



# Map Fusion Method Based on Image Stitching for Multi-robot SLAM

Qirong Tang<sup>(✉)</sup>, Kun Zhang, Pengjie Xu, Jingtao Zhang<sup>(✉)</sup>, and Yuanzhe Cui

Laboratory of Robotics and Multibody System, School of Mechanical Engineering,  
Tongji University, Shanghai 201804, People's Republic of China

**Abstract.** Compared with the single-robot SLAM, the SLAM task completed by a multi-robot system in cooperation has the advantages of more accuracy, more efficiency and more robustness. This study focuses on the map fusion problem in the multi-robot SLAM task, which is to fuse the local maps created by multiple independent robots into an integrated map. A multi-robot SLAM map fusion method based on image stitching is therefore proposed. A single robot uses lidar SLAM to build a local environment map and upload it to a central node. The central node then maps each local map from a two-dimensional occupancy grid map to a grayscale image. The SuperPoint network is used to extract the depth features from the grayscale images, and the transformation relationships between the local maps are calculated via the feature matching. The matching topology graph is used to realize the final map fusion. It carries out experimental verification in the indoor environment on three mobile robots, which were developed by our own, and the experiment proved that the method has good real-time performance and robustness. After obtaining the global map, some new robots were placed in the environment, and realized the task of multi-robot target search by using the relocalization function.

**Keywords:** Multi-robot system · Simultaneous localization and mapping · Map fusion · Image stitching

## 1 Introduction

Simultaneous localization and mapping (SLAM) technology is the main solution to the key problem in the field of intelligent robot research that how the robot can obtain the map of an unknown environment while realizing its own localization in the map simultaneously, and provide technical support for subsequent tasks.

Due to the difficulties for a single robot to complete some complicated tasks, such as limited movement capability, insufficient computing power, poor anti-interference capability and so on, a multi-robot system is demanded. In the face of a large-scale complex environment, multiple robots can be dispersed into various sections of regions of the environment. A single robot uses SLAM technology to establish a local environment map/model. Then the system fuses these local

models into a larger-range environment model. Compared with the single-robot SLAM, the SLAM completed by a multi-robot system in collaboration has the advantages of more accuracy, more efficiency and more robustness. The main challenges in implementing a multi-robot SLAM system include bandwidth limitations, map fusion, asynchronous communications, coherent information integration, and data association between robots [1].

Most of the multi-robot SLAM methods are based on the corresponding single-robot SLAM algorithms. Earlier implementations are based on filtering methods, such as Extended Kalman Filter (EKF) [2,3]. These methods inherit some of the shortcomings of filter-based SLAMs. The current mainstream multi-robot SLAM algorithms are based on graph optimization or pose graph methods [4,5].

At present, most of the researches on the multi-robot SLAM problem are still in the simulation stage, where the information sensed and the communication conditions are often set to be completely ideal. This study focuses on the problem of multi-robot SLAM map fusion, proposes a new method based on image stitching [6], and conducts experimental verification in the real scene on the mobile robot platform developed by our own. With the help of ROS data transmission mechanism, our map fusion method does not depend on the specific single-robot SLAM algorithm, as long as the maps generated by the algorithm can be converted to grayscale images. It can process maps from any number of robots within the allowable range of computing power, and allows dynamic addition or removal of robots in the system.

The method closest to this work is proposed by Hörner [7]. We both draw on the principle of image stitching in computer vision. The difference is that this work uses depth features, and includes relocalization function using the global map. In addition, this work has been verified by experiments in the real scene.

The rest of this paper is organized as follows. Section 2 presents the method overview. The map fusion method based on image stitching for multi-robot SLAM is detailed in Sect. 3. The experiments are presented in Sect. 4, and conclusions are drawn in Sect. 5.

## 2 Method Overview

The proposed method is used to fuse 2D occupancy grid maps independently established by multiple robots, and method pipeline is already shown in Fig. 1.

