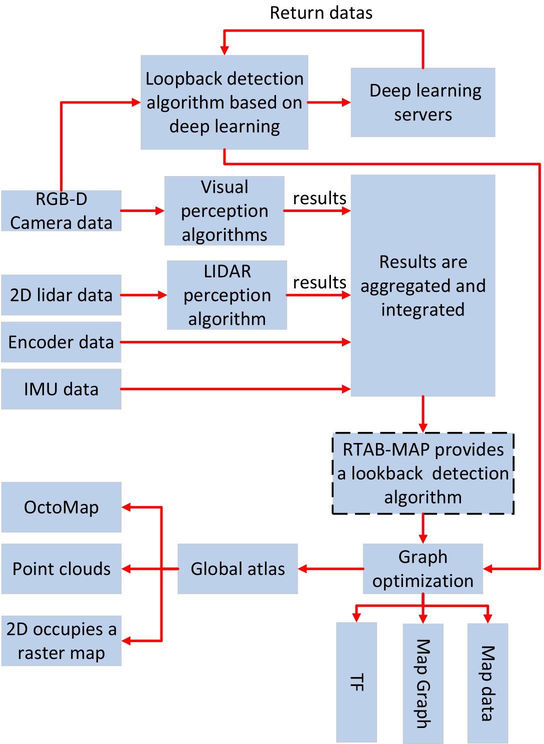

Return datas
Loopback detection algorithm based on deep learning
Deep learning servers
RGB-D Camera data
Visual perception algorithms
results
2D lidar data
LIDAR perception algorithm
results
Results are aggregated and integrated
Encoder data
IMU data
RTAB-MAP provides a lookback detection algorithm
OctoMap
Point clouds
2D occupies a raster map
Global atlas
Graph optimization
TF
Map Graph
Map data

FIGURE 7. Multi-sensor SLAM framework with loop closure detection algorithm.

Second, a comparative experiment of SLAM of the three multi-sensor fusion frameworks studied above is carried out on the improved Gmapping algorithm and analyzed the experimental results. Finally, a comparison experiment of SLAM mapping between the original RTAB-MAP algorithm and the deep learning-based RTAB-MAP loop closure detection algorithm is carried out, and the experimental results are analyzed in detail.

A. EXPERIMENTAL PLATFORM AND EXPERIMENTAL ENVIRONMENT

1) EXPERIMENTAL PLATFORM

FIG.8 below shows the Mecanum kinematics principle mobile robot, equipped with NVIDIA Jetson Nano motherboard, using the STM32F405 control board as the motion master to control the motion of the robot. In terms of sensors, it is equipped with Silan RPLIDAR A1 lidar, which has a measurement radius of 12M and a scanning frequency of 16HZ. IMU accelerometer gyroscope sensor is used to measure the three-axis attitude angle (or angular rate) and acceleration of the object. 520 encoder geared motor, accumulating mileage from the moment the robot starts to move. LeTV LeTMC-520 RGB-D depth camera with RGB pixels is 1080P, a depth resolution of 640 × 480, and a video frame rate of 30FPS.

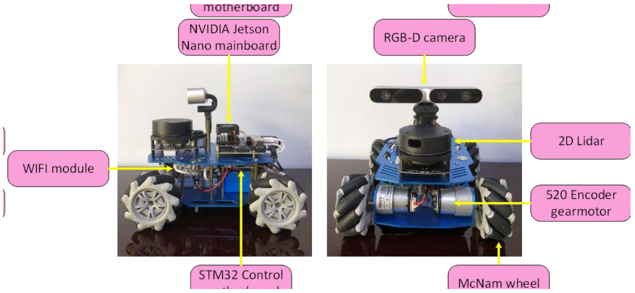


FIGURE 8. Mecanum Kinematics principle of mobile robot.

All software of the Mecanum robot runs in the Ubuntu system, and the version used in this paper is Ubuntu18.04. The robot's operating system is ROS Melodic, which can run in the Ubuntu18.04 environment. Moreover, many third-party libraries are used in the experiments, such as Pangolin, Eigen, nanoflann, PCL, OpenCV, Octomap, G2O, and Sophus, Etc. After the connection between the PC (Personal Computer) side and the robot side is established through the LAN, the software rviz can be used on the PC to inspect the robot's SLAM process visually, and the software Gazebo can be used to simulate the robot's running experiments.

The experimental platform is a server for the deep learning laboratory. The hardware configuration includes two Intel Xeon processor E5 series CPUs, four RTX 2080 Ti graphics cards, one 1TB SSD solid state drive, and two DDR4 32G memories. The software environment is a 64-bit Ubuntu18.04 operating system, which is configured with CUDA10.0 and cuDNN7.5 provided by NVIDIA. The PyTorch deep learning framework is used as the basic framework of the experiment, and the development tool is Python3.7.

