Robotica (2016) volume 34, pp. 2516–2531. © Cambridge University Press 2015
doi:10.1017/S0263574715000168

Merging grid maps of different resolutions by scaling registration

Liang Ma†, Jihua Zhu‡*, Li Zhu‡, Shaoyi Du† and Jingru Cui‡

†*Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, P. R. China*
‡*School of Software Engineering, Xi'an Jiaotong University, P. R. China*

(Accepted February 17, 2015. First published online: March 20, 2015)

SUMMARY
This paper considers the problem of merging grid maps that have different resolutions. Because the goal of map merging is to find the optimal transformation between two partially overlapping grid maps, it can be viewed as a special image registration issue. To address this special issue, the solution considers the non-common areas and designs an objective function based on the trimmed mean-square error (MSE). The trimmed and scaling iterative closest point (TsICP) algorithm is then proposed to solve this well-designed objective function. As the TsICP algorithm can be proven to be locally convergent in theory, a good initial transformation should be provided. Accordingly, scale-invariant feature transform (SIFT) features are extracted for the maps to be potentially merged, and the random sample consensus (RANSAC) algorithm is employed to find the geometrically consistent feature matches that are used to estimate the initial transformation for the TsICP algorithm. In addition, this paper presents the rules for the fusion of the grid maps based on the estimated transformation. Experimental results carried out with publicly available datasets illustrate the superior performance of this approach at merging grid maps with respect to robustness and accuracy.

KEYWORDS: Multi-robot systems; Grid map merging; Scaling transformation; Iterative closest point; Image registration.

1. Introduction

Simultaneous localization and mapping (SLAM) is one of the most important and fundamental issues in robotics, and has been the object of much attention since the seminal work presented in ref. [1].

In previous literature,[2] most related studies have focused on the problem of building a single map for small- or medium-scale environments only. To build large scale maps, many researchers have taken an approach that uses cooperative mapping by multi-robot systems[3–11] because they have a superior performance in terms of mission efficiency and robustness. Therefore, it is necessary to merge the maps produced by different robots that explore various parts of the same environment.

A grid map is a kind of probabilistic map[2,12] that does not need to extract any special environmental features and can model arbitrary types of environments. Consequently, a grid map is one of the most popular map representations. So far, most merging approaches can only merge grid maps at the same resolution and are unable to determine whether one of the two maps needs to be magnified in order to be merged with the other. However, grid maps could be built at different resolutions by various robots in multi-robot systems. Suppose, for example, it is necessary to build a grid map for an abandoned mine.[35] Because a mine belongs to a large environment, multi-robot systems could be adopted to build the grid map. In general, some areas of the mine may include more details than others. For reasons of compactness, the map could be built at a fine resolution in the area with more details and at a coarse resolution in the area with fewer details. Thus, it is necessary to consider the problem of merging grid maps at different resolutions.

* Corresponding author. E-mail: zhujh@mail.xjtu.edu.cn

In fact, grid maps can be seen as images and the merging problem can be viewed as an image registration issue that can also be solved by the iterative closest point (ICP) algorithm or its variants.[13–19] For a pair of grid maps at different resolutions, edge points are extracted and the merging problem is transformed into a scaling registration of partially overlapping 2D point sets. For this particular registration issue, a trimmed mean-square distance function[14] is designed by introducing the backward distance measurement. Subsequently, the scaling transformation[16–18] involved in this objective function is calculated by the proposed TsICP algorithm, which converges monotonically to a local minimum from any given initial transformation. To obtain the optimal transformation, it is necessary to provide an initial scaling transformation for the TsICP algorithm. Accordingly, a good initial transformation can be estimated by geometrically consistent feature matches that are picked out by the RANSAC algorithm[20] from all the SIFT features[21,22] extracted and matched for the grid maps to be merged. In addition, this paper also presents the rules for the fusion of grid maps based on the obtained scaling transformation. In this way, the proposed approach merges grid maps of different resolutions.

The remainder of this paper is organized as follows. After reviewing the methods for merging maps that have the same resolution in Section 2, Section 3 presents a novel scaling registration approach for merging grid maps that have different resolutions. Section 4 then provides the implementation details of the proposed approach. In Section 5, the proposed approach is tested and evaluated on real robot datasets. Finally, the conclusion is drawn in Section 6.

2. Merging Grid Maps of the Same Resolution

As map merging is a good solution for building large scale maps, it has attracted increasing attention. Subsequently, different merging approaches have been proposed for the corresponding map representations.[3–11]

For maps represented by occupied cells, Carpin *et al.*[4] viewed map merging as an optimization problem and proposed a stochastic search approach to obtain a suitable rigid transformation to align two grid maps. In a subsequent work, Birk and Carpin[6] introduced a more sophisticated method to guide the search and presented some mechanisms to detect failures. Both of these approaches acquire the optimal solution. However, their computational requirements are sizeable because of the nature of exhaustive search. Researchers later proposed two related approaches[7,8] that share the same basic idea: when two robots meet during the mapping, their relative pose is determined and the two individual maps are merged into a combined one. The main difference between the approaches is whether the robots meet randomly or search each other out. The first approach cannot merge maps until robots can see each other, while the second approach requires robots to communicate with each other continuously. Meanwhile, Carpin[9] proposed an effective method based on the Hough transform.[23] Although this approach is extremely fast, it is not very accurate because of the discretization phenomena in the Hough transform. Recently, Zhu *et al.*[11] viewed grid map merging as a point set registration problem[13–19] that can be solved by the trimmed ICP algorithm.[14,15] In addition, Blanco *et al.*[24] proposed a novel matching method from the perspective of image registration. This method extracts and matches local features between two grid maps that can be used to estimate the rigid transformation for merging grid maps.

All of the above mentioned approaches can only merge grid maps that have the same resolution. However, the representation of multiple resolutions is necessary for building large scale maps.[25] Therefore, it is required to merge grid maps that have different resolutions.

3. Merging Grid Maps of Different Resolutions

In this section, the grid map merging problem is stated and transformed into an image registration issue that can be solved by the proposed approach.

3.1. Problem definition

According to ref. [9], the goal of grid map merging is to find a rigid transformation T such that two given maps P and Q can be correctly overlapped. Usually, transformation T is the combination of a rotation matrix $\mathbf{R}$, followed by a translation $\vec{t}$ along the x and y axes of magnitudes $\vec{t}_x$ and $\vec{t}_y$,

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

(a) (b)

Fig. 1. Edge extraction from a grid map: (a) grid map and (b) extracted edges.

respectively. More specifically, rotation matrix $\mathbf{R}$ is determined by the rotation angle θ as follows:

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{1}$$

In practice, each mobile robot in a multi-robot system explores different parts of the same environment and could independently produce grid maps at different resolutions. To address the resolution issue, a scale factor should be introduced to represent the resolution ratio of two grid maps to be merged. More exactly, the goal of map merging[18–20] is to determine a scaling transformation $T = (s, \mathbf{R}, \vec{t})$ that is a combination of a scale factor s, 2D rotation matrix $\mathbf{R}$, and 2D translation vector $\vec{t}$. After determining T, it should be applied to map P to obtain the transformed map P' indicated by $P' = TP$. To obtain the correct merging results, map Q and transformed map P' should be fused appropriately.

