A multilevel fusion system for multirobot 3-D mapping using heterogeneous sensors

Yue, Yuefeng; Yang, Chule; Wang, Yuanzhe; Senarathne, P. G. C. N.; Zhang, Jun; Wen, Mingxing; Wang, Danwei

2020

Yue, Y., Yang, C., Wang, Y., Senarathne, P. G. C. N., Zhang, J., Wen, M. & Wang, D. (2020). A multilevel fusion system for multirobot 3-D mapping using heterogeneous sensors. IEEE Systems Journal, 14(1), 1341-1352. <https://dx.doi.org/10.1109/JSYST.2019.2927042>

<https://hdl.handle.net/10356/147245>

<https://doi.org/10.1109/JSYST.2019.2927042>

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: <https://doi.org/10.1109/JSYST.2019.2927042>

Downloaded on 26 Apr 2023 21:02:10 SGT

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/334623572>

A Multilevel Fusion System for Multirobot 3-D Mapping Using Heterogeneous Sensors

Article in IEEE Systems Journal · July 2019

DOI: 10.1109/JSYST.2019.2927042

CITATIONS

17

READS

346

7 authors, including:

Yufeng Yue

Beijing Institute of Technology

35 PUBLICATIONS 155 CITATIONS

[SEE PROFILE](#)

Chule Yang

Nanyang Technological University

22 PUBLICATIONS 112 CITATIONS

[SEE PROFILE](#)

Yuanzhe Wang

Nanyang Technological University

20 PUBLICATIONS 70 CITATIONS

[SEE PROFILE](#)

Jun Zhang

Nanyang Technological University

14 PUBLICATIONS 88 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:

Intelligent Transportation Control System [View project](#)

Unique Person Recognition [View project](#)

A Multi-level Fusion System for Multi-Robot 3D Mapping using Heterogeneous Sensors

Yufeng Yue, Chule Yang, Yuanzhe Wang, P.G.C.N. Senarathne, Jun Zhang, Mingxing Wen, and Danwei Wang

Abstract—Operating multiple robots in an unstructured environment is challenging due to its high complexity and uncertainty. In such applications, the integration of individual maps generated by heterogeneous sensors is a critical problem, especially the fusion of sparse and dense maps. This paper proposes a general multi-level probabilistic framework to address the integrated map fusion problem, which is independent of sensor type and SLAM algorithm employed. The key novelty of this work is the mathematical formulation of the overall map fusion problem and the derivation of its probabilistic decomposition. The framework provides a theoretical basis for computing the relative transformations amongst robots and merging probabilistic map information. Since the maps generated by heterogeneous sensors have different physical properties, an Expectation-maximization (EM) based map matching algorithm is proposed which automatically determines the number of voxels to be associated. Then, a time-sequential map merging strategy is developed to generate a globally consistent map. The proposed approach is evaluated in various environments with heterogeneous sensors, which demonstrates its accuracy and versatility in 3D multi-robot map fusion missions.

Index Terms—Collaborative Mapping, Probability Theory, Information Fusion, Multi-Robot Systems.

I. INTRODUCTION

Cooperative operation of multi-robot systems has attracted significant attention in the past two decades due to its great potential in practical applications, such as large-scale mapping [1], intelligent decision-making [2], and rapid response [3]. To provide a comprehensive description of the whole environment, it is necessary for the robots to share and fuse their individual map information with each other [4]. For example, this can be the case of path planning in unknown environments [5], or an area coverage problem with heterogeneous sensor networks [6]. As the state of the art of single robot map construction and autonomous exploration algorithms are becoming more efficient and robust, the next challenge is to extend these techniques to multi-robot systems equipped with miscellaneous sensors.

As different sensors have different advantages and limitations, each robot can carry heterogeneous sensors, e.g. laser scanners, cameras, etc. Even though 3D Lidars have a 360° field-of-view, the quality of the map they generated are not satisfactory due to the sparseness of its physical property.

*The research was partially supported by the ST Engineering NTU Corporate Lab through the NRF corporate lab@university scheme. (Corresponding author: Yuanzhe Wang)

Yufeng Yue, Chule Yang, Yuanzhe Wang, Jun Zhang, Mingxing Wen and Danwei Wang are with School of Electrical and Electronic Engineering, Nanyang Technological University(NTU), Singapore ({yyue001, yang0438, wang0951, jzhang061, mingxing001}@e.ntu.edu.sg, edwwang@ntu.edu.sg).

P.G.C.N. Senarathne was with ST Engineering-NTU Corporate Lab, Nanyang Technological University at the time of research (namal.senarathne@gmail.com).

RGB-D camera generates a denser and more detailed 3D map. Nevertheless, it suffers from limited field-of-view and the observation error grows exponentially as the sensing range increases. Therefore, using homogeneous sensors in multi-robot systems is far from enough in terms of mapping accuracy and robustness [7]. However, employing heterogeneous sensors on multiple robots can make full use of their advantages and provide a more accurate description of the environment.

For a group of robots in arbitrary environments, e.g. structured, semi-structured or unstructured environments, the fundamental challenge is to acquire a comprehensive understanding of the environment [8]. Due to limited sensing capabilities, each robot only has partial information of the surroundings. Therefore, the integration of multiple sensor knowledge together can help to obtain the overall information of the environment, which will significantly improve the perception performance of the robot. In order to ensure that each robot can respond rapidly from a holistic viewpoint, it is essential to share and fuse local 3D maps generated by heterogeneous sensors to produce a globally consistent map such that a comprehensive understanding of the whole environment can be obtained.

The motivation for this work is that the research of exploiting multi-robot map fusion framework using heterogeneous sensors has not been studied extensively. In general, the map fusion problem is composed of the estimation of relative transformations and the integration of partial map information, which is usually tackled separately [9]. Existing approaches usually employ point set registration algorithms to achieve map matching [4], [10], where the hard one-to-one correspondence relationship is assigned. Since the intrinsic characteristics (density/range/field-of-view) of the map varies from different sensors, accurate estimation of the relative transformation is a challenging problem. In addition, most researchers focus on estimating relative transformation and directly applying the estimated transformation to stitch the partial maps [11], which ignores the propagation and integration of the transformation and map uncertainties. Distributed settings should also be considered to increase the robustness of the system under periodic communication, narrow bandwidth and limited computational power.

This paper proposes a novel multi-level framework that allows joint local and global map matching and merging. This will provide a theoretical basis for multi-robot mapping missions. The main contributions of this paper are listed as follows:

- A general probabilistic framework is proposed to address the integrated map fusion problem, which is independent of sensor type and SLAM algorithm. The framework jointly estimates the posterior of computing the relative

transformations (map matching) and merging probabilistic map information (map merging).

- A probabilistic map matching algorithm is proposed which automatically decides the number of voxels to be associated for maps produced by heterogeneous sensors.
- A time-sequential map merging algorithm is designed which addresses the distributed multi-robot map fusion problem. In particular, uncertainties embedded in individual maps and caused by the relative transformation are propagated and integrated.

The rest of the paper is organized as follows. Section II reviews relative work on collaborative mapping. Section III gives an overview of the system framework. Section IV outlines the multi-robot probabilistic map matching algorithm. Section V introduces the time-sequential map merging algorithm. Section VI presents the results of the experiments. Section VII concludes the paper and discusses several possible future directions.