2) EXPERIMENTAL ENVIRONMENT

In order to better compare the SLAM effects of different multi-sensor fusion frameworks, the experimental environment is selected in a closed room, which can reduce the impact of drastic environmental changes on the experiment. FIG.9 shows the central part of the experimental environment.

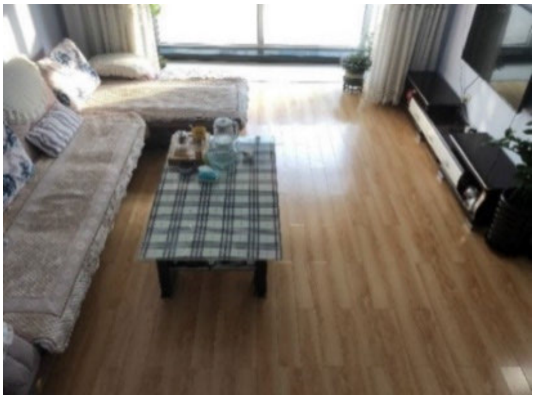


FIGURE 9. The central part of the experimental environment.

B. IMPROVED GMAPPING ALGORITHM EXPERIMENT

When using the original and improved algorithms in this paper to conduct a comparative experiment, the displacement distance is the same, which can ensure that the experimental results are more strongly compared. When using the original Gmapping algorithm, the state of the particles during the robot's forward process is shown in FIG.10(a)(b).

When using the improved Gmapping algorithm, the state of the particles in the forward process of the robot is shown in FIG.11(a)(b).

The time used to run the original Gmapping algorithm and the improved Gmapping algorithm in this paper to make the particles reach the same convergence state is shown in Table 2.

TABLE 3. Time for particles to converge to the same state for the two algorithms.

algorithm	Particle convergence time /s
Gmapping original algorithm	5.42
Gmapping improved algorithm	3.26

By observing FIG.10(a)(b), it can be concluded that when the Gmapping algorithm is used, the state of the particles has roughly converged after the robot moves about two map units. When again observed in Fig. 11(a)(b) using the improved Gmapping algorithm in this paper, the particles converge after the robot moves the same distance as the original algorithm experimental process. Then, according to Table 3, it can be concluded that the particle convergence speed of the improved Gmapping algorithm in this paper is increased by 39.85%. To sum up, the experiment proves that the particle convergence speed of the improved Gmapping algorithm in this paper is faster than the original algorithm, and a better improvement result is obtained.

Yan et al. [34] made two improvements based on the Gamping algorithm and conducted related experiments. The first improvement is to combine the Gmapping algorithm with AF (Firefly Algorithm) to increase its speed by 7.32%; the second improvement is to combine the Gmapping algorithm Combining with AF (firefly algorithm) and AS (adaptive sampling), the speed is increased by about 14%, and the improved method proposed in this paper increases the speed by 39.85%, which fully proves the effectiveness of the improved method proposed in this paper.

C. MULTI-SENSOR FUSION SLAM EXPERIMENT

The primary purpose of the experiments in this subsection is to compare the three multi-sensing SLAM frameworks studied in the previous section. In this part of the experiment, the improved Gmapping algorithm in this paper is used.

The primary purpose of the experiments in this subsection is to compare the three multi-sensing SLAM frameworks studied in the previous section. In this part of the experiment, the improved Gmapping algorithm in this paper is used.

- The first part is the SLAM experiment of laser and encoder fusion. The experimental results are shown in FIG.12(a)(b)(c).
- The second part is the SLAM experiment of the fusion of laser, encoder and IMU. The experimental results are shown in FIG.13(a)(b)(c).
- The third part is the SLAM experiment of the fusion of vision, laser, encoder and IMU. First, FIG.14(a)(b)(c) is the colour image, depth image and depth-point cloud data map captured by the camera. The images captured by these cameras can provide much information for the system to complete the SLAM function. The experimental results are shown in FIG.15(a)(b)(c), which shows three perspectives of the mapping results.

First, comparing the experimental results in FIG.12 and FIG.13, it can be concluded from observation that when the IMU data is not fused, the edge contours of the 2D grid maps constructed by the three algorithms are blurred and noisy. The 2D raster map constructed after the system integrates the IMU data has more apparent edges and reduced noise. Then, comparing the experimental results of FIG.13 and FIG.14, after the system integrates the data provided by the RGB-D camera, it can reconstruct the surrounding scene in 3D based on building a 2D grid map. The map constructed in this way can contain more environmental information.

