Research Article

Design of Multisensor Mobile Robot Vision Based on the RBPF-SLAM Algorithm

Yiyi Cai and Tuanfa Qin

1School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, Guangdong, China
2The Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning 530004, Guangxi, China

Correspondence should be addressed to Tuanfa Qin; tfqin@gxu.edu.cn

Received 21 June 2022; Accepted 16 August 2022; Published 31 August 2022

Academic Editor: Wenlong Hang

Copyright © 2022 Yiyi Cai and Tuanfa Qin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The existing algorithms are mainly based on the laser or visual inertial odometer, which has problems such as large particle recommendation distribution error, particle consumption, and long running time of the algorithm. An SLAM scheme based on multisensor fusion is proposed, which integrates multisource sensor data including lidar, camera, and IMU and constructs a set of real-time mapping and positioning system with better robustness and higher accuracy. The experimental results show that the proposed SLAM method uses less running time, and the accuracy of map construction is better.

1. Introduction

SLAM (simultaneous localization and mapping) is the key to realize autonomous localization and navigation of the robot in an unknown environment by acquiring environment information through the multisensor carried by the robot itself, estimating its own position and posing to complete map construction. At present, single robot SLAM technology has a certain degree of robustness, and it is widely used in automatic driving vehicles for high-precision map construction and location in unknown environment, which provides accurate trajectory information for automatic drive such as navigation planning modules [1,2]. The camera is low cost and easy to install, which is often used in UAV or low-cost automatic driving vehicle. However, the monocular camera has small FOV [3] angle and is sensitive to light conditions, so it is difficult to obtain the accurate depth information of the motion scene in small scale. In addition, lidar is not affected by ambient illumination, and can accurately and dynamically measure the environmental depth information without considering the error caused by extra calculation. It has a 360° field of view in the horizontal direction and can detect the environmental information within 200 m as far as possible. In recent years, with the gradual reduction of lidar cost, more and more automatic driving vehicle platforms choose lidar as one of the main sensing components of the system. However, it produces distortion in the process of high-speed motion, which has poor adaptability to the scenes with repeated characteristics such as tunnel and highway. Lou et al. [4] adopted the feature extraction scheme of lidar point cloud based on curvature to improve the accuracy and robustness of the system, but it is easy to diverge and fail in high-speed and feature-deficient scenes.

The principle of the particle filter which is a nonparametric filter is based on Bayesian reasoning and gives importance to sampling, which can deal with nonlinear and multipeak distribution problems [5]. Thrun Montemerlo [6,7] first proposed a SLAM method based on the particle filter, which assigns weights to particles in the space state, which makes positioning more accurate. However, as the number of particles increases, the complexity of map construction also increases. Li et al. [8] proposed a RBPF-SLAM method based on adaptive resampling, which improved the
2 Mathematical Problems in Engineering

operation efficiency of SLAM. Luo et al. [9] combined the odometer and radar data to optimize the proposed distribution function, thereby effectively reducing the uncertainty of robot pose in the prediction stage and reducing the number of particles required for SLAM.

SLAM based on the single sensor is unstable when the indoor environment is poor, and the robot moves too fast or turns too fast. For example, the scanning observation range of the lidar sensor is limited, and it is easily affected by the complex geometric structure in the environment. The camera has certain requirements for lighting conditions around the robot. The encoder motor will produce accumulated errors after long working time. Based on the above methods, a SLAM system based on tightly coupled vision lidar-IMU and a state estimation scheme based on nonlinear sliding window optimization is proposed by integrating 2D lidar, IMU, wheeled odometer, and other sensors. First, IMU is used to correct the odometer data and establish the robot motion model. Then, the lidar observation data are introduced to modify the robot motion data, and the proposed distribution is jointly optimized to improve the accuracy of the model. At the same time, aiming at the particle dissipation problem of the RBPF-SLAM algorithm, the particle resampling strategy is improved.

2. Research Status

Lidar is not affected by ambient light and can accurately and dynamically measure ambient depth information without considering the error introduced by additional calculations. It has a 360° field of view in the horizontal direction and can detect environmental information up to 200 m. However, the lidar will produce distortion in the process of high-speed movement, and it is not suitable for the scenes with repeated characteristics. The fusion scheme usually adopts the scheme of inertial measurement unit (IMU) or lidar-IMU. There are usually two kinds of tight coupling schemes based on inertial sensors [10], one is based on EKF (extended Kalman filter) or ESKF (error state Kalman filter), such as MSCKF and FAST-LIO, and the other is based on optimization, and the representative algorithms are VINS-MONO and LIO-SAM. In the above scheme, IMU can be used to correct the point cloud motion distortion of lidar and provide a priori motion estimation for interframe matching. The scheme of integrating with IMU enables the system to deal with the problem of environmental information missing in a short time so as to improve the robustness and accuracy of the system under long-term operation.