II. RELATED WORK

In the past few decades, researchers have proposed various effective solutions to address multi-robot mapping problems. In [2] [12], comprehensive reviews of different solutions for multi-robot mapping are summarized. These techniques range from front-end 2D topology extraction [13] [9], sparse feature descriptors [14], full 3D dense map registration [15] [16], to back-end optimization of splitting maps into sub-maps [17], [18] and spatial constraints graph structures [19] [11]. Here, our work focuses on the front-end and makes the assumption that the local map generated by each robot is consistent.

A. Heterogeneous Map matching

The difficulty of multi-robot matching task largely depends on estimating the relative positions amongst the robots. Most of the existing work on multi-robot map fusion only addresses the map matching part, where relative transformations between partial maps are estimated with various registration algorithms. Depending on the types of the generated maps, different approaches for map matching can be found in the literature. For 2D grid maps, the topology of Delaunay triangulation is extracted to match feature maps [13]. Similarly, topology-based approaches like Voronoi diagram [9] is applied. However, above-mentioned methods are not adequate for sophisticated 3D environments and result in loss of information.

As the availability of high-quality 3D sensors increases, 3D registration algorithms have been widely applied, such as the classical Iterative Closest Point (ICP) [20], its variant Generalized Iterative Closest Point (GICP) [21] and Normal Distribution Transform (NDT) [22]. To apply dense registration algorithms in map matching, probabilistic map information is incorporated with the geometric map information to improve the accuracy of the estimated transformation in [4]. However, the algorithm assigns hard one-to-one correspondences.

Inspired by the probabilistic characteristics of expectation maximization (EM), several work considers a more generalized scenario of correspondence between point sets with noise. Ref [23] proposes an EM-based ICP to improve the registration

precision by considering the Gaussian noise between the point sets. In [24], the author further exploits the data association relationship and proposes an algorithm to align a sparse point cloud with a dense one, where each point is associated with a set of points. Then, EM-based data association is extended to include semantic information and features, as well as internal measurements in [25]. Our map matching algorithm is similar to [24], however, the paper only hard codes fixed number of nearest neighborhood. In particular, it does not discuss how to automatically tune the multiple data association strategy to balance the accuracy and efficiency. In addition, the problem is not fully considered in a multi-robot mapping framework, where additional problems will be introduced. A structural descriptor to match dense point-cloud maps and vision keypoint maps is proposed in [14]. However, the 3D laser usually tends to generate sparse data and may not provide enough information to produce the descriptors.

B. Probabilistic Information Fusion

To distribute data among robots, a serial sequential strategy is proposed to communicate with robots in [12]. In DDF-SAM [26], a robust distributed communication framework is proposed, which only provides simulations results under simplified assumptions. Ref [27] proposes a distributed scheme for multi-robot data-gathering under limited communication and buffer size.

Once the map is received and relative transformation is calculated, the source map is transformed to the coordination frame of the target map by a transformation function. In order to merge the partial maps to a globally consistent map, the data perceived using individual robots has to be integrated into a global map. Most approaches directly apply this transformation to stitch the partial maps and ignore the uncertainty propagation and merging.

In general, there are two types of uncertainties that need to be integrated when performing map merging, which are uncertainty embedded in the map and uncertainty caused by the relative transformation. To propagate the transformation uncertainty, the covariance of transformation function is derived in [17]. Ref [28] proposes a closed-form estimate of ICP's covariance. Then, the covariance of the transformed point is calculated by deriving the Jacobian of the transformation function in [9]. Map uncertainty can be represented by the occupancy probability in [29]. In [30], two 2D Gaussian distributions are multiplied to a compounded Gaussian distribution. Alternatively, the maps are fused on the voxel-wise level by directly taking the average of the probabilities in [15]. In [4], the relative transformation is evaluated based on a probabilistic measurement model and map dissimilarities are combined with relative entropy filter.

Aforementioned approaches either assume accurate transformation and only consider merging map uncertainty [15] [4] [30], or ignore the map uncertainty and only focus on transformation uncertainty [9] [17]. In addition, existing approaches do not couple the map matching output with map merging, where two sub-problems are solved separately.

Fig. 1. The diagram of the multi-robot mapping system.

III. DISTRIBUTED MULTI-ROBOT MAP FUSION

In this section, the system architecture of multi-robot map fusion with single robot level and multi-robot level is presented, where the problem is formulated to provide a theoretical basis for the factorization into two subproblems, namely map matching and map merging.

A. System Architecture

The overview of the system architecture is depicted in Fig. 1. The focus of this section is to build a general framework for map fusion, which can be divided into two sub-problems, i.e., map matching and map merging. **First level** is perception level, the robot equipped with heterogeneous sensors will generate its map. For example, the laser scanner tends to generate a 360 degrees sparse map, while the RGB-D camera generates a denser map with the narrow field-of-view. **Second level** is the local map level, the maps produced by different sensors are fused locally on each robot to combine the strength of each sensor. **Third level** is global map level, the robots communicate with each other to transmit and fuse the local maps into a global map. Our main focus is on the second level and third level.

B. System Framework Definition and Formulation

Problem 1 Given a group of robots $\mathcal{R} \triangleq \{r\}_{1:n_r}$, each robot r is equipped with a set of heterogeneous sensors $\mathcal{S} \triangleq \{s\}_{1:n_s}$, the map produced by robot r using sensor s is denoted as $m_{r,s}$, the problem is to estimate the fused global map M_r and the set of relative transformation $\mathcal{T}_r$ with a fully distributed network.

Let $\mathcal{R} = \{r_1, r_2, \dots, r, \dots\}$ (the number of robots is n_r) be the set of robots in an environment. The neighboring robots of robot r is defined as $\mathcal{R}_r = \{r_n, r_n \in \mathcal{R}_r\}$, where neighboring robots $\mathcal{R}_r$ is the set of robots in communication with r and $\mathcal{R} = \{r \cup \mathcal{R}_r\}$ if the communication range covers all robots (see Fig. 2).

In the system, the input is the perception level and the output is the global map level. Local map level is an intermediate

Fig. 2. Top: Robot platform and the fused 3D map created using the two robots. Bottom: The diagram of the group of robots.

level that acts as communication node and single robot fusion master. The benefits of introducing the local map level come at following advantages, firstly: the size of the fused local map is smaller than the summation of individual sensor maps, which reduces the communication burden; secondly, the fused map on single robot level helps the localization and exploration of a single robot, while the maps produced by different sensors are fused to compensate for the limitation of each type of sensor; thirdly, performing map fusion at the perception level brings the computational complexity of $\mathcal{O}(R^2 S^2)$, while the local map level decreases to $\mathcal{O}(R^2 + R S^2)$. Given the maps of perception level $m_{r,s}$ and estimating fused global map $M_{\mathcal{R}_r}$ conditional on available local maps M_r , the multi-level framework is factorized into estimating local fused map M_r .

$$\underbrace{\prod_{r \in \mathcal{R}} p(m_r, T_{S_r} | m_{r,1:n_s})}_{\text{local map level}} \cdot \underbrace{p(M_r, T_{r,\mathcal{R}_r} | m_r, m_{\mathcal{R}_r})}_{\text{global map level}} \quad (1)$$

For local map level, each individual robot r is assumed to estimate its trajectory based on local observations and generate its partial map m_r^s , then robot r fuses the sensor maps $m_r^{1:n_s}$ to generate M_r . For global map level, robot r receives the set of local maps $M_{\mathcal{R}_r}$ from nearby robots, where $M_{\mathcal{R}_r} = \{M_{r_n}, r_n \in \mathcal{R}_r\}$.