It can be seen from the horizontal comparison experimental results that the edge contour of the map constructed

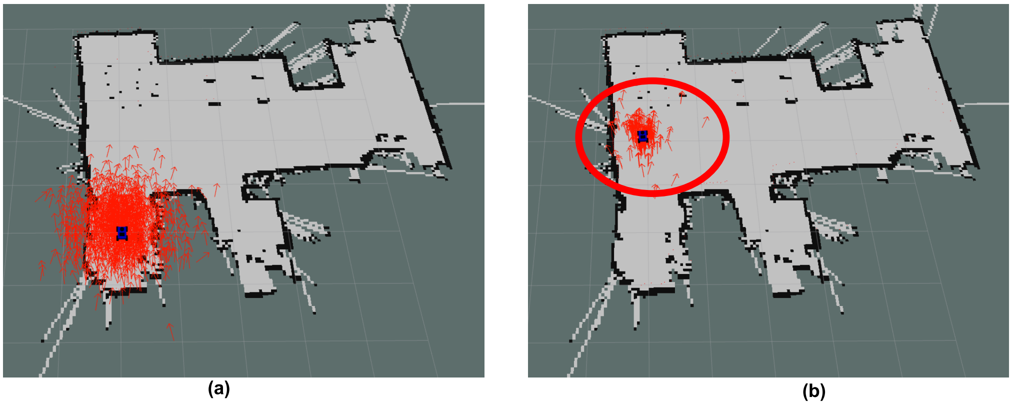


FIGURE 10. Particle state of the original algorithm. (a) The particle state at the initial position of the robot in the original algorithm; (b) The particle state after the robot moves forward in the original algorithm.

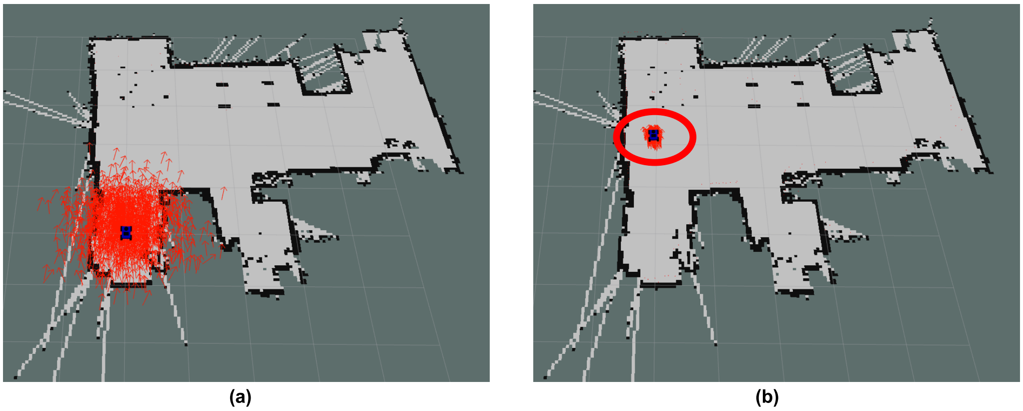


FIGURE 11. Particle state of the improved algorithm. (a) The particle state at the initial position of the robot in the improved algorithm; (b) The particle state after the robot moves forward in the improved algorithm.

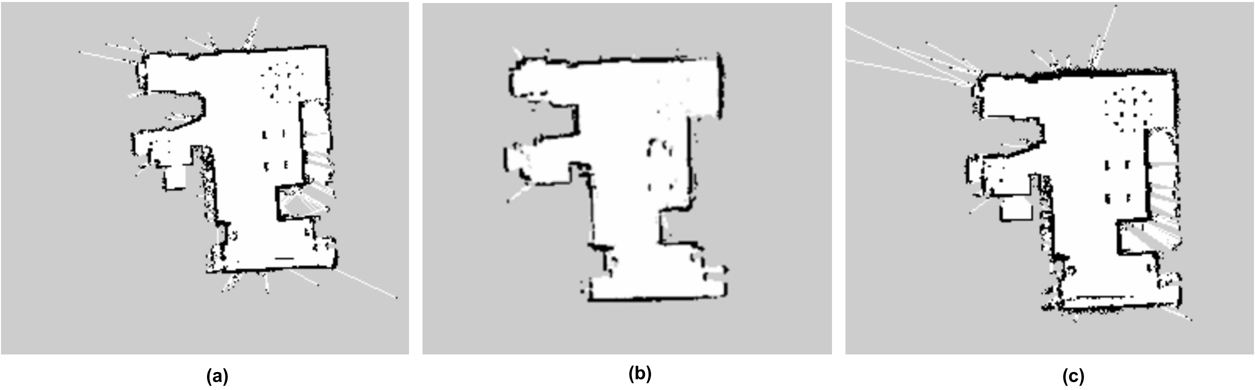


FIGURE 12. Experimental results of laser and encoder fusion SLAM. (a)Gmapping experimental result; (b)Cartographer experimental result; (c) HectorSLAM experimental result.

by the Cartographer algorithm based on graph optimization is relatively blurred. The construction of the HectorSLAM algorithm will contain more noise at the edges of the map. The Gmapping algorithm, with more explicit map boundaries and less noise, is better than the HectorSLAM algorithm and the Cartographer algorithm in the performance of indoor mapping.