3.2. Mathematical model

During robot mapping, the environment is divided into a large number of regular cells. Each cell in a grid map contains an occupied probability to indicate its status. The value of the occupied probability is stored in a matrix that can be displayed in the form of image. In other words, one cell of a grid map can be viewed as an image pixel. Accordingly, the problem of grid map merging can be viewed as an image registration issue. Given two overlapping grid maps, non-reference map P, and reference map Q, the corresponding edge point sets $P \triangleq \{\vec{p}_i\}_{i=1}^{N_p}$ and $Q \triangleq \{\vec{q}_j\}_{j=1}^{N_q}$ can be obtained by applying an edge detection algorithm, where N_p and N_q denote the number of edge points extracted from P and Q, respectively. Here, the value of point $\bar{p}_i$ or $\bar{q}_j$ can be measured by the row and column number of the corresponding edge point in the grid map. Figure 1 illustrates the edge point set extracted from a grid map that has the following relationship:

$$\vec{q}_{c(i)} \triangleq s\mathbf{R}\vec{p}_i + \vec{t}, \tag{2}$$

where $\vec{q}_{c(i)}$ indicates the correspondence to edge point $\vec{p}_i$ in the reference map.

Similarly to the method of rigid registration,[13] the issue of grid map merging can be formulated as a least squares (LS) problem based on the forward distance measurement

$$\min_{\substack{s,\mathbf{R},\vec{t} \\ c(i)\in\{1,2,..,N_q\}}} \left\| s\mathbf{R}\vec{p}_i + \vec{t} - \vec{q}_{c(i)} \right\|_2^2 s.t. \quad s > 0, \mathbf{R}^T\mathbf{R} = \mathbf{I}_{2\times 2}, \det(\mathbf{R}) = 1. \tag{3}$$

However, this formulation is impractical for real grid map merging. Because the goal of map merging is to find the minimum of the objective function, it is easy to obtain the results shown in Fig. 2(a). As shown in this figure, when the scale factor s tends towards zero, all edge points in the non-reference map P gather together and find the same correspondence in the reference map Q. This means that the above LS formulation is an ill-posed problem[18] that can result in unexpected results. In addition, the grid maps being merged are produced by robots exploring different parts of the same environment,

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

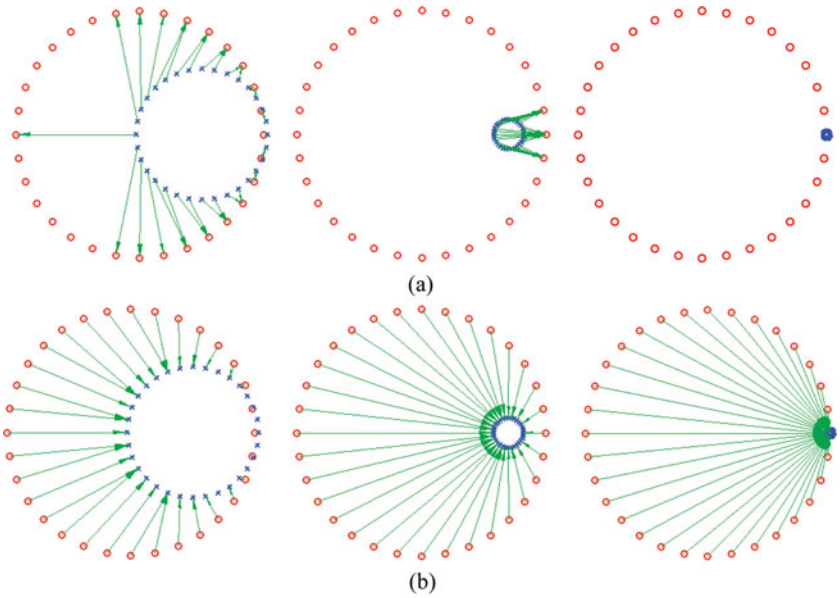


Fig. 2. Change in different objective functions as $s \to 0$, where the small red circles denote edge points of Q and the blue crosses indicate edge points of P: (a) Eq. (3) based on the forward distance measurement and (b) Eq. (4) based on the backward distance measurement.

hence these two point sets only partially overlap, and the non-common areas should be discarded to obtain an accurate scaling transformation for map merging.

To avoid an ill-posed problem, the grid map merging issue should be viewed as a scaling registration problem, where the backward distance measurement can be introduced into the objective function as follows:

$$\min_{\substack{s,\mathbf{R},\vec{t} \\ c(j)\in\{1,2,..,N_p\}}} \left\| s\mathbf{R}\vec{p}_{c(j)} + \vec{t} - \vec{q}_j \right\|_2^2 s.t. \quad s > 0, \mathbf{R}^T\mathbf{R} = \mathbf{I}_{2\times 2}, \det(\mathbf{R}) = 1. \tag{4}$$

In Eq. (4), when s tends towards zero, its numerator increases because of the backward distance measurement. The effectiveness of Eq. (4) is displayed in Fig. 2(b), where the proposed objective function increases as s tends towards zero. Accordingly, the ill-posed problem inherited in Eq. (3) can be avoided by the backward distance measurement.

As edge point sets extracted from grid maps partially overlap, the overlap percentage ξ should be considered in the objective function. Suppose Q_ξ denotes a subset of edge points Q that are extracted from the common area between the two given grid maps. The overlap percentage is calculated as follows:

$$\xi = \frac{\left|Q_\xi\right|}{|Q|}, \tag{5}$$

where $|\cdot|$ represents the cardinality of a set that is equal to the number of elements in the set. Subsequently, we propose the following objective function to merge two grid maps that have different resolutions:

$$\min_{\substack{s,\mathbf{R},\vec{t},\xi,Q_\xi \\ c(j)\in\{1,2,..,N_p\}}} \frac{1}{\left|Q_\xi\right|(\xi)^{1+\lambda}} \sum_{\vec{q}_j\in Q_\xi} \left\| s\mathbf{R}\vec{p}_{c(j)} + \vec{t} - \vec{q}_j \right\|_2^2 s.t. \quad \mathbf{R}^T\mathbf{R} = \mathbf{I}_{2\times 2}, \det(\mathbf{R}) = 1$$

$$\xi \in [\xi_{\min}, 1], Q_\xi \subseteq Q, \left|Q_\xi\right| = \xi\,|Q|\,, \tag{6}$$

where λ is the control parameter and $\xi_{\min}$ denotes the allowed minimum value of ξ.

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

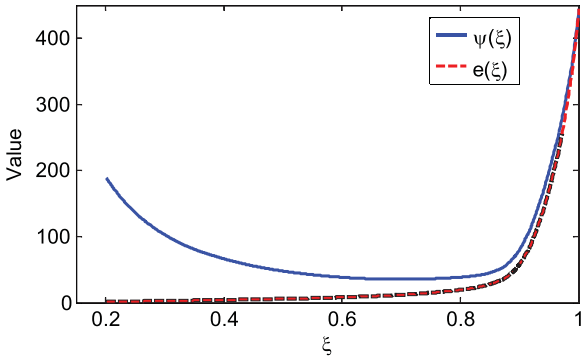


Fig. 3. Typical shape of the objective function $\psi(\xi)$ and MSE $e(\xi)$.

To show how the non-common areas can be rejected by introducing ξ, Eq. (6) can be represented as follows:

$$\psi(\xi) = \frac{e(\xi)}{(\xi)^{1+\lambda}}, \tag{7}$$

where $e(\xi)$ denotes the MSE

$$e(\xi) = \frac{1}{|Q_\xi|} \sum_{\vec{q}_j \in Q_\xi} \left\| s\mathbf{R}\vec{p}_{c(j)} + \vec{t} - \vec{q}_j \right\|_2^2. \tag{8}$$

Suppose the correspondence $\vec{p}_{c(j)}$ for edge point $\vec{q}_j$ has been established and all the established point pairs have been sorted in ascending order by their distances. The typical shape of the objective function $\psi(\xi)$ and MSE $e(\xi)$ is shown in Fig. 3, where $e(\xi)$ increases gradually until ξ passes the true overlap percentage. Accordingly, $\psi(\xi)$ first decreases and then increases when the true overlap percentage is exceeded. Therefore, the true overlap percentage can be obtained by finding the minimum of $\psi(\xi)$. After the optimal overlap percentage has been determined, the non-common areas are easily rejected using the sorted order of point pairs. The above process is implemented in the TsICP algorithm.