Assume that there are  $n$  robots participating in the mapping task. The map built by a single robot through lidar SLAM is called a local occupancy grid map (short for local map)  $Lmap_i$  ( $i = 1, 2, \dots, n$ ), and the fused occupancy grid map is called a global map  $\mathcal{M}$ . The local environment maps are represented in the form of a common occupancy grid. One establishes a mapping relationship  $\mathcal{F} : OccupancyGridMap \rightarrow GrayscaleImage$  from occupancy grid map to grayscale image. Obviously, the proposed method also supports other formats of map, as long as the map format can be mapped to a grayscale image.

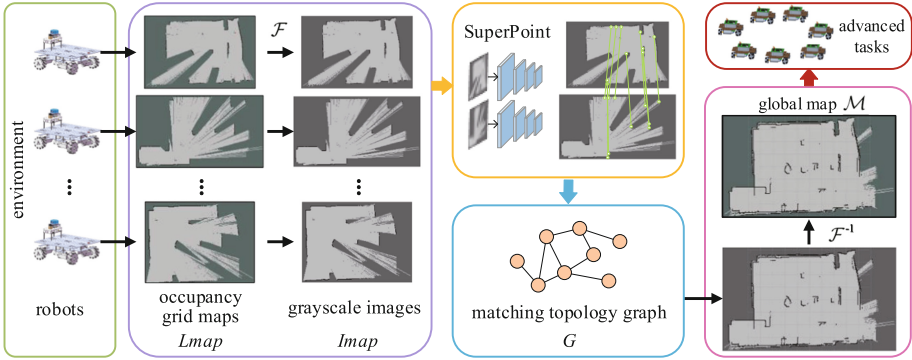


Fig. 1. Method pipeline

After  $Lmap_i$  has been converted into grayscale image  $Imap_i$  ( $i = 1, 2, \dots, n$ ), one then uses the neural network SuperPoint [8] to extract feature points and calculate descriptors, which can obtain higher image matching accuracy compared with traditional ORB [9] or SIFT [10]. After feature matching, it estimates the coordinate transformation relationships between local maps, that is, solves the homography matrices. The local maps and their matching relationships form a weighted graph, which is called the matching topology graph  $G$ . The largest spanning tree  $Tr$  is established in the largest connected component  $H$  of  $G$ , and map fusion is achieved by exploring  $Tr$ . Finally, the inverse mapping  $\mathcal{F}^{-1}$  needs to be used to convert the global map in the form of grayscale image back into an occupancy grid map  $\mathcal{M}$ .

The obtained global map can then provide support for the robot to perform other advanced tasks in the area, such as target search and path planning.

### 3 Multi-robot SLAM Map Fusion

#### 3.1 Estimate the Coordinate Transformations Between Local Maps

In order to fuse the local maps, calculating the transformations between the local occupancy grid maps established by each robot is the necessity. The method follows these steps:

- (1) Convert the local grid map  $Lmap_i$  ( $i = 1, 2, \dots, n$ ) built by each robot into a grayscale image  $Imap_i$  ( $i = 1, 2, \dots, n$ ). The value of each cell is within the range of  $[0, 100]$ . This value represents the probability of obstacles in the cell. If the probability is unknown, it is represented by  $-1$ . The local map in this form is converted to a grayscale image, and if the value is  $-1$ , it maps to 255 in the pixel value of the grayscale image, and one can get a standard 8-bit depth grayscale image.
- (2) Use the SuperPoint network to extract the feature points and calculate feature descriptors of each local map  $Imap_i$ .

- (3) Feature matching is carried out for each pair of local maps afterwards. The brute force matching algorithm is used. If there are a large number of robots, the parallel hierarchical clustering algorithm will be used to speed up the matching.
- (4) Solve the coordinate transformation matrix  $T_{i,j}$  between each pair of local grid maps ( $Imap_i, Imap_j$ ). Use the RANSAC [11] method to filter the matched features, use the SVD method to solve the  $T_{i,j}$ , and calculate the matching confidence  $c_{i,j}$  of the corresponding match

$$c_{i,j} = \frac{number\_inliers_{i,j}}{8 + 0.3 * number\_matches_{i,j}}, \quad (1)$$

where  $number\_inliers_{i,j}$  is the number of inliers found in the RANSAC method, and  $number\_matches_{i,j}$  is the number of matched feature points between each pair of local maps ( $Imap_i, Imap_j$ ).