In addition, from the experimental results, it can be concluded that the information obtained by the 2D lidar, encoder and IMU can only build a 2D grid map and cannot build a 3D environment map. Moreover, because the laser-based multi-sensor SLAM system is challenging to complete the loop closure detection, it will lead to poor global consistency of the constructed map. According to the experimental analysis,

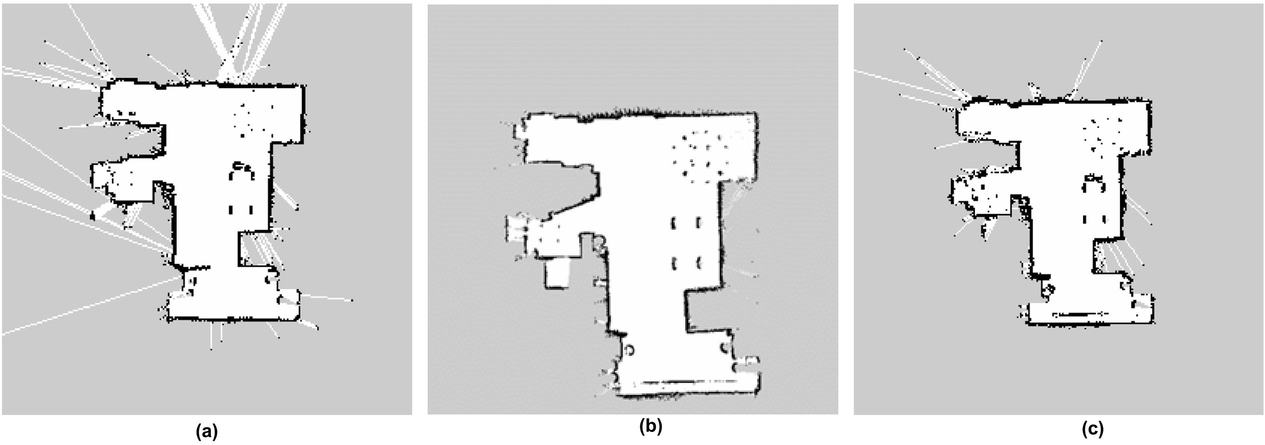


FIGURE 13. Experimental results of laser and encoder fusion SLAM. (a)Gmapping experimental result; (b)Cartographer experimental result; (c) HectorSLAM experimental result.

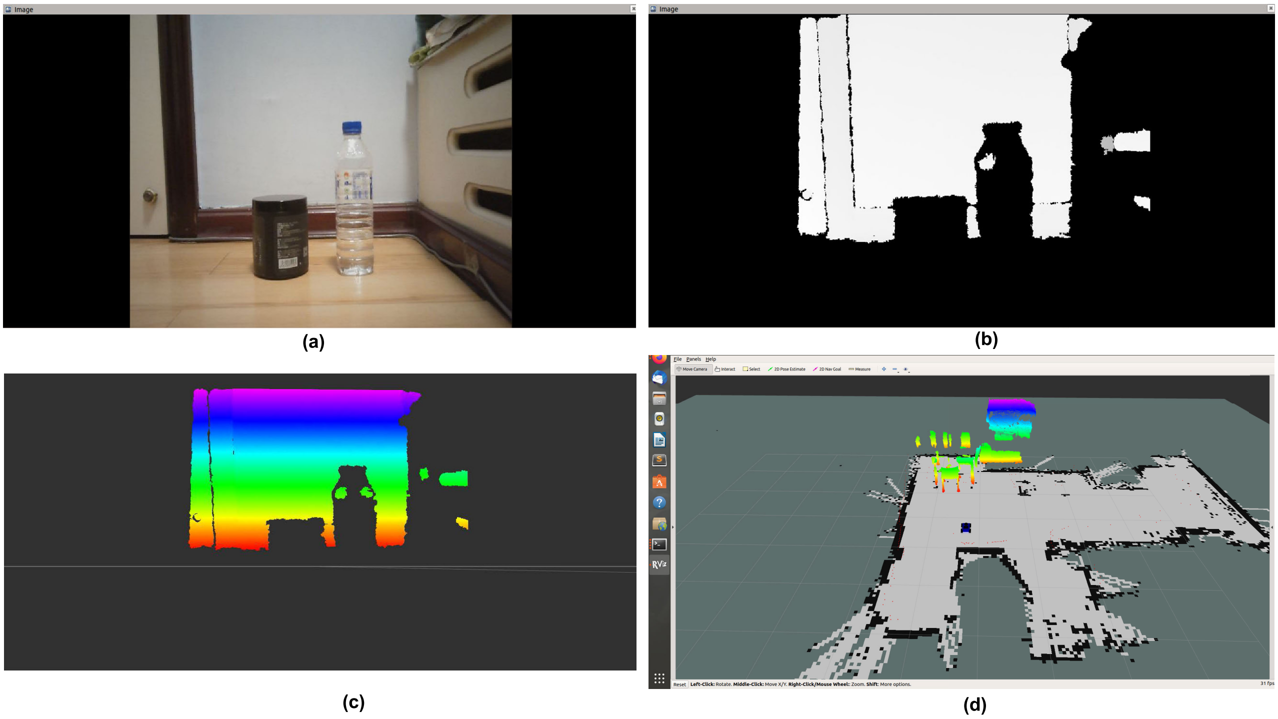


FIGURE 14. Image captured by the camera. (a) Color image; (b) Depth image; (c)Depth point cloud data graph; (d) Depth-point cloud data graph at runtime.

it can be concluded that the overall performance of the SLAM system integrated with vision, laser, encoder and IMU is better.