3.3. TsICP algorithm

The solution of Eq. (6) is clearly important for grid map merging. In practice, it can be solved by the TsICP algorithm, a variant of the original ICP algorithm. Given the initial transformation $T_0 = (s_0, \mathbf{R}_0, \vec{t}_0)$, one iteration of this algorithm consists of three steps

Step 1: Assign the correspondence for edge point $\vec{q}_j$ using the $(k-1)$th transformation

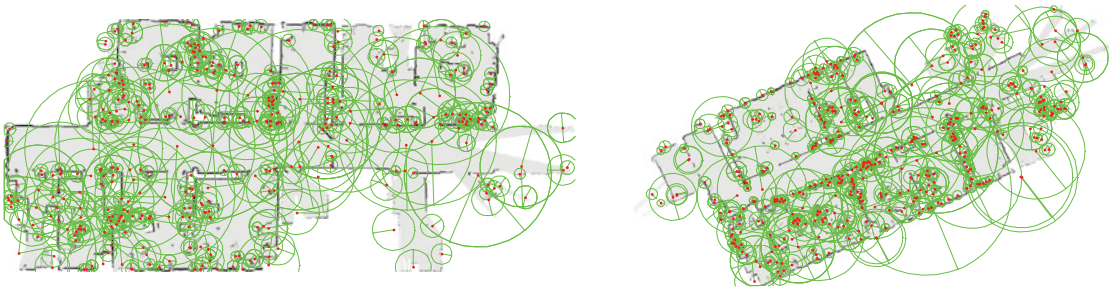
$$c_k(j) = \underset{i \in \{1,2,\ldots,N_p\}}{\arg\min} \left\| s_{k-1}\mathbf{R}_{k-1}\vec{p}_i + \vec{t}_{k-1} - \vec{q}_j \right\|_2 \quad (j = 1, 2, \ldots, N_q). \tag{9}$$

In fact, Eq. (9) represents the nearest neighbor search problem that can be solved by many efficient methods.[26–29] In this paper, the nearest neighbor search method based on the *k-d* tree is adopted to assign correspondence $\vec{p}_{c_k(j)}$ to edge point $\vec{q}_j$. In addition, the distance between $\vec{q}_j$ and $(s_{k-1}\mathbf{R}_{k-1}\vec{p}_{c_k(j)} + \vec{t}_{k-1})$ should be preserved for subsequent processing.

Step 2: Update the kth overlapping parameter ξ_k and its corresponding subset Q_{ξ_k}

$$(\xi_k, Q_{\xi_k}) = \underset{\xi_{\min} \le \xi \le 1}{\arg\min} \sum_{\vec{q}_j \in Q_\xi} \left\| s_{k-1}\mathbf{R}_{k-1}\vec{p}_{c_k(j)} + \vec{t}_{k-1} - \vec{q}_j \right\|_2^2 / (|Q_\xi|(\xi)^{1+\lambda}). \tag{10}$$

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press



(a) (b)

Fig. 4. SIFT features extracted from the potentially merging maps: (a) reference map and (b) non-reference map.

According to ref. [15], Eq. (10) can be solved in a sequence processing manner by sorting the edge point pairs $\{\vec{q}_j, (s_{k-1}\mathbf{R}_{k-1}\vec{p}_{c_k(j)} + \vec{t}_{k-1})\}_{j=1}^{N_q}$ by their distances in ascending order and letting

$$\psi(\xi) = \sum\nolimits_{\vec{q}_j \in Q_\xi} \left\| s_{k-1}\mathbf{R}_{k-1}\vec{p}_{c_k(i)} + \vec{t}_{k-1} - \vec{q}_j \right\|_2^2 / (|Q_\xi| \xi^{1+\lambda}). \tag{11}$$

At each iteration, a pair of sorted points is added to compute the corresponding value of $\psi(\xi)$. By traversing all sorted edge point pairs, it is easy to obtain the minimum value $\psi(\xi_k)$ that corresponds to the optimal overlap percentage ξ_k. The points involved in the front $\xi_k N_q$ sorted point pairs are then selected to update the corresponding point subset Q_{ξ_k}

Step 3: Calculate the current scaling transformation $T_k = (s_k, \mathbf{R}_k, \vec{t}_k)$

$$(s_k, \mathbf{R}_k, \vec{t}_k) = \arg\min_{s,\mathbf{R},\vec{t}} \sum_{\vec{p}_j \in Q_{\xi_k}} \left\| s\mathbf{R}\vec{p}_{c_k(j)} + \vec{t} - \vec{q}_j \right\|_2^2. \tag{12}$$

Although the calculation of the rigid transformation between two point sets is a well-studied problem,[30] the calculation of the scaling transformation is more difficult. To solve Eq. (12), the singular value decomposition technique[16] is utilized.

The TsICP algorithm can be proven to be locally convergent in theory and a detailed proof is presented in the Appendix.

3.4. Analysis of the initial transformation

Because the TsICP algorithm is locally convergent, it requires a good initial transformation to obtain the desired global minimum. Hence, an effective method is presented to provide the initial transformation for the TsICP algorithm. This method first extracts local features from the grid maps and then matches these features between map pairs to estimate the initial transformation. To do this, the local features should satisfy the two following conditions: (1) the same features can be independently detected in both grid maps, and (2) for a feature in one map, the corresponding feature can be correctly recognized in the other map.

As the grid maps to be merged are built in different resolutions and coordinate systems, the extracted features should be invariant to scale, rotation, and translation changes. Therefore, we decided to adopt the SIFT feature, which is geometrically invariant under different scaling transformations. As Fig. 4 shows, a large number of SIFT features can be extracted from the pair of grid maps to be merged. For a pair of such maps, it is easy to find a set of feature matches. Because of sensor noise and the precision of the SLAM algorithm, it is inevitable that some feature matches will be geometrically inconsistent. Accordingly, the RANSAC algorithm is used to choose the inliers from all the feature matches.

Suppose N pair-wise features $\{F_{i,Q}, F_{i,P}\}_{i=1}^{N}$ have been extracted and matched for two grid maps to be merged, where $\vec{f}_{i,Q}$ and $\vec{f}_{i,P}$ represent the positions of SIFT features $F_{i,Q}$ and $F_{i,P}$, respectively. As two SIFT features involved in one consistent match correspond to the same position in the

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

environment, it is obvious that

$$\|s_0\mathbf{R}_0\vec{f}_{i,P} + \vec{t} - \vec{f}_{i,Q}\|_2 \approx 0, \tag{13}$$

where $\mathrm{T}_0 = (s_0, \mathbf{R}_0, \vec{t}_0)$ denotes the best initial scaling transformation. According to Eq. (12), there are at least two feature matches required to estimate the initial transformation. Setting $r = 0$ and $N_{\text{best}} = 0$, the RANSAC algorithm can be used to estimate the initial transformation as follows:

Step 1: Let $r = r + 1$, and choose two feature matches $\{F_{i,Q}, F_{i,P}\}_{i=m,n}$ at random from all feature matches;
Step 2: Substitute $\{\vec{f}_{i,Q}, \vec{f}_{i,P}\}_{i=m,n}$ into Eq. (12) and obtain the scaling transformation $\mathrm{T}_r = (\mathrm{s}_r, \mathbf{R}_r, \vec{t}_r)$;
Step 3: Apply $\mathrm{T}_r = (s_r, \mathbf{R}_r, \vec{t}_r)$ to all SIFT features $\{F_{i,P}\}_{i=1}^{N}$ of map P and calculate the distance $d_{i,r} = \|s_r\mathbf{R}_r\vec{f}_{i,P} + \vec{t} - \vec{f}_{i,Q}\|_2$ for all feature matches;
Step 4: Given threshold d_{thr}, find the number N_r of geometrically consistent feature matches that satisfy $d_{i,r} \le d_{thr}$. If $N_r > N_{\text{best}}$, set $N_{\text{best}} = N_r$ and store these feature matches;
Step 5: Repeat Steps 1–4, until $r = 200$. The N_{best} geometrically consistent feature matches can then be used to estimate the initial scaling transformation using Eq. (12).