- (5) Eliminate matches with a confidence less than 1, and form a matching topology graph  $G$ . The vertices of the  $G$  are the local maps  $Imap_i$  built by single robots, and the edges are  $T_{i,j}$ ,  $c_{i,j}$ ,  $number\_inliers_{i,j}$  and  $number\_matches_{i,j}$ .

Each time a robot updates the local map, it will upload the updated incremental map to the central node instead of uploading the entire local map, which can effectively reduce the amount of data transmission. Then, the central node will regenerate a corresponding grayscale image.

### 3.2 Map Fusion Based on Matching Topology Graph

After the matching topology graph is established, if any of the local maps has strong matching relationships with the others, the coordinate system of this local map will be fixed as the world coordinate system. The coordinate transformations will be performed on the remaining maps according to the results of feature matching. The global map is thus established.

If there are a large number of robots or there are maps that need to be eliminated, such as local maps with large errors, isolated local maps and so on, the graph method will be used for map fusion [7]. It is very common that some local grid maps cannot be successfully matched. In order to cover as large environmental area as possible, the weighted maximum connected component  $H$  in the established matching topology graph  $G$  is considered, and only the local maps included by  $H$  will be fused. This study selects the coordinate system of one of the local maps as the global coordinate system in  $H$ . The maximum spanning tree  $Tr$  is established in  $H$ , and by exploring  $Tr$ , the transformations between the local maps and the global coordinate system are finally determined. As for each local map, it can obtain the final transformation result by synthesizing paired transformations along the path. After all the transformations are completed, the map fusion is realized.

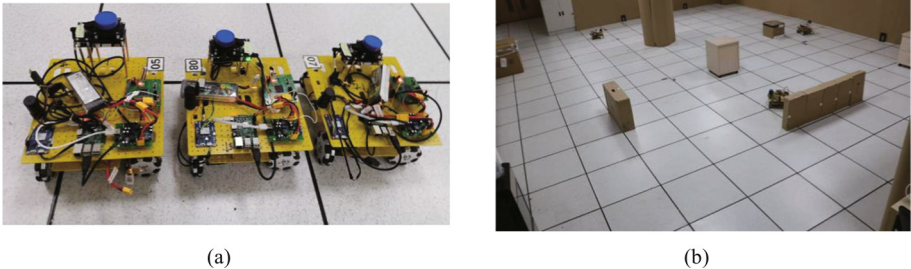
### 3.3 Robot Relocalization

Relocalization with maps built by lidar SLAM is often more difficult than the visual relocalization problem, because the information stored in the map established by lidar SLAM is not rich enough. At present, the relocalization problem using lidar usually uses the loop detection part of Cartographer [12] or Karto-SLAM [13]. In this study, after obtaining the global map  $\mathcal{M}$ , the newly entering robot  $k$  uses its own lidar to build a local grid map  $Rmap_k$ , and uses the method described in Sect. 3.1 to perform feature matching between the local grid map  $Rmap_k$  and the global map  $\mathcal{M}$ . The coordinate transformation between the local grid map and the global map is estimated, and then the position of the newly entering robot in the global coordinate system is obtained, so as to realize the relocalization and provide support for the robot to realize the tasks such as navigation and target search.

## 4 Experiments

### 4.1 Multi-robot Map Fusion Experiment

A real world experiment using three ground omnidirectional mobile robots as shown in Fig. 2(a) is conducted. The experimental area is an indoor environment, as shown in Fig. 2(b). The three mobile robots, each carrying a laser range scanner, are distributed in different areas to measure the environment and perceive environmental information. In the experiment, it uses remote control to make those robots move in a certain part of the field and build local maps. The map established by each robot covers only one part of the entire environment. The fused map on the central node includes the contents of all three local maps.