D. COMPARATIVE EXPERIMENT OF LOOP CLOSURE DETECTION ALGORITHM BASED ON DEEP LEARNING

The loop closure detection experiment in this paper uses the dataset CityCentre from Oxford University, a total of 2474 outdoor scene images, the size is 640*480, and the format is.jpg. The indicators for evaluating the loop closure detection algorithm in this paper mainly include the Precision-Recall (P-R) curve and the Average Precision, (AP).

In this section, AlexNet is selected as the network for extracting low-dimensional features, and MobileNetV3-Large is used as the network for extracting high-dimensional features. Both networks are pre-trained on ImageNet. Use the cosine similarity to judge the similarity of the output features to perform loop closure detection. The experimental results show that when $\alpha = 0.8$, $\beta = 0.6$, $\gamma = 0.95$, the accuracy of loop closure detection is the highest, and the detection efficiency is improved with the reduction of calculation amount.

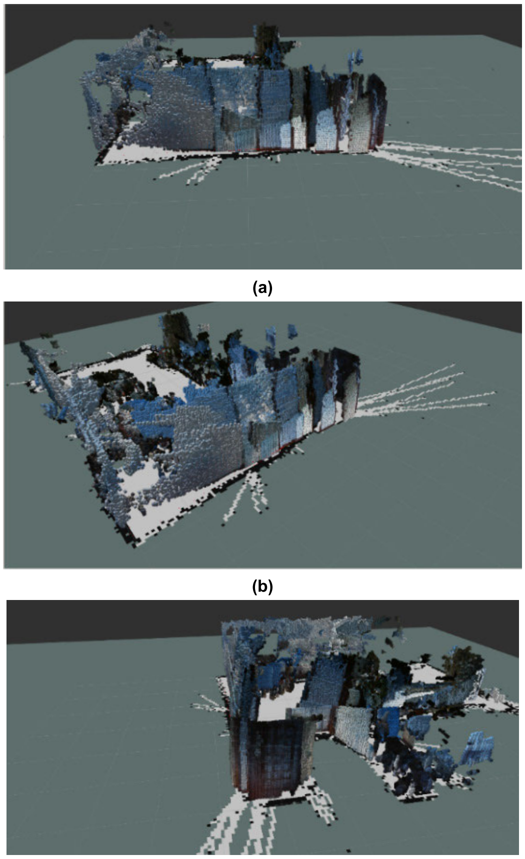


FIGURE 15. Experimental results of vision, laser, encoder and IMU fusion SLAM. (a) Due south; (b) Diagonal direction; (c) Due east.

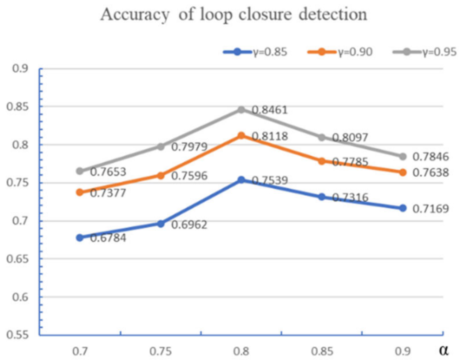


FIGURE 16. Line diagram of loop closure detection experiment.

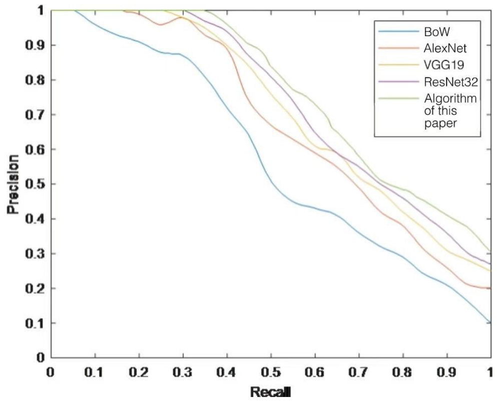


FIGURE 17. Experimental results using the CityCentre dataset.