Theoretically, the calculation of the initial transformation only requires two consistent feature matches. However, if the number of matches is less than three, there is no way to check and obtain the correct initial transformation.

3.5. Fusion of grid maps

The proposed approach is a convenient means to obtain an accurate scaling transformation for two grid maps to be merged. The scaling transformation can then be applied to the non-reference map to produce the transformed map. To obtain a suitable merging result, the reference map and transformed map should be correctly fused. Before map fusion, it is necessary to define an empty map M with its size determined by the superposition of the reference map and the transformed map. In addition, the origin of the empty map should be attached to the origin of the reference map. Because the merging map is a fusion of the reference and transformed maps, the occupied probability p_{ij}^{M} of pixel (i, j) in M can be determined by

$$p_{ij}^{M} = h\big(p_{ij}^{Q}, p_{i'j'}^{P}\big), \tag{14}$$

where $h(\cdot, \cdot)$ represents the fusion function, p_{ij}^{Q} denotes the status of cells in Q with index (i, j), and $p_{i'j'}^{P}$ denotes the status of cells in P with index (i', j'), which can be calculated as follows:

$$\begin{pmatrix} i' \\ j' \end{pmatrix} = \left\lfloor (s\mathbf{R})^{-1} \cdot \left(\begin{pmatrix} i \\ j \end{pmatrix} - \vec{t} \right) \right\rfloor. \tag{15}$$

In the grid map, the value of occupied probability p_{ij} can be used to judge the status for each cell. The corresponding fusion rules are displayed in Table I. Suppose $p_{ij} < 0.495$ indicates that the cell at (i, j) is free, $p_{ij} \in [0.495, 0.505]$ denotes an unknown status, and $p_{ij} > 0.505$ indicates that it is occupied. Accordingly, the fusion function $h(\cdot, \cdot)$ can be formulated as:

$$h\big(p_{ij}^{Q}, p_{i'j'}^{P}\big) = \begin{cases} \max(p_{ij}^{Q}, p_{i'j'}^{P}) & \max(p_{ij}^{Q}, p_{i'j'}^{P}) > 0.5 \\ \min(p_{ij}^{Q}, p_{i'j'}^{P}) & \max(p_{ij}^{Q}, p_{i'j'}^{P}) \le 0.5. \end{cases} \tag{16}$$

In this way, it is easy to obtain suitable merging maps.

4. Implementation

In practical application, the overlap percentage between two grid maps should be larger than a certain value. Otherwise, there is no guarantee that enough consistent feature matches for estimating the initial transformation will be obtained, and the TsICP algorithm may become trapped in a local minimum. In practice, we found that setting $\xi_{\min} = 0.35$ in Eq. (6) guarantees enough matches to

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

Table I. Rules for grid map fusion.

Q \ P	Unknown	Free	Occupied
Unknown	Unknown	Free	Occupied
Free	Free	Free	Occupied
Occupied	Occupied	Occupied	Occupied

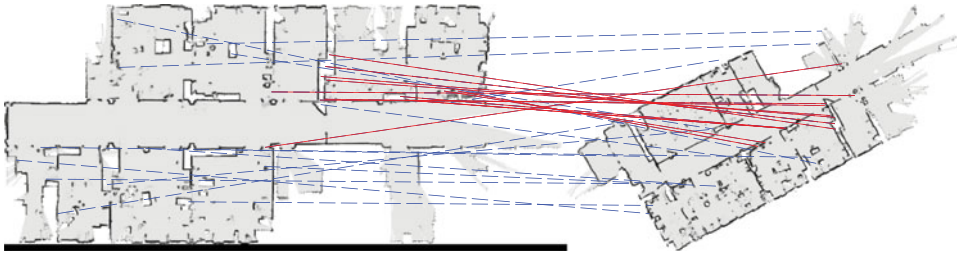


Fig. 5. SIFT features extracted and matched for a pair of grid maps, where solid red lines denote the geometrically consistent feature matches and dashed blue lines indicate the inconsistent feature matches.

acquire robust merging results. From the above discussion, we outline the proposed approach to merge grid maps of different resolutions in Algorithm 1.

Algorithm 1. Merging grid maps of different resolutions.

Input: Reference map Q and non-reference map P

(1) Extract SIFT features from Q and P and acquire feature matches $\{F_{i,Q}, F_{i,P}\}_{i=1}^{N}$;
(2) Estimate initial transformation $\mathrm{T}_{\mathrm{best}}$ by the method in Section 3.4;
(3) Extract the edge point sets $P \triangleq \{\vec{p}_i\}_{i=1}^{N_p}$ and $Q \triangleq \{\vec{q}_j\}_{j}^{N_q}$, set $\mathrm{T}_0 = \mathrm{T}_{\mathrm{best}}$, and calculate the optimal transformation by the method in Section 3.2;
(4) Based on this transformation, obtain the merged map M by the method in Section 3.5;

Output: Merged map M

5. Experimental Results

To verify the performance of the proposed approach, experiments were performed on two public datasets: "Fr079.log"[31] and "Loop25.log,"[32] which were recorded by mobile robots equipped with a laser range finder and odometer. Both were recorded in indoor environments. To simulate multi-robot systems, one dataset can be divided into two parts that can then be utilized to build the corresponding grid maps for testing the proposed approach. The proposed approach was compared with two other related approaches that appeared in refs. [9] and [24]. For simplicity, these two approaches are abbreviated as "Carpin" and "Blanco," respectively. All code was written in MATLAB and run on the same PC.

5.1. Validation

To verify its validity, the proposed approach was tested on Fr079.log. By dividing this dataset into two parts, the SLAM algorithm was utilized to build two grid maps at different resolutions. To merge this pair of grid maps, SIFT features were extracted and matched. Figure 5 displays the matched SIFT features extracted from the pair of grid maps, where the resolution of the reference (left) map is 5 cm and the resolution of the non-reference (right) map is 8 cm. As shown in Fig. 5, there are many geometrically inconsistent feature matches between these two maps because of sensor noise and the precision of the SLAM algorithm. To estimate the initial transformation, geometrically consistent feature matches were chosen from all the feature matches. Subsequently, the RANSAC algorithm was used.

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

Fig. 6. Map fusion results for different scaling transformations: (a) scaling transformation estimated by all geometrically consistent SIFT matches and (b) scaling transformation obtained by the TsICP algorithm.

As enough consistent feature matches were obtained from the RANSAC algorithm, it was easy to estimate the initial transformation by solving Eq. (12). Subsequently, the TsICP algorithm was used to calculate an accurate scaling transformation. Based on the obtained scaling transformation, the fusion rule was applied to produce the merged grid maps. Figure 6 displays the map fusion results for different scaling transformations for the pair of grid maps. As Fig. 6(a) shows, the scaling transformation estimated by some geometrically consistent SIFT matches was not very accurate because of the noise in the SIFT positions. The TsICP algorithm used this transformation as an initial scaling transformation to obtain a better one. As Fig. 6(b) shows, the scaling transformation obtained by the TsICP algorithm is very accurate. In addition, Fig. 6 also depicts that the proposed fusion rule can lead to reasonable merging results for maps that have different resolutions, verifying the effectiveness of the proposed approach.