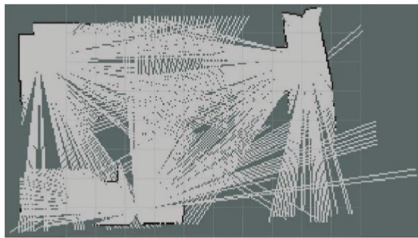


**Fig. 2.** (a) Mobile robot running single-robot SLAM. In order to avoid the lidar interference between each other, the lidars of the three robots are set up in different heights. (b) Experimental area.

Each robot uploads the environmental information obtained to the central node. In this experiment, the central node is only responsible for controlling the movement of the mobile robots and for performing map fusion, and single-robot

SLAM is implemented independently by mobile robots and does not depend on the central node. The central node is a laptop computer with Intel Core i7-8750H and 8G RAM in this case. The communication method between the central node and the mobile robots is WiFi, and our method is run based on ROS. The data format of the local map established by a single robot in ROS is *nav\_msgs/OccupancyGrid* message, and the data format of the corresponding incremental map updates is *map\_msgs/OccupancyGridUpdate* message. Common SLAM algorithms that support this format rule in ROS include Karto-SLAM [13], Gmapping [14], and so on.

The map fusion process observed in the display interface of the center node is shown in Fig. 3.



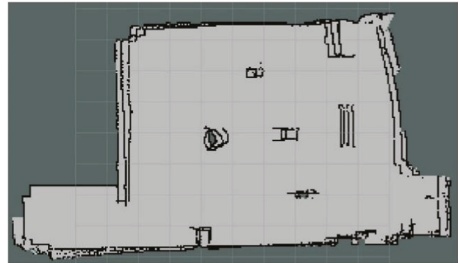
(a)  $t = 0s$ , program initialized.



(b)  $t = 60s$ , map fusion error was large.



(c)  $t = 95s$ , map was fused successfully.



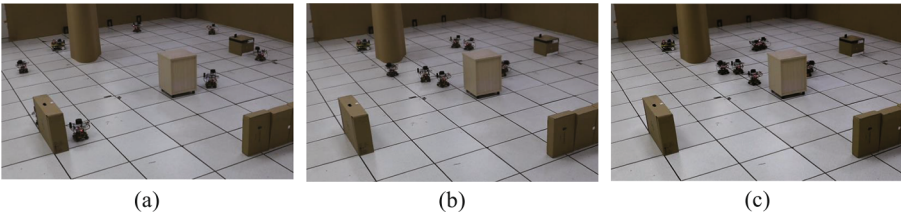
(d)  $t = 380s$ , mapping task completed.

**Fig. 3.** Map fusion in the exploration process of three ground mobile robots.

As shown in Fig. 3, after 380 s of exploration, the three ground mobile robots finally completed the task of building a global map. On the display interface of the central node, the process of localization and mapping by the mobile robots could be observed. Initially, the fusion error between the three local maps was relatively large due to the fact that there were few feature points. After 95 s, the fusion of the mobile robot maps achieved a good effect. After 380 s, each of the three robots covered most of the entire site, the global map contained enough area information, and the task was completed.

## 4.2 Multi-robot Target Search Experiment

In the same field, 7 new mobile robots are placed, and the configuration of these 7 mobile robots and the lidars they carried are different from the three robots used for mapping. In this experiment, the robots themselves do not have the localization function. It only detects environmental information and uploads it to the central node. The central node uses the global map that has been obtained before to provide localization service for the robots and guide them to perform collaborative search task. Seven mobile robots use the improved PSO algorithm for target search, and the specific implementation method refers to our previous work [15]. The experiment process is shown in Fig. 4.



**Fig. 4.** Multi-robot target search experiment. (a) Seven robots were in their initial positions. (b) Seven robots were searching. (c) Target search task completed.