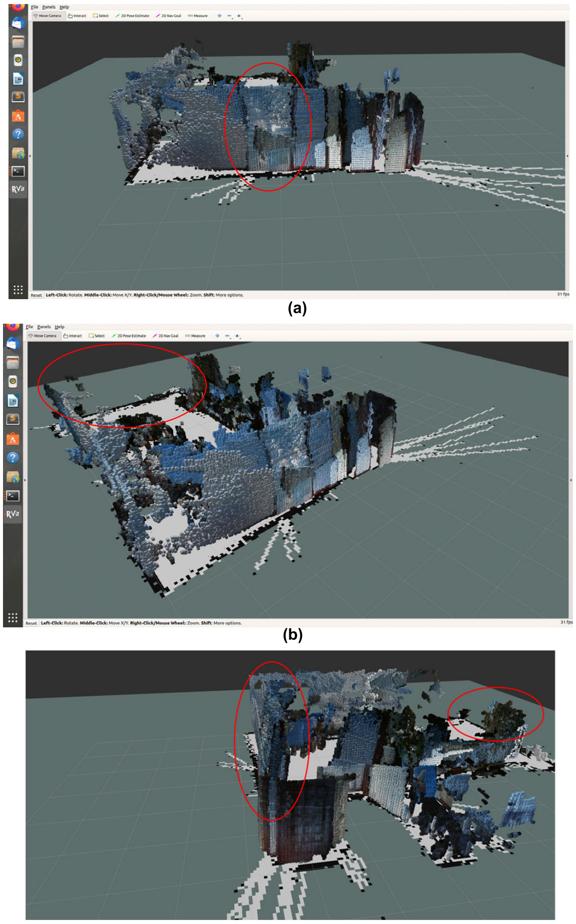
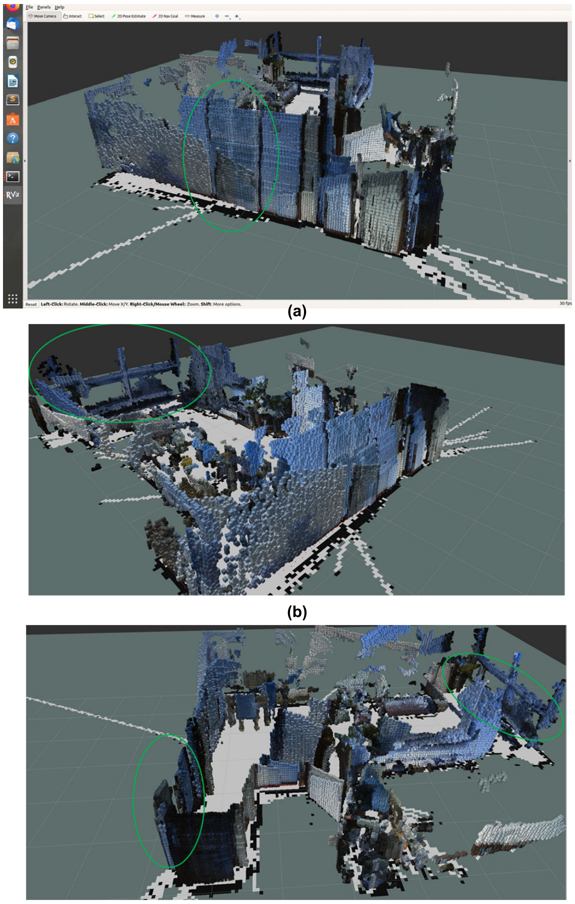


FIGURE 18. RTAB-MAP mapping result. (a) Due south; (b) Diagonal direction; (c) Due east.

In the comparison experiment, the traditional loop closure detection algorithm is the SIFT-based Bag of Words (BoW) algorithm, and the deep learning-based loop closure detection algorithm includes AlexNet, VGG19, ResNet32 and the deep learning-based loop closure detection algorithm proposed in this paper. By observing the P-R curve shown in FIG.17, it can be seen that when the recall rate is lower than 0.4, the accuracy of the algorithm proposed in this paper is 1. In comparison, the loop closure detection algorithm proposed in this paper has the best effect. From the average accuracy rate shown in Table 4, the accuracy rate of the algorithm proposed in this paper is 31.26% higher than that of the traditional algorithm BoW; compared with the deep learning-based AlexNet algorithm, VGG19 algorithm, ResNet32 algorithm, the accuracy rates Increased by 14.21%, 3.05%, and 1.56%. In contrast, the overall performance of the loop closure detection algorithm proposed in this paper is the best.



(a)

(b)

(c)

FIGURE 19. RTAB-MAP mapping results after replacing the loopback detection algorithm. (a) Due south; (b) Diagonal direction; (c) Due east.

E. TWO TYPES OF LOOP CLOSURE DETECTION ALGORITHM's COMPARISON EXPERIMENTS

Run RTAB-MAP original algorithm on the experimental platform for the SLAM mapping experiment, and the results are shown in FIG.18.

TABLE 4. Experimental algorithm loop closure detection accuracy.

algorithm	Accuracy on CityCentre
BoW	0.643
AlexNet	0.739
VGG19	0.819
ResNet32	0.831
The algorithm proposed in this paper	0.844

In the experimental platform, the loop closure detection algorithm introduced above is used to replace the traditional loop closure detection algorithm provided by RTAB-MAP to conduct the SLAM mapping experiment. The results are shown in FIG.19.