To illustrate its convergence domain, the TsICP algorithm can be tested on the pair of grid maps displayed in Fig. 5 at different initial scaling transformations that can be obtained by applying synthetic noise to the scaling transformation estimated from all geometrically consistent SIFT matches. During the experiments, rotation and scale was disturbed by values uniformly drawn from the interval $[-8^\circ,\ 8^\circ]$ and $[0.92,\ 1.08]$, respectively. Consequently, the TsICP algorithm was used to minimize the objective function denoted by Eq. (6). The final value of the objective function for different initial scaling transformations is displayed in Fig. 7. As this figure shows, the TsICP algorithm is locally convergent over a wide domain. When the disturbance falls into the wide area close to $(0^\circ,\ 1)$, the TsICP algorithm always converges to the desired global minimum. Thus, the proposed approach can obtain the optimal solution when it is provided with a good initial scaling transformation.

For the proposed approach, more than two geometrically consistent feature matches are required to estimate the initial scaling transformation. Otherwise, there is no guarantee that a good initial scaling transformation will be obtained. Therefore, the proposed approach should be tested on two potentially mergeable grid maps both at different resolutions and rotations.

For the second experiment, Fr079.log was divided into two parts that were used to build two grid maps at different resolutions that varied within the interval^{2,10} cm. To view the two parts of Fr079.log in a more intuitive way, one corresponding map at 5 cm and another map at 8 cm are displayed in Fig. 5. For each pair of grid maps, geometrically consistent feature matches were detected by the proposed approach. The merged results for grid map pairs at different resolutions are illustrated in Fig. 8. As this figure shows, the proposed approach is susceptible to the resolution. To merge two grid maps that have a large resolution difference, the proposed approach is probably unable to detect enough geometrically consistent feature matches to estimate the initial scaling transformation. To address this issue, it is reasonable to set the reference map to a medium resolution (such as 5 cm) to enable the grid maps to merge for small- or medium-resolution differences.

In fact, a homogeneous transformation can include all three elements in the scaling transformation, hence the method can also use a homogeneous transformation to merge grid maps. Because a

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

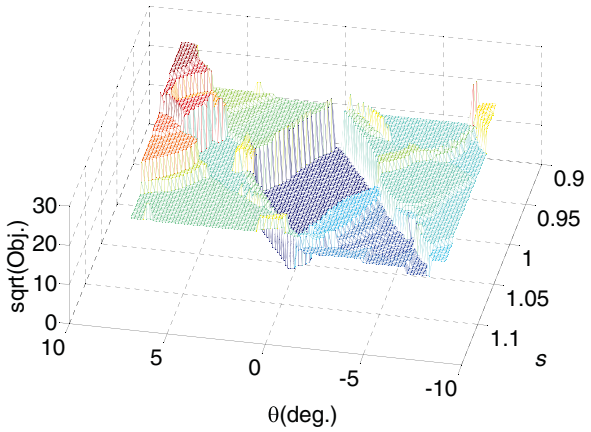


Fig. 7. Objective value obtained by the TsICP algorithm for different scaling transformations.

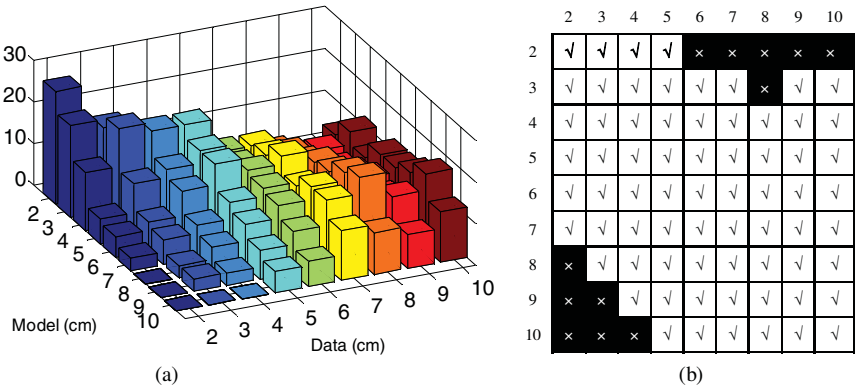


Fig. 8. Merged results of the grid map pairs at different resolutions, where the numbers in the vertical and horizontal directions denote the resolutions of the reference and non-reference maps, respectively: (a) number of geometrically consistent feature matches between grid map pairs and, (b) merge results for grid map pairs, where light cells indicate a successful merge and dark cells indicate a failed merging.

homogeneous transformation includes eight free parameters, this method requires four consistent feature matches to estimate the initial transformation. However, a scaling transformation includes four free parameters, and the proposed approach only needs two consistent feature matches to estimate the initial transformation. As there may not be enough matches to merge grid maps that have different resolutions, the proposed approach can improve the robustness of map merging.

In addition, the proposed approach was applied to merge the reference map with the rotated non-reference maps at different resolutions. Figure 9 displays the merging results for all grid map pairs that were merged, indicating whether the proposed approach can successfully merge a map pair or not. As depicted in Figs. 9 and 10, the proposed approach is robust to rotation.

5.2. Comparison

In this section, the proposed approach is compared with two other merging approaches. As few approaches can merge grid maps at different resolutions, the comparison was implemented in two ways: (1) merging grid maps that have the same resolution, and (2) merging grid maps that have different resolutions. In the following experiments, the resolution of all the reference maps was 5 cm.

5.2.1. Grid maps of the same resolution. Here, the proposed approach was tested on four pairs of grid maps at the same resolution and compared with the results of the other two approaches. For each of the data sets, the SLAM algorithm was utilized to build two pairs of grid maps to be merged at the same resolution (5 cm): one pair with a low overlap percentage and the other with a high overlap percentage. These pairs of grid maps were then merged by different approaches. To compare these approaches, the runtime, value of the objective function, and other related information are listed in

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

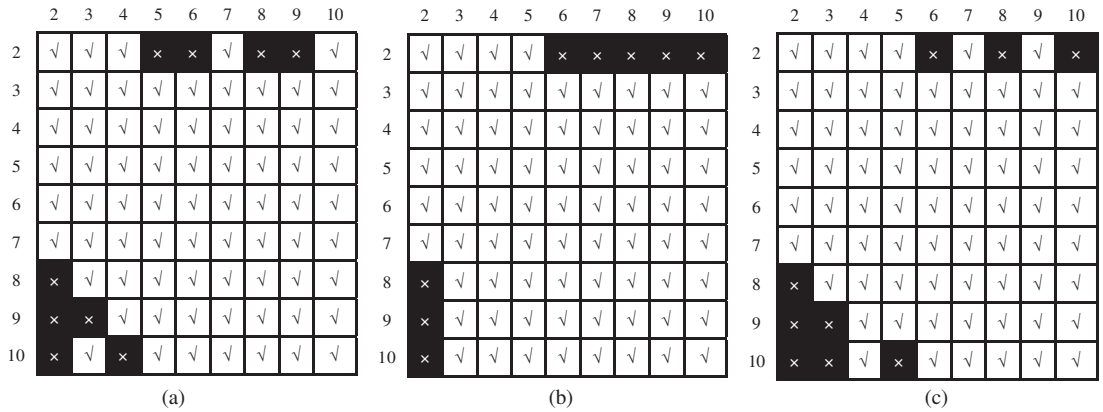


Fig. 9. Merge results for the reference map and rotated non-reference map: (a) rotated by 90° , (b) rotated by 180° , and (c) rotated by 270°.

Tabel II. Comparison result of different merging approaches tested on grid maps at the same resolution.

			Carpin[9]			Blanco[24]			Our method		
Dataset	ξ	Res. (cm)	Obj.	T (s)	Suc.	Obj.	T (s)	Suc.	Obj.	T (s)	Suc.
Fr079	Low	5/5	35.3365	5.9895	N	12.2950	8.4938	Y	4.9409	2.9248	Y
	High	5/5	5.7399	9.4423	Y	3.0650	7.6949	Y	1.8167	4.6624	Y
Loop25	Low	5/5	77.1152	6.8712	N	3.6602	6.3585	Y	0.7914	3.8805	Y
	High	5/5	16.4362	10.1076	Y	1.2788	11.4999	Y	0.7538	7.2926	Y

Table III. Comparison of different merging approaches tested on grid maps that have different resolutions.