C. Map Fusion Definition and Formulation

Problem 2 For both local and global robot levels, map fusion problem is a joint estimation of relative transformation and merged map. Assuming the relative transformation estimation and map merging are conditionally independent, the map fusion problem is factorized into estimating a set of relative transformation (map matching) and generating the global map conditioned on all available partial maps (map merging).

To approximate the joint distribution, both local and global map fusion in Eq. (1) can be factorized into a product of relative transformation posterior and the fused map posterior. To simplify the derivation, only global map level is derived in detail, where it is the same for the local level. Then, Eq. (1) is written as follows:

$$p(M_{\mathcal{R}_r}^{\mathcal{R}_r}, T_{r,\mathcal{R}_r} | m_r, m_{\mathcal{R}_r}) = \underbrace{p(T_{r,\mathcal{R}_r} | m_r, m_{\mathcal{R}_r})}_{\text{map matching}} \cdot \underbrace{p(M_{\mathcal{R}_r}^{\mathcal{R}_r} | T_{r,\mathcal{R}_r}, m_r, m_{\mathcal{R}_r})}_{\text{map merging}} \quad (2)$$

This decouples multi-robot map fusion problem into two separate estimation problems. The key advantages of this factorization is summarized as follows: firstly, it benefits from using MAP estimation for robots relative transformation estimation, making it possible to apply probabilistic nonlinear least square minimization models. Secondly, it benefits from utilizing a time-sequential relative entropy filter based map merging algorithm. Thirdly, the output of the map matching result can be used to propagate the uncertainty for map merging, which tightly concatenates the two sub-problems.

IV. MULTI-ROBOT MAP MATCHING

In this section, we present a general framework for map matching. To address the challenge of matching maps produced by heterogeneous sensors, the estimation of relative transformation based on expectation maximization is proposed.

A. Mathematic Formulation

For robot r , it receives partial maps sequentially, which means the relative transformation is a serial processing upon receiving of new partial maps. Since the relative transformation T_{r,r_n} is also pairwise independent of each pairs of robots, the posterior of relative transformation $p(T_{r,\mathcal{R}_r} | m_r, m_{\mathcal{R}_r})$

can be factorized into a product of the posterior of relative transformation between each pair, as shown in Eq. (3):

$$p(T_{r,\mathcal{R}_r} | m_r, m_{\mathcal{R}_r}) = \prod_{r_n \in \mathcal{R}_r} p(T_{r,r_n} | m_r, m_{r_n}) \quad (3)$$

where $T_{r,r_n} \in SE(3)$ is a rigid transformation with rotation $R_{r,r_n} \in SO(3)$ and translation $t_{r,r_n} \in \mathcal{R}^3$. Instead of simply considering T_{r,r_n} as a fixed unknown parameter, we model it as a 6D random variable. Then, T_{r,r_n} can be modeled as a multivariate Gaussian distribution and solved by MAP estimation, where $T_{r,r_n} \sim (\hat{T}_{MAP}, \hat{\Sigma}_{MAP})$. To solve the MAP estimation problem, we apply Bayes rule to factorize Eq. (3) into Eq. (4), which is the product of maximum likelihood estimation and prior estimation:

$$p(T_{r,r_n} | m_r, m_{r_n}) = p(m_r | T_{r,r_n}, m_{r_n}) \cdot p(T_{r,r_n}), \quad (4)$$

where term $p(m_r | T_{r,r_n}, m_{r_n})$ is a common maximum likelihood estimation (MLE) problem that aims to find the most likely relative transformation T_{r,r_n} by matching the two partial maps m_r and m_{r_n} . By specifying the initialization, we can incorporate prior information to evaluate the posterior even with a poor initial input, which can be given by feature descriptor correlation. If no prior information is available, the prior distribution is assumed to be a uniform distribution. In this case, we have $\hat{T}_{MAP} = \hat{T}_{MLE}$, where MLE is a special case of MAP estimation. To establish voxel-wise correspondences for map merging, the dense scan-matching technique is applied to solve MLE.

B. 3D Occupancy Map Matching

For dense registration of two maps m_r and m_{r_n} , the common solution is to fix the map of current robot m_r as model map, and the neighboring map m_{r_n} as scene map, where the rigid transformation T_{r,r_n} is estimated by transforming m_{r_n} to the coordinate frame of m_r and minimizing the error metric iteratively. To estimate T_{r,r_n} , data association needs to be estimated firstly, which is usually addressed by performing the nearest neighbor search of Euclidean distance or Mahalanobis distance. For pure registration purpose, the matches are not parameters of interest and are usually ignored after registration.

In our previous work [4], we apply a novel map registration algorithm called occupancy iterative closest point(OICP), which establishes voxel-to-voxel correspondences between all overlapped partial maps. Here, we further exploit the data association issue with different cases. For the partial map, we define $m_r = \{m_r^i\}_{i=1}^{N_{m_r}}$, $m_{r_n} = \{m_{r_n}^j\}_{j=1}^{N_{m_{r_n}}}$, where N_{m_r} and $N_{m_{r_n}}$ are the number of voxels of partial maps. Since the voxel-wise correspondences are usually assumed to be independent, the probability density function of $p(m_r | m_{r_n}, T_{r,r_n})$ can be factorized into Eq. (5).

$$p(m_r | m_{r_n}, T_{r,r_n}) = \prod_i p(m_r^i | m_{r_n}^{c_i}, T_{r,r_n}) \quad (5)$$

where $m_{r_n}^{c_i} = \{m_{r_n}^j, j \in c_i\}$ is the set of candidate corresponding voxels of m_r^i . The problem of calculating the corresponding points $m_{r_n}^{c_i}$ of m_r^i is often referred as data association. To estimate the hidden data association, we introduce EM, which is

(a) An example of heterogeneous map fusion

(b) Density of RGB-D camera

(c) Density of laser scanner robot

(d) Density histogram of RGB-D camera

(e) Density histogram of laser scanner robot

Fig. 3. The demonstration of heterogeneous map fusion, where the density of heterogeneous maps are compared.

an algorithm well-suited for models containing latent variables and estimating them iteratively by marginalizing the likelihood function.