RTAB-MAP can build a 3D map of the surrounding environment. Comparing the experimental results, at the red circle mark in FIG.18 and the green circle mark in FIG.19, after using the loop closure detection algorithm proposed in this paper to replace the loop closure detection algorithm in RTAB-MAP, it can be clearly observed that the constructed Environment maps are more regular. FIG.19 does not produce the distortion at the red circle in FIG.18, which means that a globally consistent map can be better constructed.

Moreover, the environment map constructed by the latter can reproduce added environmental information.

V. CONCLUSION

In this paper, we propose a multi-sensor fusion SLAM algorithm framework based on improved Gmapping. Through theoretical research and experimental verification, it is proved that the framework applied to mobile robots can make it work indoors with high robustness and precision. We see that the SLAM system integrated by vision, laser, encoder and IMU is generally more stable and accurate. Secondly, the improved Gmapping algorithm is significantly better than the Cartographer algorithm and HectorSLAM algorithm in terms of indoor mapping performance, and the particle convergence speed is 39.85% higher than the original Gmapping algorithm, which significantly reduces the memory and calculation amount occupied by the algorithm. Finally, the loop closure detection algorithm proposed in this paper is superior to the traditional algorithm BoW, AlexNet algorithm, VGG19 algorithm, and ResNet32 algorithm in terms of accuracy, and is superior to the original RTAB-MAP algorithm in terms of the construction effect of the environment map.

To address the shortcomings of this paper in future research, first, we will test our system in more complex scenarios, and improve the test analysis to compare more different algorithms to achieve better results. Second, a multi-geometry-based dynamic object detection method will be added to cooperate with image processing algorithms to further improve the adaptability and robustness of the system in dynamic environments. Third, on this basis, explore the multi-sensor fusion SLAM and multi-frame fusion Gmapping algorithm based on the tightly coupled method of extended Kalman filtering to further improve the performance of SLAM.

REFERENCES

[1] M. Yang, "Overview on issues and solutions of SLAM for mobile robot," *Comput. Syst. Appl.*, vol. 27, no. 7, pp. 1–10, Jul. 2018.

[2] M. Labbé and F. Michaud, "RTAB-map as an open-source LiDAR and visual simultaneous localization and mapping library for large-scale and long-term online operation," *J. Field Robot.*, vol. 36, no. 2, pp. 416–446, 2019.

[3] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2D LiDAR SLAM," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 1271–1278.

[4] J. Zhang and S. Singh, "Visual-LiDAR odometry and mapping: Low-drift, robust, and fast," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 2174–2181.

[5] X. Wang, "Mobile robot for SLAM research based on LiDAR and binocular vision fusion," *Chin. J. Sensors Actuators*, vol. 31, no. 3, pp. 394–399, Mar. 2018.

[6] K. Konolige, G. Grisetti, R. Kümmerle, W. Burgard, B. Limketkai, and R. Vincent, "Efficient sparse pose adjustment for 2D mapping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 22–29.

[7] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE Trans. Robot.*, vol. 31, no. 5, pp. 1147–1163, Oct. 2015.

[8] S. Kumar and R. M. Hegde, "Multi-sensor data fusion methods for indoor localization under collinear ambiguity," *Pervasive Mobile Comput.*, vol. 30, pp. 18–31, Aug. 2016.

[9] X. Ding, Y. Wang, D. Li, L. Tang, H. Yin, and R. Xiong, "Laser map aided visual inertial localization in changing environment," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 4794–4801.

[10] Y. Sun, M. Liu, and M. Q.-H. Meng, "Active perception for foreground segmentation: An RGB-D data-based background modeling method," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 4, pp. 1596–1609, Oct. 2019.

[11] L. Zhang, L. Wei, P. Shen, W. Wei, G. Zhu, and J. Song, "Semantic SLAM based on object detection and improved octomap," *IEEE Access*, vol. 6, pp. 75545–75559, 2018.

[12] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An open-source slam system for monocular, stereo, and RGB-D cameras," *IEEE Trans. Robot.*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017.

[13] B. Bescos, J. M. Fácil, J. Civera, and J. L. Neira, "DynaSLAM: Tracking, mapping, and inpainting in dynamic scenes," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 4076–4083, Oct. 2018.

[14] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.

[15] G. Wan, X. Yang, R. Cai, H. Li, Y. Zhou, H. Wang, and S. Song, "Robust and precise vehicle localization based on multi-sensor fusion in diverse city scenes," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 4670–4677.

[16] H. Xue, H. Fu, and B. Dai, "IMU-aided high-frequency LiDAR odometry for autonomous driving," *Appl. Sci.*, vol. 9, no. 7, p. 1506, Apr. 2019.

[17] D. Wisth, "Unified multi-modal landmark tracking for tightly coupled LiDAR-visual-inertial odometry," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 1255–1262, Apr. 2021.

[18] L. Meng, C. Ye, and W. Lin, "A tightly coupled monocular visual LiDAR odometry with loop closure," *Intell. Service Robot.*, vol. 15, no. 1, pp. 129–141, Mar. 2022.