Based on the inertial fusion scheme, the multimode sensor fusion mapping and positioning scheme using camera-lidar-IMU is gradually developed, which has higher accuracy and robustness. Graeter et al. [11] used lidar point cloud projection to assist in recovering the depth information of visual feature points, which solves the problem of low accuracy and stability of depth information estimation. Shao et al. [12] used point cloud matching to further optimize the calculation results after detecting loopback information and obtaining matching results so as to solve the problem of low accuracy of visual matching scheme based on PnP (Perspective-n-Point) calculation. LVI-SAM [13] designed two subsystems of lidar inertial odometer and visual inertial odometer using the fusion scheme optimized by factor graph, where after the integration with IMU is completed in the subsystem, the optimization results are output. However, there are still some problems in the above scheme. The main conclusions are as follows:

1. The original point cloud is projected to the pixel plane, and then the filtering and plane fitting will lead to the same point cloud blocks being extracted and fitted for many times in the area with dense feature points, which reduces the operation efficiency of the system.

2. Multiple subsystems independently optimize the repeated use of IMU information, resulting in IMU bias absorbing different system residuals in different subsystems, which destroys the consistency of data and affects the stability of the system.

3. Finally, due to the small range of visual perception, the single vision loop detection scheme cannot detect the loop information correctly in the intersection or wide lane driving environment.

3. SLAM System Based on Multisensors

3.1. Overall Structure. The overall architecture of the proposed multisensor fusion SLAM system is shown in Figure 1. It is mainly composed of the following parts:

1. The front-end data preprocessing part includes feature point extraction, optical flow tracking based on sparse pyramid, feature point depth estimation, lidar point cloud feature extraction, odometer calculation of frame local map, and optimization of visual feature point depth information by using point cloud features.

2. In the back-end state estimation part, various kinds of observation data provided by the front-end are used to construct the corresponding constraints, such as the visual reprojection error constraint constructed by visual feature points, the preintegration constraint constructed by IMU, and the relative attitude and prior position constraint between key frames provided by the lidar odometer.

3. Loop detection part is based on vision-lidar coupling.

3.2. Optimized RBPF-SLAM Algorithm. The traditional RBPF-SLAM algorithm provides the robot motion model and observation model through the odometer and lidar sensor information, respectively, and takes the motion model as the recommended distribution of the particle filter separately, so there is a large cumulative error in the estimation of robot pose, which makes the posteriori distribution of ground map of particle computing deviate from the real environment greatly, which leads to the particle memory explosion problem. In this paper, a multisensor SLAM method is proposed.
3.2.1. Data Fusion. In the working environment, the robot SLAM algorithm calculates the velocity of the robot through an odometer to obtain the real-time pose of the robot. But the odometer itself will have a certain cumulative error. The IMU sensor is composed of an accelerometer, a magnetometer, and a gyroscope, which can provide stable robot attitude information. In this paper, the odometer error is corrected by using the characteristics of IMU sensor with high precision and fast response in a short time.

For a mobile robot, the collected wheel odometer data and IMU data are modeled, and the space state quantity expression is defined as follows:

\[ x_t = [X_t, Y_t, \theta_t, v_t, \omega_t]^T, \]

where \( X_t, Y_t \) are displacement quantities of mobile robot in \( X \) and \( Y \) directions, respectively. \( \theta \) is the attitude angle of the robot. The pose at time \( t + 1 \) can be expressed as

\[
x_{t+1} = [X_t + v_t (\cos \theta_t) \Delta t, Y_t + v_t (\sin \theta_t) \Delta t, \theta_t + \omega_t \Delta t, v_t, \omega_t].
\]

Therefore, the motion equation of the system is as follows:

\[
x_{t+1} = \varphi_{t+1} x_t + \Gamma W_t = f(x_t) + w_t.
\]

Set the threshold \( \theta_0 \) to calculate the difference between the attitude angle measured by the current IMU sensor and odometer. If the difference is greater than the given threshold, the attitude angle measured by IMU is used to estimate the robot’s attitude; otherwise, the attitude angle is calculated and the robot’s attitude is estimated by weighted average. The fusion process of IMU and the odometer is shown in Figure 2.