There are two steps for EM: the E-step efficiently estimates the hidden variables by evaluating the expectation, while the M-step solves the rigid transformation based on E-step using optimization algorithms. Here, we define a binary matching matrix J to represent the estimation of hidden data association between the corresponding voxels, where $J_{i,j} = 1$ if m_r^i corresponds to $m_{r_n}^j$ and equals 0 otherwise. For each row i , the matched voxels of m_r^i is the set of $J_{i,c_i} = \{J_{i,j} = 1, j \in c_i\}$. The EM form of Eq. (5) is rewritten as:

$$\begin{aligned}
 p(m_r | m_{r_n}, T_{r,r_n}) &= \sum_j p(m_r, J | m_{r_n}, T_{r,r_n}) \\
 &= \prod_i \sum_{c_i} p(m_r^i, J_{i,c_i} | m_{r_n}^{c_i}, T_{r,r_n}) \\
 &= \prod_i \sum_{c_i} \underbrace{p(J_{i,c_i} | m_r^i, m_{r_n}^{c_i}, T_{r,r_n})}_{\text{E step: data association}} \cdot \underbrace{p(m_r^i, J_{i,c_i} | T_{r,r_n}, m_{r_n}^{c_i})}_{\text{M step: optimization}} \quad (6)
 \end{aligned}$$

The negative log-likelihood to be minimized with EM is given by

$$f(T_{r,r_n}) = \min_{T_{r,r_n}} \sum_i \sum_{c_i} p(J_{i,c_i} | m_r^i, m_{r_n}^{c_i}, T_{r,r_n}) \cdot \ln p(m_r^i, J_{i,c_i} | T_{r,r_n}, m_{r_n}^{c_i}) \quad (7)$$

C. E-step

- **Homogeneous Sensors:** Since the density of map is similar for the same type of sensors, one-to-one hard assignment is established for computational efficiency. For each voxel m_r^i , there will be one and only one voxel $m_{r_n}^j$ correspond to it, where $J_{i,c_i} = J_{i,j}$.

$$\forall m_r^i \in m_r, \exists! m_{r_n}^j \in m_{r_n} \Rightarrow J_{i,j} = \omega_{ij}$$

- **Heterogeneous Sensors:** If heterogeneous sensors need to be aligned, one-to-one matching might fail due to wrong matches or insufficient correspondences. Hence, multiple soft correspondences should be established, where each voxel m_r^i corresponds to multiple voxels in $m_{r_n}^{c_i}$

$$\forall m_r^i \in m_r, \exists m_{r_n}^{c_i} \in m_{r_n} \Rightarrow J_{i,c_i} = \{J_{i,j} = \omega_{ij}, j \in c_i\}$$

The problem comes at two hands: (1) automatically determining the policy for the candidate correspondences searching. If it is the case of heterogeneous sensors, how to determine the value of c_i to balance the efficiency and accuracy. (2) For heterogeneous sensors, the candidate $J_{i,j} \in J_{i,c_i}$ should have different weights to contribute to the objective function.

As can be seen in Fig. 3, the density of RGB-D camera is higher than laser scanner. In addition, the histogram of the RGB-D map complies a Gaussian distribution, which means the density is distributed more evenly. However, the histogram of the laser is more like the Gamma distribution. Based on the observation, we propose a local density map based automatic tuning strategy for probabilistic data association, as shown in Algorithm 1. For the given maps, we firstly generate their density maps (depicted in Fig. 3). If the means of the density maps are similar, a one-to-one correspondence is applied. If not, the numbers of correspondences (c_i) to be established are determined by the ratio of local density value. As shown in Fig. 4, the density of RGB-D map is distributed evenly, the numbers of voxels matched increases with the decreasing of laser map's density. Each voxel in the laser map matches different value to c_i in RGB-D map, where the histogram and density map shows the automatically tuned value of c_i .

D. M-step

Solving M-step is to minimize the log-likelihood of Eq. (6) based on the data association established in E-step. Here, probabilistic weights are assigned to each pair of correspondences. There are several different ways of updating the rotation and translation, for example, Levenberg-Marquardt, single value decomposition.

$$f(T_{r,r_n}) = \min_{T_{r,r_n}} \sum_i \sum_{c_i} \omega_{i,j} \cdot \ln p(m_r^i, J_{i,j} | T_{r,r_n}, m_{r_n}^j) \quad (8)$$

V. TIME-SEQUENTIAL MAP MERGING

After the relative transformation has been derived, the partial map m_{r_n} is transformed to the coordinate frame of m_r by transformation function $\Phi(T_{r,r_n}, m_{r_n})$, to generate $r_{r_n} = \Phi(T_{r,r_n}, m_{r_n})$. The process of combing the information from

(a) Numbers of voxels matched in RGB-D map for each voxel in laser matched map (b) Histogram of numbers of voxels matched in RGB-D map for each voxel in laser matched map

(c) The edges show the multiple data association established. Red points represent laser map, while black points are RGB-D map.

Fig. 4. Density based multiple data association strategy.

partial maps to form an enhanced global map is usually referred as map merging and is formulated as Eq. (9):

$$p(M_r^{\mathcal{R}_r} | T_{r,\mathcal{R}_r}, m_r, m_{\mathcal{R}_r}) = p(M_r^{\mathcal{R}_r} | m_r, \Phi_{r_n \in \mathcal{R}_r}(T_{r,r_n}, m_{r_n})) \quad (9)$$

Algorithm 1 Probabilistic Map Matching

Require: The maps $m_r = \{m_r^i\}_{i=1}^{N_{m_r}}$, $m_{r_n} = \{m_{r_n}^j\}_{j=1}^{N_{m_{r_n}}}$, and initial transformation function T_0

Ensure: Iteratively minimize the error function

1. generate density map d_r for m_r and d_{r_n} for m_{r_n}

for k from 1 to termination criteria **do**

for i from 1 to N_{m_r} criteria **do**

if $\text{round}(\frac{\text{mean}(d_r)}{\text{mean}(d_{r_n})}) = 1$ **then**

2. one-to-one matching, where $J_{i,c_i} = J_{i,j}$

else

3. multiple data association, where $c_i = \text{round}(\frac{d_r^i}{d_{r_n}^j})$

end if

4. assign weight: $\omega_{ij} = \exp(-(T^k * m_r^i - m_{r_n}^j)^2)$, $j \in c_i$

5. iteratively optimize the error function: $\hat{T}_{r,r_n}^{k+1} = \text{argmin} \sum_i \sum_j \omega_{ij} \cdot \ln p(m_r^i, J_{i,j} | T_{r,r_n}^k, m_{r_n}^j)$

end for

end for

During real missions, maps are generated and transmitted sequentially. This results in robot r receiving its neighboring robots' maps in some permutation π of $\mathcal{R}_r$. Hence, the map merging can be performed serially, where reception of a new partial map would trigger a pair-wise map merging. Then, Eq. (9) can be factorized into Eq. (10).

$$p(M_r^{\mathcal{R}_r} | m_r, \Phi_{r_n \in \mathcal{R}_r}(T_{r,r_n}, m_{r_n})) = \prod_{r_n=1}^{N_r} p(M_r^{\pi(1,r_n)} | M_r^{\pi(1,r_n-1)}, \Phi(T_{r,\pi(r_n)}, m_{\pi(r_n)})) \quad (10)$$

Note that $M_r^{\pi(1,r_n)}$ is the partial global map of robot r generated by fusing the maps of robots $\pi(1)$ to $\pi(r_n) \in \mathcal{R}_r$ with m_r after the reception of map $m_{\pi(r_n)}$ from neighboring robot $\pi(r_n) \in \mathcal{R}_r$. The final global fused map $M_r^{\mathcal{R}_r}$ of robot r , can be retrieved after the end of this serial process, hence $M_r^{\mathcal{R}_r} = M_r^{\pi(1,N_r)}$.