[19] R. Li, S. Wang, Z. Long, and D. Gu, "UnDeepVO: Monocular visual odometry through unsupervised deep learning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Brisbane, QLD, Australia, May 2018, pp. 7286–7291.

[20] C. Zonghai, "Monocular visual odometer based on recurrent convolutional neural network," *Robotics*, vol. 41, no. 2, pp. 147–155, 2019.

[21] Y. Yu, "A loop closure detection method for visual SLAM based on deep learning," *Comput. Eng. Des.*, vol. 41, no. 2, pp. 529–536, 2020.

[22] Y. Zhou, Y. Wang, F. Poiesi, Q. Qin, and Y. Wan, "Loop closure detection using local 3D deep descriptors," *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 6335–6342, Jul. 2022.

[23] D. Cattaneo, M. Vaghi, and A. Valada, "LCDNet: Deep loop closure detection and point cloud registration for LiDAR SLAM," *IEEE Trans. Robot.*, vol. 38, no. 4, pp. 2074–2093, Aug. 2022.

[24] S. Das, "Simultaneous localization and mapping (SLAM) using RTAB-MAP," 2018, *arXiv:1809.02989*.

[25] J. Civera and S. H. Lee, "RGB-D odometry and SLAM," in *RGB-D Image Analysis and Processing*. Cham, Switzerland: Springer, Oct. 2019, pp. 117–144.

[26] M. Hsiao, E. Westman, G. Zhang, and M. Kaess, "Keyframe-based dense planar SLAM," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Singapore, May 2017, pp. 5110–5117.

[27] F. Steinbrucker, C. Kerl, D. Cremers, and J. Sturm, "Large-scale multiresolution surface reconstruction from RGB-D sequences," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sydney, NSW, Australia, Dec. 2013, pp. 3264–3271.

[28] Y. Sun, M. Liu, and M. Q.-H. Meng, "Improving RGB-D SLAM in dynamic environments: A motion removal approach," *Robot. Auton. Syst.*, vol. 89, pp. 110–122, Mar. 2017.

[29] M. Klingensmith, S. S. Sirinivasa, and M. Kaess, "Articulatedrobotmotion for simultaneous localization and mapping (ARM-SLAM)," *IEEE Robot. Autom. Lett.*, vol. 1, no. 2, pp. 1156–1163, Jul. 2016.

[30] R. Scona, M. Jaimez, Y. R. Petillot, M. Fallon, and D. Cremers, "StaticFusion: Background reconstruction for dense RGB-D SLAM in dynamic environments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Brisbane, QLD, Australia, May 2018, pp. 1–9.

[31] J. Yang, D. Guo, K. Li, Z. Wu, and Y.-K. Lai, "Global 3D non-rigid registration of deformable objects using a single RGB-D camera," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 4746–4761, Oct. 2019.

[32] J.-H. Kim, C. Cadena, and I. Reid, "Direct semi-dense SLAM for rolling shutter cameras," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 1308–1315.

[33] H. Li, C. Tian, L. Wang, and H. Lv, "A loop closure detection method based on semantic segmentation and convolutional neural network," in *Proc. Int. Conf. Artif. Intell. Electromech. Autom. (AIEA)*, May 2021, pp. 269–272.

[34] Y. Han, W. Wei, C. Jinhua, D. Didi, and W. Rujia, "Research on SLAM Gmapping based on AF and AS algorithm optimization," *J. Jiangsu Inst. Technol.*, vol. 28, no. 2, pp. 93–101, 2022.

CHENGJUN TIAN received the Doctor of Engineering degree from the Changchun University of Science and Technology, in 2011. He is currently an Associate Professor with the Changchun University of Science and Technology. His current research interests include pattern recognition and intelligent systems. He is also a member of the Education and Training Committee of the China Simulation Society, the Director of the Jilin Province Automation Society, and the Director of the Jilin Province Robotics Society.

HAOBO LIU received the bachelor's degree from the Jilin Institute of Chemical Technology, Jilin, China, in 2021. He is currently pursuing the master's degree with the Changchun University of Science and Technology, Jilin. His current research interests include SLAM and computer vision.

ZHE LIU received the bachelor's degree from the Jiangxi University of Science and Technology, Jiangxi, China, in 2021. She is currently pursuing the master's degree with the Changchun University of Science and Technology, Jilin, China. Her current research interests include SLAM and computer vision.

HONGYANG LI received the bachelor's degree from the Tianjin University of Science and Technology, Tianjin, China, in 2016, and the master's degree from the Changchun University of Science and Technology, Jilin, China, in 2021. He is currently engaged in SLAM and NLP related work.

YUYU WANG received the bachelor's degree in engineering from the School of Electronic Information Engineering, Changchun University of Science and Technology, Changchun, China, in 2020. He is currently pursuing the master's degree in control science and engineering. His current research interests include computer vision and motion planning of robotic arms.