## 5 Conclusions

In this study, a multi-robot SLAM map fusion method based on image stitching is proposed. In the method, the occupancy grid map is mapped to grayscale image, and the transformation relationships between maps are estimated through depth feature matching. On this basis, the map fusion is performed using the matching topology map. This method can realize the fusion of multiple local maps consuming fewer computing resources. We have carried out experiments on three mobile robots developed by our own in an indoor environment, and the experiments have proved that the method has good real-time performance and robustness. After obtaining the global map, some new robots are used in the environment and the multi-robot target search task is realized successfully by using the relocalization function.

Future works might include using more robots to conduct experiments in real scenes and to realize autonomous robotic exploration.

**Acknowledgements.** This work is supported by the projects of National Natural Science Foundation of China (No.61873192; No.61603277; No.61733001), the Quick Support Project (No.61403110321), and Innovative Project (No.20-163-00-TS-009-125-01). Meanwhile, this work is also partially supported by the Fundamental Research

Funds for the Central Universities and the Youth 1000 program project. It is also partially sponsored by International Joint Project Between Shanghai of China and Baden-Württemberg of Germany (No. 19510711100) within Shanghai Science and Technology Innovation Plan, as well as the projects supported by China Academy of Space Technology and Launch Vehicle Technology. All these supports are highly appreciated. The authors also would like to thank Zhongqun Zhang for his helpful suggestions.

## References

1. Lázaro, M., Paz, L., Piniés, P., Castellanos, J., Grisetti, G.: Multi-robot SLAM using condensed measurements. In: Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1069–1076. Tokyo, Japan (2013)
2. Roumeliotis, S., Bekey, G.: Distributed multirobot localization. *IEEE Trans. Robot. Autom.* **18**(5), 781–795 (2002)
3. Sasaokas, T., Kimoto, I., Kishimoto, Y., Takaba, K., Nakashima, H.: Multi-robot SLAM via information fusion extended Kalman filters. *IFAC-PapersOnLine* **49**(22), 303–308 (2016)
4. Dellaert, F., Kipp, A., Krauthausen, P.: A multifrontal QR factorization approach to distributed inference applied to multi-robot localization and mapping. In: Proceedings of the 20th National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, Vol. 3, pp. 1261–1266. Pittsburgh, USA (2005)
5. Chang, H., Lee, C., Hu, Y., Yung-Hsiang, L.: Multi-robot SLAM with topological/metric maps. In: Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1467–1472. San Diego, USA (2007)
6. Brown, M., Lowe, D.: Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vis.* **74**(1), 59–73 (2007)
7. Hörner, J.: Map-merging for multi-robot system. Bachelor’s thesis, Charles University in Prague (2016)
8. DeTone, D., Malisiewicz, T., Rabinovich, A.: SuperPoint: Self-supervised interest point detection and description. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 337–349. Salt Lake City, USA (2018)
9. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: An efficient alternative to SIFT or SURF. In: Proceedings of the 2011 International Conference on Computer Vision, pp. 2564–2571. Barcelona, Spain (2011)
10. Lowe, D.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
11. Fischler, M., Bolles, R.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **24**(6), 381–395 (1981)
12. Hess, W., Kohler, D., Rapp, H., Andor, D.: Real-time loop closure in 2D LIDAR SLAM. In: Proceedings of the 2016 IEEE International Conference on Robotics and Automation, pp. 1271–1278. Stockholm, Sweden (2016)
13. Konolige, K., Grisetti, G., Kümmerle, R., Burgard, W., Limketkai, B., Vincent, R.: Efficient sparse pose adjustment for 2D mapping. In: Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 22–29. Taipei, China (2010)



14. Grisetti, G., Stachniss, C., Burgard, W.: Improved techniques for grid mapping with Rao-Blackwellized particle filters. *IEEE Trans. Robot.* **23**(1), 34–46 (2007)
15. Tang, Q., Yu, F., Xu, Z., Eberhard, P.: Swarm robots search for multiple targets. *IEEE Access* **8**, 92814–92826 (2020)