A. Uncertainty Propagation and Transformation

In general, there are two types of uncertainties need to be integrated when performing map merging.

1) *Map uncertainty*: in real environments, uncertainties will occur due to localization error and sensor measurement noise. In map level, these uncertainties are represented with the occupancy probability value. Since the same object can be observed in different perspectives by different robots, voxels representing this object could have different occupancy probabilities in separated maps. Hence, it is vital to consider these occupancy probability dissimilarities when fusing them into a global map.

2) *Relative transformation uncertainty*: This uncertainty is related to the relative transformation estimation results. The more accurate the estimation, the smaller the uncertainty. This type of uncertainty also needs to be considered in map merging process. The relative estimation uncertainty is propagated by the map matching algorithm and will be represented as a covariance matrix.

Our approach is to propagate the uncertainty for both single robot and multi-robot levels, then the corresponding probability distributions are merged to output an improved estimate of the observed object. The assumption is based on that sensor errors are independent and complies a normal distribution. Assume that $q = [q_x \ q_y \ q_z \ L_q]$ denotes the position of the extracted voxel center $[q_x \ q_y \ q_z]$ and the occupancy probability L_q . The mean of q is $\mu_q = [\mu_{q_x} \ \mu_{q_y} \ \mu_{q_z}]$, the covariance is a diagonal matrix with the same entry $\Sigma_q = \text{diag}[\sigma_{q_x q_x}^2 \ \sigma_{q_y q_y}^2 \ \sigma_{q_z q_z}^2]$. Voxel q is transformed by the following function to voxel r .

$$r = f(\alpha, \beta, \gamma, t_x, t_y, t_z, q_x, q_y, q_z) = Rq + t \quad (11)$$

where α, β, γ are the rotation angles and t_x, t_y, t_z are the translation elements. Since the transformation function $f(\cdot)$ is nonlinear, then the transformed voxel r will be non-Gaussian. To approximate r with Gaussian, the transformation $f(\cdot)$ is linearized using the first-order Taylor expansion:

$$p(r) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp \left\{ -\frac{1}{2}(r - \mu_r)^T \Sigma_r^{-1} (r - \mu_r) \right\} \sim \mathcal{N}(r; \mu_r, \Sigma_r) \quad (12)$$

,where μ_r can be directly calculated using Eq. (11). Assume that the covariance of transformation is $\Sigma_f =$

$diag[\sigma_{tx}^2, \sigma_{ty}^2, \sigma_{tz}^2, \sigma_{\alpha\alpha}^2, \sigma_{\beta\beta}^2, \sigma_{\gamma\gamma}^2]$. Through deriving the Jacobian of transformation function, the covariance of transformed voxel Σ_r is estimated in Eq. (13).

$$F_{\alpha\beta\gamma} = \frac{\partial f(\cdot)}{\partial(\alpha, \beta, \gamma, t_x, t_y, t_z)}, F_{q_x q_y q_z} = \frac{\partial f(\cdot)}{\partial(q_x, q_y, q_z)}$$

$$\Sigma_r = \underbrace{F_{\alpha\beta\gamma} \Sigma_{t_x t_y t_z} F_{\alpha\beta\gamma}^T}_{\text{transformation uncertainty}} + \underbrace{F_{q_x q_y q_z} \Sigma_q F_{q_x q_y q_z}^T}_{\text{map uncertainty}} \quad (13)$$

Algorithm 2 Uncertainty Propagation and Map Merging

Require: The maps m_r and $m_{r'}$, and transformation function T

Ensure: Transform the map m_r and fuse them

for all voxels q **do**

1. represent the source voxel as multi-Gaussian distribution: $\mathcal{N}_q(\mu_q, \Sigma_q)$
2. transform the voxel q to r : $\mu_r = T(\mu_q)$
3. calculate Σ_r using Eq. (13)
4. represent the transformed voxel r with $\mathcal{N}_r(\mu_r, \Sigma_r)$
5. merge r with corresponding voxel p to get voxel c : $\mathcal{N}_c(\mu_c, \Sigma_c) \sim \mathcal{N}_r(\mu_r, \Sigma_r) \cdot \mathcal{N}_p(\mu_p, \Sigma_p)$

end for

B. Uncertainty merge

We employ our previous work on map merging [4] to execute the sequential pair-wise uncertainty merging. For the two types of uncertainty, the merging process should preserve all valuable information of the partial maps while decreasing the uncertainty of the fused map. A relative entropy filter divergence is applied to merge map measurements based on Kullback Leibler, which decreases the total uncertainty of the global map.

In order to integrate multidimensional Gaussian distribution, we apply a three-dimensional statistical fusion method based on matrix operations. Let $\mathcal{N}_x(\mu, \Sigma)$ denote the density function of voxel x , then,

$$\mathcal{N}_c(\mu_c, \Sigma_c) \sim \mathcal{N}_r(\mu_r, \Sigma_r) \cdot \mathcal{N}_p(\mu_p, \Sigma_p) \quad (14)$$

,where

$$\mu_c = (\Sigma_r^{-1} + \Sigma_p^{-1})^{-1} (\Sigma_r^{-1} \mu_r + \Sigma_p^{-1} \mu_p) \quad (15)$$

$$\Sigma_c = (\Sigma_r^{-1} + \Sigma_p^{-1})^{-1} \quad (16)$$

Then we extend the steps to the entire voxels in the map, where Algorithm 2 show the overall process. Firstly, the source voxel q and target voxel p is represented as a 3D Gaussian distribution. In line 2-4, q is transformed to r according to the given transformation function T , where covariance of transformed voxel is calculated using Eq. (13). Then the final merging happens in line 5, where the uncertainty of both source and target map is fused using Eq. (14).

Fig. 5. The experiment setting up: NTU Indoor, NTU Mixed, Changi Centre, NTU Forest respectively.

VI. EXPERIMENTAL RESULTS

The performance of the proposed multi-robot map fusion framework is validated through extensive real-world experiments. Both qualitative and quantitative results are presented. The former is obtained by visual inspection of the fused map in different experimental settings. Then, the latter is expressed by calculating the root mean square error and the running time.

A. Evaluation Protocol

Experimental Environment: Four real experiments were conducted, where two robots were equipped with heterogeneous sensors. The resolution of the 3D occupancy grid map was set to be 0.1m. The communication between robots was established by long-range Wi-Fi with limited bandwidth. The experiment settings are:

a) NTU Indoor: two Husky robots equipped with ASUS Xtion Camera and Velodyne 3D Lidar in an indoor environment of Nanyang Technological University.

b) NTU Mixed: two Husky robots equipped with ASUS Xtion Camera and Velodyne 3D Lidar in the mixed indoor and outdoor environment of Nanyang Technological University.

c) Changi Centre: two Husky robots equipped with Velodyne 3D Lidar during Asia Unmanned Exhibition in Changi Exhibition Centre.

d) NTU Forest: two Husky robots equipped with Velodyne 3D Lidar in a forest of Nanyang Technological University.