In the traditional RBPF-SLAM algorithm, the motion model of the robot SLAM is constructed by the information of odometer. In view of the fact that the odometer is easily disturbed by external information, the precise lidar data are fused into the proposed distribution function. By combining the motion model obtained from odometer and IMU data with the observation model based on lidar, a new proposed distribution function is obtained, and its function is as follows:

\[
p(x_t | m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1}) = \frac{p(z_t | m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1})}{p(z_t | m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1})}.
\]

Formula (3) is substituted into the particle weight calculation formula to obtain

\[
w_t^{(i)} = w_{t-1}^{(i)} \eta p(z_t | m_{t-1}^{(i)}, x_{t-1}^{(i)}) p(x_t | x_{t-1}^{(i)}, u_{t-1})
\frac{p(x_t | m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1})}{p(x_t | m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1})}
\]

\[
= w_{t-1}^{(i)} \cdot p(z_t | m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1}).
\]
By analyzing the Formula (5), it is found that the variance of the important weight of laser information matching is very small. If the precise lidar scanning matching information is added to represent the recommended distribution, the robot model can be modified to limit the sampling range. In addition, the proposed distribution with laser information is closer to the real target distribution, which effectively reduces the number of particles required.

3.2.2. Improvement of the Particle Sampling Strategy. After the improved proposed distribution function is obtained, in order to alleviate the problem of particle dissipation, the particle resampling strategy is improved, where the corresponding distribution function value of each particle is calculated by constructing a new distribution function, and the value of the distribution function is kept in the [1/2, 1], and the first three particles are selected as the new particle set. Finally, resampling is carried out in the specified interval. The specific steps are as follows.

1. Sort the particle weights in the descending order.
2. Construct distribution function \( H(w_k) \) to calculate the particle distribution function. The smaller the particle weight, the larger the corresponding distribution function value. If the distribution function value is between [1/2, 1], the particle weight is small and the deviation from the actual robot pose is large. Therefore, the particles whose distribution function value is less than 1/2 are reserved.
3. According to the distribution function value, the top three particles are selected, and then a particle is randomly selected, and its corresponding position in the new particle set is \( A \), and the corresponding position of the particle with distribution function value of 1/2 in the new particle set is \( B \). Finally, resampling is performed between \([A, B]\).

The implementation process of the improved RBPF-SLAM algorithm is as follows:

Step 1: Adopt the EKF method to fuse odometer and IMU data, so as to establish the robot motion model.
Step 2: Use the iterative nearest neighbor (ICP) algorithm to scan and match the lidar sensor according to the map \( m_{i-1}^{(i)} \), so as to obtain the observation model, and the robot motion model and observation model are fused to obtain a new proposed distribution function.
Step 3: Sample the particles obtained from Step 2.
Step 4: Calculate and update the importance weight of particles.
Step 5: Perform resampling according to the improved particle resampling strategy.
Step 6: Calculate \( m_{i}^{(i)} \) and update the map according to robot pose \( x_{i}^{(i)} \) and observation information \( z_{i} \).

3.3. Optimization of Loop Detection. When the SLAM system runs for a long time or in a large scene, it will inevitably lead to the problem of global consistency of the constructed map. Loop detection can effectively correct the trajectory offset of the system and make the map have better consistency. The widely used visual loop detection method is to calculate and store the descriptors of the visual feature points under each frame of the image, so it is convenient to store and retrieve the historical feature information by the bag of words model. However, due to the narrow view angle of the image, it is unable to correctly identify the scene of the loop, as shown in Figure 3, and the blue line is the vehicle trajectory and the gray line is the vehicle direction.

In the KITTI test scene, although the vehicle passes through the same intersection, due to the problem of view field, single vision cannot correctly detect the loop information in the intersection or opposite driving scene. Mechanical lidar has a 360° angle of view in the horizontal direction, but it is difficult to achieve efficient loop information detection with geometric storage. Therefore, a loop detection system coupled with lidar and vision is proposed. By using the advantages of both sensors, the system can achieve more accurate loopback information judgment with lower computational power consumption.