			Carpin[9]			Blanco[24]			Our method		
Dataset	ξ	Res. (cm)	Obj.	T (s)	Suc.	Obj.	T (s)	Suc.	Obj.	T (s)	Suc.
Fr079	Low	5/2	1.4949×10^4	14.2515	N	2.3590×10^4	33.6656	N	4.2264	10.4738	Y
	High	5/2	8.2760×10^4	16.8074	N	1.2601×10^4	16.1521	N	1.2701	10.9501	Y
Loop25	Low	5/10	93.5330	5.0153	N	618.4076	3.7298	N	12.0835	2.8694	Y
	High	5/10	251.7081	6.1079	N	487.0942	4.7689	N	4.2348	3.4103	Y

Table II. To visualize the results in a more intuitive way, two examples are displayed in Figs. 10 and 11, where the pair of maps depicted in Fig. 10 have a low overlap percentage and the pair of maps shown in Fig. 11 have a high overlap percentage.

As these results show, Carpin's approach is not very robust and is less accurate than the other two approaches because of the discretization phenomenon inherent in the Hough transformation. Because the other two approaches take an image registration approach and share similar principles, they perform almost the same when merging grid maps that have the same resolution. Compared with Blanco's approach, the proposed approach is more accurate because it accounts for the overlap percentage in the objective function.

5.2.2. Grid maps of different resolutions. Here, the proposed approach was tested on four pairs of grid maps that have different resolutions and compared with the results of the other two approaches. These four pairs of grid maps are similar to those used in Section 5.2.1, except for the resolution of the non-reference maps. Accordingly, these pairs of grid maps were merged by the different approaches. To compare these approaches, the runtime, the value of the objective function, and other related information are also listed in Table III. Note that the map resolutions are not the input for any merging approaches. To visualize the results in a more intuitive way, two examples of the map merging results are displayed in Figs. 12 and 13, where the pair of maps depicted in Fig. 12 has a high overlap percentage and the pair of maps shown in Fig. 13 has a low overlap percentage.

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

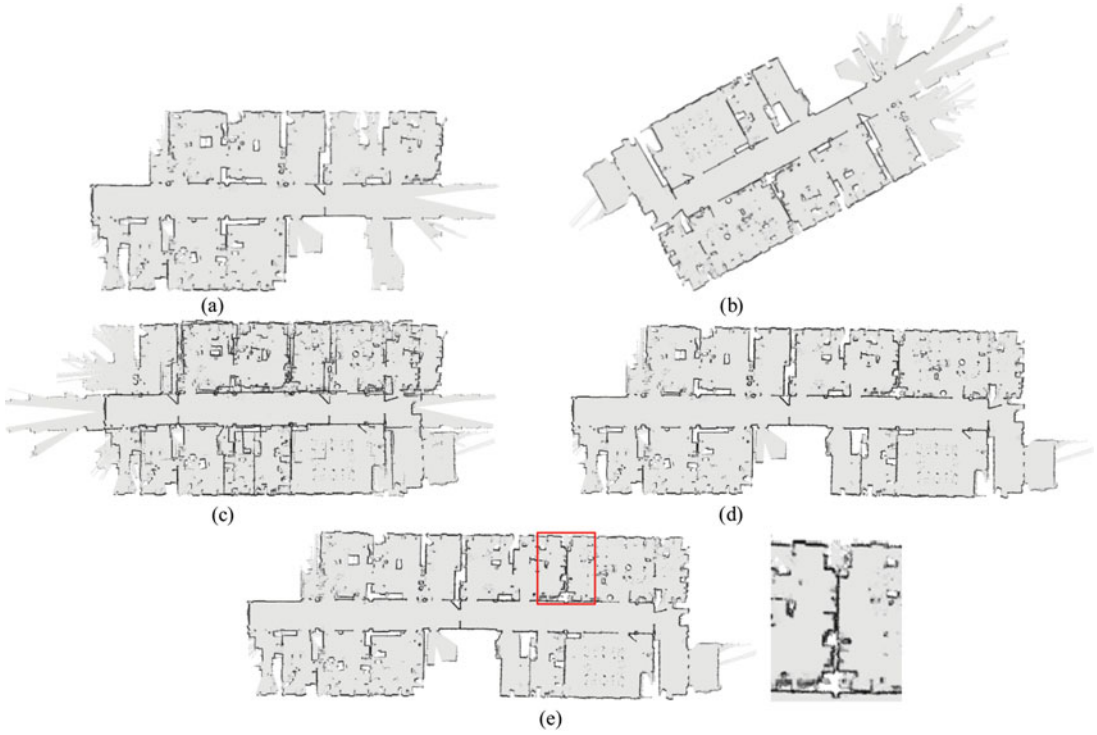


Fig. 10. Merge results of two maps generated from Fr079.log: (a) reference map (5 cm), (b) non-reference map (5 cm), (c) result of Carpin, (d) result of Blanco, and (e) result of the proposed method with enlarged detail.

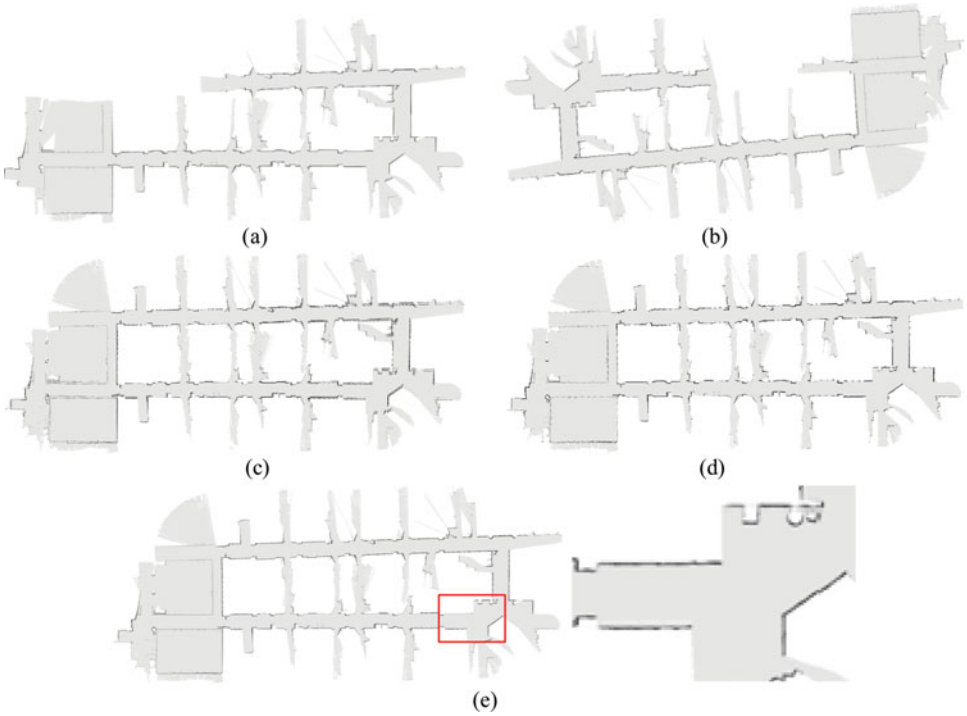


Fig. 11. Merge results of two maps generated from Loop25.log: (a) reference map (5 cm), (b) non-reference map (5 cm), (c) result of Carpin, (d) result of Blanco, and (d) result of the proposed method with enlarged detail.