Fig. 5 and Table. I summarize the setting and the main characteristics of each experiment.

Error Metric: The error metric applied is MSE (Mean Squared Error), which is the mean squared Euclidean distance between the matched pairs.

Comparison Baseline: For each experiment, we compare two existing registration algorithms with the proposed approach. First, we compare the proposed algorithm with OICP [4], which is one of the one-to-one data association correspondence methods. Second, NDT is compared with the proposed

TABLE I
SUMMARIZE OF THE EXPERIMENT ENVIRONMENT.

Dataset	Sensor	Location	Dynamics	Type
NTU Indoor	Asus & Velodyne	Indoor	Dynamic	Structured
NTU Mixed	Asus & Velodyne	Mixed	Static	Structured
Changi Centre	Velodyne	Indoor	Dynamic	Semi-Stru
NTU Forest	Velodyne	Outdoor	Static	Unstructured

(a) At $t=0s$, Velodyne map is very sparse. The left image shows the successful registration of multiple probabilistic data association (average associated voxels=3.57), while the right shows the large offset between two maps.

(b) At $t=20s$, Velodyne map gets denser. It can be seen that the density of Velodyne map is increasing. Thus, the average associated voxels of multiple data association strategy decrease to 2.15.

Fig. 6. The experiment of heterogeneous map matching with different data association strategies.

algorithm [22] and implemented based on the standard Point Cloud Library [31].

B. NTU Indoor

The indoor environment is the second floor of a building in the university which contains challenging objects such as reflective glasses and moving people (see Fig. 5a). In order to test the algorithm's flexibility of using different sensors with varying characteristics, two robots were teleoperated to explore the environment with different sensors (Asus Xtion PRO and Velodyne VLP-16). Map fusion was implemented at two levels. First, the dense map produced by Asus Xtion and sparse map produced by Velodyne were fused to generate the local fused map. Then, the local fused maps were sent to the neighboring robot to generate the global fused map.

At the local level, the Velodyne map is very sparse in the initial position (See Fig. 6a), while the Asus Xtion produces dense map at the same time. The matching result of the sparse map and dense map presented in Fig. 6a illustrates the superiority of the multiple data association strategy. At $t = 0$, each voxel in the sparse map correlates average 3.47 voxels in the dense map to establish more robust correspondences. With the increasing density of the Velodyne map, the average correlated voxels drops to 2.15 at $t = 20s$ and two algorithms produce comparable results. In summary, despite the heterogeneity of sensors and challenging environments, the proposed method

Fig. 7. The result of multi-level map fusion result in NTU indoor experiment. Top: individual map produced by heterogeneous sensors. Middle: the locally fused map by two different robots. Bottom: the global map by fusing two local maps.

Fig. 8. The result of multi-level map fusion result in NTU Mix experiment. Top: individual map produced by heterogeneous sensors. Middle: the locally fused map by two different robots. Bottom: the global map by fusing two local maps.

is capable of generating a more consistent map that combines sparse 3D LiDAR map with dense RGB-D map. (see Fig. 6b). In the global map, the local maps produced by two robots were fused into a global consistent map in Fig. 7 (top part). As shown in Fig. 7, the overlapped area between the two maps combines the advantages of the two sensors, which produces a more detailed and abundant fused map.

C. Mixed Ground Floor

The mixed environment depicted in Fig. 5b is at the ground floor of a building in the university, where the indoor environment is the lobby, and the outdoor environment is the affiliated car park. Two robots were teleoperated to explore the environment with different sensors (Asus Xtion PRO and Velodyne VLP-16). Fig. 8 shows the fusion of maps on local level and global level. Robot 1 generates a map in an indoor environment, where the Asus Xtion PRO captures most of the information in the lobby. As shown in the figure, the algorithm combines the map information from two heterogeneous maps to generate an enhanced and more consistent global map.

Fig. 9 shows the local magnified result of map fusion, which details the highlighted part marked by the red ellipse in Fig. 8. Due to the limited field of view and range of sensing, the

Fig. 9. Left: magnified map fusion results in NTU Mixed experiment. Right: zoom in of the uncertainty merging.

Kinect generated a dense map without the wall, while the Velodyne created a less dense map consisting of most of the environmental information. After map fusion, the fused map is dense and compact. The right part of Fig. 9 shows the zoom-in results of the detailed uncertainty merging process. The 3D Gaussian sphere denotes the uncertainty of each map voxel, which is inversely proportional to the occupancy probability. After fusion, the probability information of the Kinect map and the lidar map are fused, and the probabilistic Gaussian distribution of the global map is obtained (See Fig. 9 bottom right).

D. Large Exhibition Center

The Changi Center depicted in Fig. 5c is at a convention center during the Asia Unmanned System Exhibition in Singapore, where the red arrow and trajectory is robot 1, the black represents robot 2. Two robots started their mapping from a common neighborhood and explored the environment by traversing two different paths. It differs from the previous two experiments since it only equips Velodyne 3D Lidar. As shown in Fig. 10, the fusion starts from very sparse Velodyne maps to denser Velodyne maps, where the fusion process requires the data association to dynamically adjust the number of voxels to be associated. In the initial position, a more thorough and extensive correspondences will be established to calculate relative transformation. Then, the average number of voxels to be associated is decreased as the maps grow denser. Finally, the two local maps are fused into a more consistent and enhanced global map. More specifically, the global map filled the gaps in the partial maps due to loss of data and sensor limitations.

E. Forest Environment

Experiment conducted in the forest environment is presented in this part. Each robot was equipped with a Velodyne VLP-16

(a) $t=0s$, the initial position of the two robots. Two sparse locals maps are successfully fused into a global map.

(b) $t=76s$, half of the mission. The local maps become denser, and global map are fused.

(c) $t=145s$, the end of the mission. The two robots back to the start position and the global map of the full environment is presented.

Fig. 10. The result of the collaborative mapping system at exhibition center.

Lidar. The forest is a completely unstructured 3D environment that contains trees and slopes (see Fig. 5d). It differs from the previous three experiments since it is a fully unstructured environment. Two robots started from a nearby place, while robot 1 moved straight forward and robot 2 climbed to a small hill and turned back to the start position. As presented in Fig. 11, the map produced by robot 1 is semi-structured, while robot 2 produced a fully unstructured forest map. The partial maps were fused into a consistent global map. The algorithm combined the map information from partial maps and generated an enhanced and more consistent global map.

F. Analysis of Experiment Results

In this subsection, the quantitative results in respect of accuracy and efficiency are presented.

Table. II summarizes RMS error for all experiments. The trend is clear that the smallest error in Table. II exhibit the highest performance of estimation accuracy. During the process, the RMS errors for most experiments are decreasing. All the algorithms witness a large offset at the starting stage, which is caused by the low density of the map. As the maps become denser, more correspondences between the voxels have been established and the errors decrease. It can also be observed that RMS errors in Changi Center and NTU Forest are larger

TABLE II
THE ROOT MEAN ERROR IN ALL EXPERIMENTS. (THE BEST PERFORMANCE IS DENOTED IN BOLD.)