In this paper, Dbow2 [14] is used as the basis of visual loop detection. First of all, the key frames are matched in the history frames through the corresponding description operators on the optical flow tracing, and the most suitable frames are searched, and the historical frames which are outside the time or distance threshold with the current frame are eliminated. Then, the attitude data of the remaining historical frame data are obtained, through the PnP (Perspective-n-Point) algorithm, the position relationship between the current frame and the history frame is obtained.
Due to the low accuracy of the depth information of the visual feature points, the loop is a low-precision constraint information. Finally, to improve the accuracy of loop constraint, the attitude information of the current frame relative to the history frame of the visual loop is sent to the lidar odometer as the initial value.

The lidar odometer searches for the optimal matching according to the principle of minimum distance in the stored historical key frames and then matches according to the initial value provided by the visual loop, so as to further optimize the attitude constraints, and finally achieve high-precision loop detection results. In addition, to improve visual looping detection perspective defects, on the basis of visual loop detection, a low-cost lidar loop detection mechanism is added, where the lidar odometer will store the position and pose of historical key frames in real time. When the position information of the latest frame point cloud is close to the historical track, the loop detection mechanism will search the historical information and match it. Usually, in the case of no visual loopback detection, it may be invalid due to the large scene. However, in this paper, the coupling of two kinds of loopback methods is used to deal with loop detection in ordinary scenes through visual looping, so as to reduce the track error of the system under long-time operation. At this time, the accuracy of lidar through the spatial range loop detection can be greatly improved, which makes up for the loop failure caused by the visual angle problem.

4. Experiment and Analysis

4.1. Experimental Parameters. The open-source simulation dataset is used to test the performance of the algorithm, and the two SLAM methods are simulated. In the simulation experiment, the size of the simulation environment is 300 m, the maximum speed of the robot is 4 m/s, the maximum observation range of the lidar sensor is 20 m, the speed error \( \sigma v = 0.3 \) m/s, the distance error is 0.1 m, and the angle error is 1°. All simulations are completed in the same host, where the system type is 64 bit operating system, the processor is Intel i5 8300, the running memory is 8.00 GB, and the simulation platform is MATLAB 2016a.

The visual inertial fusion algorithm VINS-mono is selected to compare with the laser inertial fusion algorithm LIO-SAM to verify that the multimodal sensor fusion has better accuracy and environmental adaptability. In addition, the optimization strategy of depth information of visual feature points is compared and verified by the LIMO algorithm of visual lidar fusion. The scene in the campus environment is selected and the closed-loop data was recorded about 6 km in length with the system, where the campus environment adopted is more complex, including steep up and down slopes, more dense buildings, and various dynamic obstacles.

4.2. Results and Discussion

4.2.1. Error Analysis. The comparison results between different trajectories are shown in Figure 4. The proposed method has better trajectory consistency and end-to-end error compared with other methods.

Due to the large outdoor scene, it is difficult to estimate the depth information of feature points, and the scale information of the initial results of VINS-Mono cannot accurately estimate the scale information, which leads to the
The proposed method can constrain the visual scale information and carrier velocity to a certain range because of the use of lidar to assist the initialization of visual inertia, which ensures that the system does not crash due to outliers. Due to the small visual field angle, the loop detection strategy of LIMO and VINS-Mono cannot effectively record a large range of feature information, which leads to loop detection failure and obvious deviation between the track and the true value. In addition, it can be seen that visual feature points are used for state estimation in large scenes or high-speed environments, where it is easy to have errors in the estimation of depth information and scale information of feature points, which leads to trajectory divergence. While the proposed method can optimize the depth information of feature points, the fused laser odometer can further improve the accuracy of environmental scale information estimation.

4.2.2. Error Comparison. In addition, MATLAB is used to make the error curves of the robot in the simulation environment, as shown in Figure 5.

With the increase of time, the error curve of traditional algorithm fluctuates greatly. The error curve obtained by the SLAM algorithm based on multisensor fusion is relatively stable. It shows that the accuracy of the proposed algorithm is higher than that of the traditional algorithm. When the number of particles in the traditional algorithm is poor, the running time is longer. In this paper, the algorithm maintains the diversity of particles, effectively reduces the particle dissipation, improves its efficiency, which can obtain the accurate real-time attitude of the robot, and effectively improves the accuracy of the map.