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

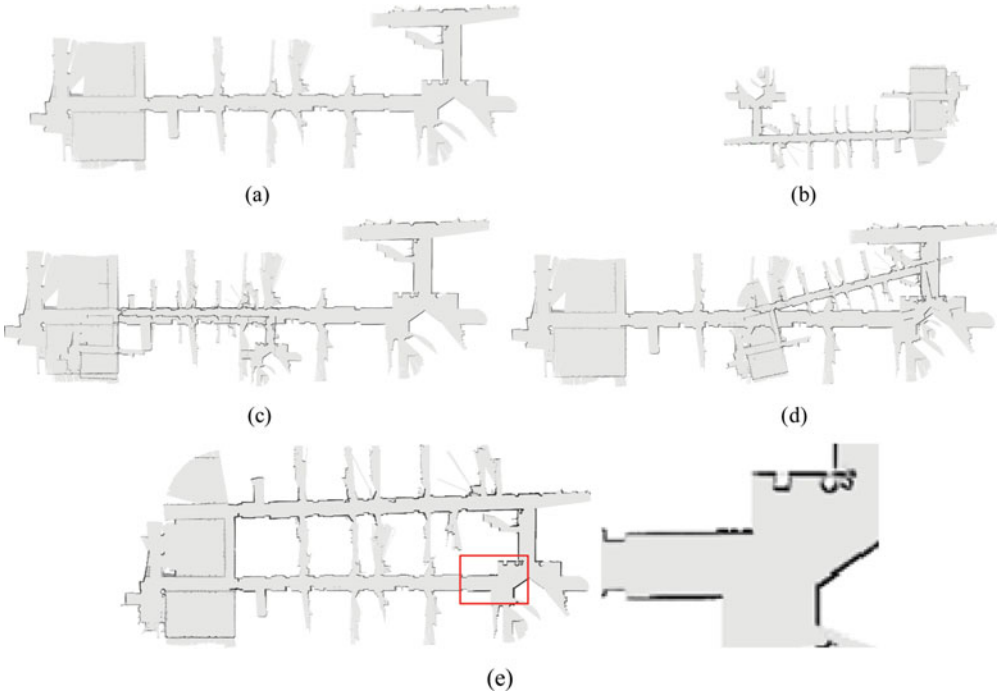


Fig. 12. Merge results of two grid maps generated from Loop25.log: (a) model grid map (5 cm), (b) data grid map (10 cm), (c) result of Carpin, (d) result of Blanco, and (d) result of the proposed method with enlarged detail.

Fig. 13. Merge results of the proposed approach applied to two grid maps that have different resolutions: (a) reference map (5 cm), (b) non-reference map (2 cm), and (d) result of the proposed method with enlarged detail.

As these results show, both the other approaches failed to merge all pairs of grid maps even when they had a high overlap percentage. By introducing the scale factor into the objective function, the proposed approach can merge all pairs of grid maps, even those at different resolutions. Because the non-reference map should be amplified or reduced to align well with the reference map, the corresponding merged map may appear to be artificial, where the width of the edge represented by the occupied cells may be inconsistent over different areas.

In a word, the proposed approach is superior for merging grid maps than the other approaches.

6. Conclusions

This paper is, to the best of our knowledge, the first that proposes an approach for merging 2D grid maps at different resolutions. The main contributions of this paper are as follows: (1) The grid map

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

merging problem is viewed as a special case of image registration, and the corresponding objective function is designed taking into account the scaling transformation for merging grid maps that have different resolutions. (2) To solve the objective function, the TsICP algorithm is proposed and proven to be locally convergent in theory. (3) To obtain the optimal solution, the TsICP algorithm is given an initial scaling transformation estimated by the RANSAC algorithm to select the consistent SIFT feature matches extracted from the map pair. (4) Based on the optimal scaling transformation, the rules of grid map fusion are presented to acquire the appropriate map merging results. The proposed approach has been implemented and tested on real robot data sets, and the experimental results illustrate that it is effective for merging grid maps that have both the same and different resolutions.

The proposed approach includes some limitations. First, the pair of grid maps to be merged must contain a certain degree of overlap. Otherwise, it is difficult to obtain appropriate merging results. However, we note that most merging approaches proposed so far share this limitation as well. Second, if the pair of maps to be merged includes a large resolution difference, the proposed approach may also fail to estimate the initial merging parameters. To avoid large resolution differences, the host robot in a multi-robot system can build its reference map at a medium resolution. Our future work will focus on addressing the second limitation.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant Nos. 61203326, u1261111, and 91320301. China Postdoctoral Science Foundation under Grant Nos. 2012M512004, the Fundamental Research Funds for the Central Universities under Grant Nos. xjj2012089.

References

1. R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *Int. J. Robot. Res.* **5**(4), 56–68 (1986).
2. S. Thrun, W. Burgard and D. *Fox, Probabilistic Robotics* (MIT Press, Cambridge, MA, USA, 2005).
3. S. Thrun and Y. Liu, "Multi-Robot SLAM with Sparse Extended Information Filers," *Robotics Research* (Springer-Verlag, Berlin, Germany, 2005) pp. 254–266.
4. S. Carpin, A. Birk and V. Jucikas, "On map merging," *Robot. Auton. Syst.* **53**(1), 1–14 (2005).
5. W. H. Huang and K. R. Beevers, "Topological map merging," *Int. J. Robot. Res.* **24**(8), 601–613 (2005).
6. A. Birk and S. Carpin, "Merging occupancy grid maps from multiple robots," *IEEE Proc.* **94**(7), 1384 (2006).
7. A. Howard, L. E. Parker and G. S. Sukhatme, "Experiments with a large heterogeneous mobile robot team: exploration, mapping, deployment and detection," *Int. J. Robot. Res.* **25**(5–6), 431–447 (2006).
8. D. Fox, J. Ko, K. Konolige, B. imketai, D. Schulz and B. Stewart, "Distributed multirobot exploration and mapping," *IEEE Proc.* **94**(7), 1325–1339 (2006).
9. S. Carpin, "Fast and accurate map merging for multi-robot systems," *Auton. Robot.* **25**(3), 305–316 (2008).
10. M.P. Lina, J.D. Tardos, J. Neria, "Divide and Conquer: EKF SLAM in O(n)," *IEEE Trans. Robot.* **24**(5), 1107–1120.
11. J. Zhu, S. Du, L. Ma, Z. Yuan and Q. Zhang, "Merging grid maps via point set registration," *Int. J. Robot. Autom*. **28**(2), 180–191 (2013).
12. H. D. Whyte and T. Bailey, "Simultaneous localization and mapping (SLAM): Part I. Essential algorithms," *IEEE Robot. Autom. Mag.* **13**, 99–108 (2006).
13. P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992).
14. D. Chetverikov, D. Stepanov and P. Krsek, "Robust Euclidean alignment of 3D point sets: the trimmed iterative closest point algorithm," *Image Vis. Comput.* **23**(3), 299–309 (2005).
15. J. M. Phillips, R. Liu and C. Tomasi, "Non-common Area Robust ICP for Minimizing Fractional RMSD," *Proceedings of International Conference on 3-D Digital Imaging and Modeling (3DIM-07)*, (2007) pp. 427–434.
16. S. Du, N. Zheng, S. Ying and J. Wei, "ICP with Bounded Scale for Registration of mD Point Sets," *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, (2007) pp. 1291–1294.
17. S. Ying, J. Peng, S. Du and H. Qiao, "A scale stretch method based on ICP for 3D data registration," *IEEE Trans. Autom. Sci. Eng.* **6**(3), 559–565 (2010).
18. J. H. Zhu, N. N. Zheng, Z. J. Yuan, S. Y. Du and L. Ma, "Robust scaling iterative closest point algorithm with the bidirectional distance measurement," *Electron. Lett.* **46**(24), 1604–1605 (2010).