Dataset	Beginning			Middle			Final		
	OICP	NDT	Ours	OICP	NDT	Ours	OICP	NDT	Ours
NTU Indoor	0.3414	0.3263	0.2479	0.2592	0.2472	0.2357	0.2675	0.2718	0.2890
NTU Mixed	0.3140	0.2903	0.2791	0.3078	0.2834	0.2563	0.2818	0.2739	0.2542
Changi	0.4405	0.4487	0.4284	0.4054	0.4142	0.3998	0.3653	0.3706	0.3550
NTU Forest	0.4393	0.4698	0.4301	0.3561	0.3929	0.3374	0.3430	0.3524	0.3237

(a) $t=0s$, the initial position of the two robots. Two sparse locals maps are successfully fused into a global map.

(b) $t=120s$, half of the mission. The local maps become denser, and global map are fused.

(c) $t=215s$, the end of the mission. The global map of the full environment is presented.

Fig. 11. The result of the collaborative mapping system at NTU forest.

compared with NTU Indoor and NTU Mixed. The reason is that Changi Center and NTU Forest are highly unstructured environments, while NTU Indoor and NTU Mixed are more structured. As shown in Table. II, the errors of three algorithms increase in the final stage of NTU Indoor environment, which might be caused by the accumulated map error at the single robot level, resulting in a certain degree of error in map fusion results. It can also be noticed that the errors of NDT and our algorithm both increase more than that of OICP. More specifically, our algorithm establishes multiple data association and the error term is magnified by the disturbance, which results in a sub-optimal local minima. This is an inherent property of all dense registration algorithms as the global minima cannot be guaranteed. Overall, our algorithm provides superior or comparable performance compared with OICP and NDT in most cases.

Fig. 12. The average numbers of voxels to be associated in our algorithm during the four experiments.

Fig. 13. The processing time during the four experiments on both single robot and multi robot levels.

In NTU Indoor and NTU Mixed experiments, two robots were equipped with heterogeneous sensors. Our algorithm dynamically adjusted the data association strategy in the matching process (See Fig. 12). As a result, the RMS error of our algorithm improved significantly of 10%-20% compared with OICP and NDT in both two scenarios in most cases, as summarized in Table. II. In Changi Center and NTU Forest, only Velodyne 3D Lidar was equipped. The difference in density is not that significant, and the number of voxels to be associated are close to one (See Fig. 12). It is shown that one-to-one correspondence cannot establish sufficient correspondence. Besides, original NDT algorithm models the raw sensor data of the point cloud into Gaussian distributions and minimizes the L2 norm error metric. Since in our case the voxels are evenly distributed with a distance of 0.1m, the Gaussian distributions may not represent the environment explicitly, so the error metric loses its original superiority.

In addition to the evaluation of the error metrics, computation time is an important factor in real/semi-real time

robotics applications. Fig. 13 summarizes the computational time of the four experiments when executed on an Intel Core i7-6700HQ CPU @ 2.60GHz robot onboard computer. The proposed algorithm provides two levels of decision-making cycle for the robots, namely single robot mapping and multi-robot map fusion. The single robot mapping cycle is used to achieve real-time objectives and is generated within 100ms. For multi-robot missions, higher level objectives such as multi-robot map fusion and goal-point selection do not require high frequency performance. Considering connectivity, bandwidth and computational constraints, the distributed setting is to fuse local maps from individual robots sporadically. Therefore, the time interval between two such multi-robot map fusion processes is performed in semi-real time, which means the map fusion process is performed in a lower frequency. In multi-robot level, the computational time for NTU Forest is significantly higher than other scenarios due to the large scale and unstructured environment.

Remark 1: In the proposed framework, the two problems of map matching and map merging can be factorized into pairwise operations, which give the flexibility of using more robots. In a robot group with more than two members, map fusion can be performed in a serial manner. At one time instant, a robot only communicates with one specific robot for map fusion, while the third robot will be placed on the waiting list until the current operation is completed.

VII. CONCLUSIONS

This paper has provided a general probabilistic framework for collaborative 3D mapping, which is independent of sensor type and SLAM algorithm. The framework provides flexibility of solving map matching and map merging problem separately, which allows the team of robots to work with heterogeneous sensor type. In the paper, the Expectation-maximization (EM) based map matching algorithm is successfully applied to register 3D Lidar map and Kinect map, where the number of voxels to be associated is automatically decided based on the map characteristics. To distribute the maps among robots, the uncertainties of relative localization estimation and local map have been propagated and integrated. Then, the time-sequential map merging algorithm is utilized to transfer the local maps between the teams of robots. The proposed approach is evaluated in various environments with heterogeneous sensors, such as forest, dynamic, and mixed indoor-outdoor environments. The results verify the accuracy, efficiency, and generality for collaborative mapping missions. The proposed mapping framework can be utilized for complex environment mapping for real/semi-real time multiple mobile robots missions.

For hybrid teams of robots (UGV and UAV), the significant difference in viewpoint and robot kinematics brings the challenge of finding the corresponding regions that enables accurate relative localization. With the fast development in deep learning, it is promising to conduct the CNN-based approach to assign correspondence at semantic levels, which will increase the robustness and efficiency.