4.2.3. Construction Simulation Map. The length, width, and height of the experimental platform were 45 cm, 38 cm, and 30 cm, respectively. The mobile platform adopts two-wheel
differential chassis, which can realize all-round movement, and the sensor part is composed of the lidar, inertial measurement sensor (IMU), and wheel odometer. The scanning radius of lidar is 12 meters, which can be scanned and matched according to the frequency. IMU provides stable real-time acceleration and angular velocity, and the odometer records the distance the robot has traveled and calculates the trajectory of the robot. The construction of simulation map is shown in Figure 6.

The SLAM algorithm proposed in this paper needs lesser particles than the traditional algorithm, which shortens the time of building map and greatly improves the efficiency of the algorithm. As can be seen from Figure 6, by only using the odometer as the distribution, the cumulative error of odometer increases with time, and the robot will lose and false the wall phenomenon after a certain period of time. The map constructed by the traditional method has the phenomenon of loss and false wall, but when the multisensor fusion SLAM algorithm is used, the map constructed is relatively stable and accurate.

5. Conclusion

In this paper, a SLAM scheme based on multisensor fusion is proposed, which integrates multisource sensor data including lidar, camera, and IMU to construct a set of real-time mapping and positioning system with better robustness and higher accuracy. The experimental results show that the proposed method can significantly improve the efficiency and accuracy of drawing construction. It improves the scheme of lidar point cloud feature extraction and plane fitting and improves the efficiency and accuracy of using point cloud to optimize the depth information of visual feature points. This research is expected to be widely used in autonomous vehicles to construct high-precision maps and output positioning results in unknown environments and to provide accurate self-vehicle attitude and trajectory information for other functions of autonomous vehicles, such as navigation planning and sensing modules.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Wang, L. Yang, and Y. Cai, “3d real time vehicle tracking based on lidar,” Automotive Engineering, vol. 43, no. 7, pp. 1013–1021, 2021.
[2] C. Cui, Multi Target Tracking Algorithm Based on Infrared Particle FilterShenyang University of technology, Shenyang, China, 2017.
[3] R. Mur-Artal and J. D. Tardos, “ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras,” IEEE Transactions on Robotics, vol. 33, no. 5, pp. 1255–1262, 2017.
[4] X. Lou, H. Wang, and Y. Cai, “Research on real-time road obstacle detection and classification algorithm using 64 line lidar,” Automotive Engineering, vol. 41, no. 7, pp. 779–784, 2019.
[5] D. Liu, Research on Particle Filter Localization and Map Creation Method for mobile RobotHunan University, Changsha, China, 2013.
[6] S. Thrun and M. Montemerlo, “The graph SLAM algorithm with applications to large-scale mapping of urban structures,” The International Journal of Robotics Research, vol. 25, no. 5-6, pp. 403–429, 2006.
[7] M. Montemerlo, S. Thrun, and D. Koller, “Fast slam: a factored solution to the simultaneous localization and mapping problem,” in Proceedings of the AAAI National Conference on Artificial Intelligence, pp. 593–598, AAAI, Menlo Park, CA, USA, November 2002.
[8] M. Li, B. Hong, and R. Luo, “The improved Rao Blackwellized particle filter is used to realize simultaneous localization and map creation of mobile robot,” Journal of Jilin University (Engineering Edition), vol. 37, no. 2, pp. 401–406, 2007.
[9] Y. Luo, L. Fu, and T. Cheng, “Simultaneous localization and mapping based on improved Rao Blackwellized particle filter,” Control theory and application, vol. 32, no. 2, pp. 267–272, 2015.
[10] X. Li, J. Li, and M. Zhu, “Research on location fusion and verification algorithm based on Unscented Kalman filter,” Automotive Engineering, vol. 43, no. 6, pp. 825–832, 2021.
[11] J. Graeter, A. Wilczynski, and M. Lauer, “Limo: lidar-monocular visual odometry,” in Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 7872–7879, IEEE, Madrid, Spain, October 2018.
[12] W. Shao, S. Vijayarangan, and C. Li, “Stereo visual inertial lidar simultaneous localization and mapping,” in Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 370–377, IEEE, Macau, China, November 2019.
[13] T. Shan, B. Englot, and C. Ratti, “Lvi-sam: tightly-coupled lidar-visual-inertial odometry via smoothing and mapping,” in Proceedings of the 2021 IEEE International Conference On Robotics And Automation (ICRA), Philadelphia, May 2022.
[14] D. Gálvez-López and J. D. Tardos, “Bags of binary words for fast place recognition in image sequences,” IEEE Transactions on Robotics, vol. 28, no. 5, pp. 1188–1197, 2012.