19. J. Zhu, D. Meng, Z. Li, S. Du and Z. Yuan, "Robust registration of partially overlapping point sets via genetic algorithm with growth operator," *IET Image Process.* **8**(10), 582–590 (2014).
20. M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM* **24**(6), 381–395 (1981).
21. M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.* **74**(1), 59–73 (2007).
22. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.* **60**(2), 91–110 (2004).
23. A. Censi, L. Iocchi and G. Grisetti, "Scan Matching in the Hough Domain," *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2 (2005) pp. 2739–2744.
24. J. L. Blanco, J. González-Jiménez and J. Fernández-Madrigal, "A robust, multi-hypothesis approach to matching occupancy grid maps," *Robotica* **31**(5), 687–701 (2013).
25. K.M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss and W. Burgard, "OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems," *Auton Robot.* **34**(3), 189–206 (2013).
26. K. F. Mulchrone, "Application of Delaunay triangulation to the nearest neighbor method of strain analysis," *J. Struct. Geol.* **5**(5), 689–702 (2003).
27. M. Greenspan and M. Yurick, "Approximate k-d Tree Search for Efficient ICP," *Proceedings of International Conference on 3-D Digital Imaging and Modeling (3DIM-03)*, (2003) pp. 442–448.
28. A. Nuchter, K. Lingemann and J. Hertzberg, "Cached k-d Tree Search for ICP algorithms," *Proceedings of International Conference on 3-D Digital Imaging and Modeling (3DIM-07)*, (2007) pp. 419–426.
29. Y. Hwang, B. Han and H. K. Ahn, "A Fast Nearest Neighbor Search Algorithm by Nonlinear Embedding," *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, (2012) pp. 3053–3060.
30. A. Nuchter, J. Elseberg, P. Schneider and D. Paulus, "Study of parameterizations for the rigid body transformations of the scan registration problem," *Comput. Vis. Image Und.* **114**(8), 358–367 (2010).
31. C. Stachniss, "Robotics Datasets," [Online]. Available at: http://www.informatik.uni-freiburg.de/ stachnis/datasets.html.
32. A. Eliazar and R. Parr, "DP-SLAM," [Online]. Available at: http://www.cs.duke.edu/ parr/dpslam/.
33. J. Bibby, "Axiomatisations of the average and a further generalisation of monotonic sequences," *J. Glasgow Math.* **15**, 63–65 (1974).
34. H. C. Chae and K.Hwajoon, "The validity checking on the exchange of integral and limit in the solving process of PDEs," *Int. J. Math. Anal.* **8**(22), 1089–1092 (2014).
35. S. Thrun, S. Thayer et al., "Autonomous exploration and mapping of abandoned mines," *IEEE Robot. Autom. Mag.* **11**(4), 79–91 (2003).

Appendix: Convergence analysis

The convergence of TsICP is similar to that of original ICP. The following theorem explains this convergence in detail to prove that the TsICP algorithm converges in theory.

Theorem 1. The TsICP algorithm converges monotonically to a local minimum with respect to the MSE.

The following proof shows that, for each step of the TsICP algorithm, the trimmed MSE can be no worse than the previous step.

Proof. Given two grid maps $P \triangleq \{\vec{p}_i\}_{i=1}^{N_p}$ and $Q \triangleq \{\vec{q}_j\}_{j=1}^{N_q}$, we use $T_k = (s_k, \mathbf{R}_k, \vec{t}_k)$, ξ_k, and Q_{ξ_k} to denote the scaling transformation, overlap percentage, and overlapping part of Q, respectively. In the first step of the kth iteration, the *k-d* tree method is used to search for the nearest neighbour $\vec{p}_{c_k(j)}$ of the edge point $\vec{q}_j$. Let $\vec{p}_{c_k(j),k-1} = s_{k-1}\mathbf{R}_{k-1}\vec{p}_{c_k(j)} + \vec{t}_{k-1}$. The trimmed MSE is then defined as:

$$e_k = \frac{1}{\left|Q_{\xi_{k-1}}\right|(\xi_{k-1})^{1+\lambda}} \sum_{\vec{q}_j \in Q_{\xi_{k-1}}} \left\|\vec{p}_{c_k(j),k-1} - \vec{q}_j\right\|_2^2. \tag{A.1}$$

In the second step of the kth iteration, percentage ξ_k and corresponding subset Q_{ξ_k} are updated. The updated MSE is

$$\eta_k = \frac{1}{\left|Q_{\xi_k}\right|(\xi_k)^{1+\lambda}} \sum_{\vec{q}_j \in Q_{\xi_k}} \left\|\vec{p}_{c_k(j),k-1} - \vec{q}_j\right\|_2^2. \tag{A.2}$$

Because ξ_k and Q_{ξ_k} are updated by minimizing the trimmed MSE, it is reasonable that $\eta_k \leq e_k$. Otherwise, let $\xi_k = \xi_{k-1}$ and $Q_{\xi_k} = Q_{\xi_{k-1}}$. It is obvious that $\eta_k = e_k$. In the third step of the kth iteration, $\{\vec{q}_j\}_{j=1}^{Q_{\xi_k}}$ is registered with $\{\vec{p}_{c_k(j),k-1}\}_{j=1}^{Q_{\xi_k}}$ and the scaling transformation $\mathrm{T}_k = (s_k, \mathbf{R}_k, \vec{t}_k)$ is

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press

optimized. Let $\vec{p}_{c_k(j),k} = s_k \mathbf{R}_k \vec{p}_{c_k(j)} + \vec{t}_k$. The trimmed MSE then becomes

$$\varepsilon_k = \frac{1}{\left|Q_{\xi_k}\right| (\xi_k)^{1+\lambda}} \sum_{\vec{q}_j \in P_{\xi_k}} \left\| \vec{p}_{c_k(j),k} - \vec{q}_j \right\|_2^2. \tag{A.3}$$

Because transformation $\mathrm{T}_k = (s_k, \mathbf{R}_k, \vec{t}_k)$ is computed from Eq. (12), it is easy to determine that $\varepsilon_k \leq \eta_k$. Otherwise, let $s_k = s_{k-1}$, $\mathbf{R}_k = \mathbf{R}_{k-1}$, and $\vec{t}_k = \vec{t}_{k-1}$. It is obvious that $\varepsilon_k = \eta_k$.

In the $(k+1)$th iteration, the new nearest neighbour $\vec{p}_{c_{k+1}(j)}$ of $\vec{q}_j$ can be searched for again. Let $\vec{p}_{c_{k+1}(j),k} = s_k \mathbf{R}_k \vec{p}_{c_{k+1}(j)} + \vec{t}_k$. The new trimmed MSE can then be defined as:

$$e_{k+1} = \frac{1}{\left|Q_{\xi_k}\right| (\xi_k)^{1+\lambda}} \sum_{\vec{q}_j \in P_{\xi_k}} \left\| \vec{p}_{c_{k+1}(j),k} - \vec{q}_j \right\|_2^2. \tag{A.4}$$

Because e_{k+1} is the result of Eq. (9), then

$$e_{k+1} = \frac{1}{\left|Q_{\xi_k}\right| (\xi_k)^{1+\lambda}} \sum_{\vec{q}_j \in P_{\xi_k}} \left\| \vec{p}_{c_{k+1}(j),k} - \vec{q}_j \right\|_2^2 \leq \frac{1}{\left|Q_{\xi_k}\right| (\xi_k)^{1+\lambda}} \sum_{\vec{q}_j \in P_{\xi_k}} \left\| \vec{p}_{c_k(j),k} - \vec{q}_j \right\|_2^2 = \varepsilon_k. \tag{A.5}$$

Hence, repeating the above procedures, the following results are obtained: $0 \leq \ldots \leq e_{k+1} \leq \varepsilon_k \leq \eta_k \leq e_k \leq \ldots$ for all k.

According to the Monotonic Sequence Theorem,[33,34] "Every bounded monotonic sequence of real numbers is convergent." Hence, the TsICP algorithm always converges monotonically to a local minimum with respect to the trimmed MSE.

https://doi.org/10.1017/S0263574715000168 Published online by Cambridge University Press