REFERENCES

- [1] P. C. Niedfeldt, A. Speranzon, and A. Surana, "Distributed map fusion with sporadic updates for large domains," in *2015 IEEE International Conference on Robotics and Automation*, May 2015, pp. 2806–2813.
- [2] S. Saeedi, M. Trentini, M. Seto, and H. Li, "Multiple-Robot Simultaneous Localization and Mapping: A Review," *Journal of Field Robotics*, vol. 33, no. 1, pp. 3–46, 2016.
- [3] Y. Wang, M. Shan, Y. Yue, and D. Wang, "Vision-based flexible leader-follower formation tracking of multiple nonholonomic mobile robots in unknown obstacle environments," *IEEE Transactions on Control Systems Technology*, pp. 1–9, 2019.
- [4] Y. Yue, D. Wang, P. Senarathne, and D. Moratuwage, "A hybrid probabilistic and point set registration approach for fusion of 3D occupancy grid maps," in *2016 IEEE International Conference on Systems, Man, and Cybernetics*, Oct 2016, pp. 1975–1980.
- [5] S. Bansal, V. Tolani, S. Gupta, J. Malik, and C. Tomlin, "Combining optimal control and learning for visual navigation in novel environments," *CoRR*, vol. abs/1903.02531, 2019. [Online]. Available: <http://arxiv.org/abs/1903.02531>
- [6] J. Zhong, H. Cheng, L. He, and F. Ouyang, "Decentralized full coverage of unknown areas by multiple robots with limited visibility sensing," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, April 2019.
- [7] H. Surmann, N. Berninger, and R. Worst, "3d mapping for multi hybrid robot cooperation," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2017, pp. 626–633.
- [8] A. Gawel, R. Dub, H. Surmann, J. Nieto, R. Siegwart, and C. Cadena, "3d registration of aerial and ground robots for disaster response: An evaluation of features, descriptors, and transformation estimation," in *2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*, Oct 2017, pp. 27–34.
- [9] S. Saeedi, L. Paull, M. Trentini, and Seto, "Group mapping: A topological approach to map merging for multiple robots," *Robotics & Automation Magazine, IEEE*, vol. 21, no. 2, pp. 60–72, 2014.
- [10] J. Jessup, S. N. Givigi, and A. Beaulieu, "Robust and efficient multirobot 3-d mapping merging with octree-based occupancy grids," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1723–1732, 2017.
- [11] T. M. Bonanni, B. D. Corte, and G. Grisetti, "3-D Map Merging on Pose Graphs," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1031–1038, April 2017.
- [12] D. Fox, J. Ko, K. Konolige, B. Limketkai, D. Schulz, and B. Stewart, "Distributed multirobot exploration and mapping," *Proceedings of the IEEE*, vol. 94, no. 7, pp. 1325–1339, July 2006.
- [13] A. Cunningham, K. M. Wurm, W. Burgard, and F. Dellaert, "Fully distributed scalable smoothing and mapping with robust multi-robot data association," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, May 2012, pp. 1093–1100.
- [14] A. Gawel, T. Cieslewski, R. Dub, M. Bosse, and R. Siegwart, "Structure-based vision-laser matching," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2016, pp. 182–188.
- [15] J. Jessup, S. N. Givigi, and A. Beaulieu, "Robust and Efficient Multi-robot 3-D Mapping Merging With Octree-Based Occupancy Grids," *IEEE Systems Journal*, vol. PP, no. 99, pp. 1–10, 2015.
- [16] Y. Yue, P. G. C. N. Senarathne, C. Yang, J. Zhang, M. Wen, and D. Wang, "Hierarchical probabilistic fusion framework for matching and merging of 3-d occupancy maps," *IEEE Sensors Journal*, vol. 18, no. 21, pp. 8933–8949, Nov 2018.
- [17] C. Brand, M. J. Schu, and H. Hirsch, "Submap matching for stereo-vision based indoor/outdoor SLAM," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2015, pp. 5670–5677.
- [18] Y. Yue, D. Wang, P. Senarathne, and C. Yang, "Robust submap-based probabilistic inconsistency detection for multi-robot mapping," in *2017 European Conference on Mobile Robots (ECMR)*, Sept 2017, pp. 1–6.
- [19] M. J. Schuster, C. Brand, H. Hirschmüller, M. Suppa, and M. Beetz, "Multi-robot 6D graph SLAM connecting decoupled local reference filters," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 5093–5100.
- [20] P. J. Besl, N. D. McKay *et al.*, "A method for registration of 3-d shapes," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 14, no. 2, pp. 239–256, 1992.
- [21] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp," in *Robotics: science and systems*, vol. 2, no. 4, 2009, p. 435.
- [22] J. P. Saarninen, H. Andreasson, T. Stoyanov, and A. J. Lilienthal, "3d normal distributions transform occupancy maps: An efficient representation for mapping in dynamic environments," *The International Journal of Robotics Research*, vol. 32, no. 14, pp. 1627–1644, 2013.

- [23] J. Liu, S. Du, C. Zhang, and J. Zhu, "Probability iterative closest point algorithm for position estimation," in *International IEEE Conference on Intelligent Transportation Systems*, Oct 2014, pp. 458–463.
- [24] G. Agamennoni, S. Fontana, R. Y. Siegwart, and D. G. Sorrenti, "Point clouds registration with probabilistic data association," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2016, pp. 4092–4098.
- [25] S. L. Bowman, N. Atanasov, K. Daniilidis, and G. J. Pappas, "Probabilistic data association for semantic slam," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2, 2017.
- [26] A. Cunningham, V. Indelman, and F. Dellaert, "Ddf-sam 2.0: Consistent distributed smoothing and mapping," in *2013 IEEE International Conference on Robotics and Automation*, May 2013, pp. 5220–5227.
- [27] M. Guo and M. M. Zavlanos, "Multi-robot data gathering under buffer constraints and intermittent communication," *arXiv preprint arXiv:1706.02092*, 2017.
- [28] A. Censi, "An accurate closed-form estimate of ICP's covariance," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, April 2007, pp. 3167–3172.
- [29] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees," *Autonomous Robots*, 2013.
- [30] A. W. Stroupe, M. C. Martin, and T. Balch, "Distributed sensor fusion for object position estimation by multi-robot systems," in *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation*, vol. 2, 2001, pp. 1092–1098 vol.2.
- [31] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 1–4.

NAMAL SENARATHNE graduated with a B.Sc. (Computer Science & Engineering) from University of Moratuwa, Sri Lanka in 2009 and with a Ph.D from Nanyang Technological University in 2015. During his Ph.D. he worked on improving 2D exploration algorithms that uses the frontier heuristic. At ST Engineering - NTU Corporate Lab, he worked on research and development of efficient algorithms for UGV and UAV exploration in 3D environments, multi-robot collaborative exploration, 3D perception for navigation of mobile robots and 3D path planning for UAVs. Currently he's a Senior Specialist in Robotics at Continental Automotive Singapore Pte. Ltd.

Jun Zhang is a 3rd year Ph.D. student at the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Before joining NTU, he received the M.Eng., in Mechatronics from Huazhong University of Science and Technology (HUST), with the innovation team of HUSTGET, in Wuhan, China, 2015. Before that, he obtained the B.Eng. in Mechanical Engineering from HUST in 2012. Currently, his research focuses on place recognition and localization, and multi-sensor calibration.

Yufeng Yue received the B.Eng. degree in automation from Beijing Institute of Technology, Beijing, China, in 2014, the Ph.D. degree from Nanyang Technological University, Singapore, in 2019. He is currently a Research Fellow with Advanced Robotics Lab, Nanyang Technological University, Singapore. He was a visiting scholar with University of California, Los Angeles in 2019. His research interests include mapping, navigation, coordination and reasoning for multi-robot systems in complex environments.

Mingxing Wen has received his B.Eng. degree of Automation and M.Sc. degree of Computer Control and Automation from Beijing Institute of Technology, Beijing, China in 2015 and Nanyang Technological University, Singapore in 2016 respectively. After working as a Research Associate in NTU, he is currently pursuing the Ph.D. in School of Electrical and Electronic Engineering, NTU. His research focuses on the robot learning, and reinforcement learning applications in autonomous system.

Chule Yang received his B.Eng. degree in Electrical Engineering from Wuhan University, Wuhan, China, in 2014. He received his M.Sc. degree in Computer Control and Automation from Nanyang Technological University, Singapore, in 2015. He is currently a Ph.D. student with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include multimodal perception, probabilistic reasoning and intelligent decision making for autonomous system.

Danwei Wang is leading the autonomous mobile robotics research group. He received his Ph.D and MSE degrees from the University of Michigan, Ann Arbor in 1989 and 1984, respectively. He received his B.E degree from the South China University of Technology, China in 1982. Since 1989, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Currently, he is professor and the co-director of the ST Engineering NTU Corporate Laboratory. He is a senator in NTU Academics Council. He

has served as general chairman, technical chairman and various positions in international conferences, such as ICARCV and IROS conferences. He is an associate editor for the International Journal of Humanoid Robotics and served as an associate editor of Conference Editorial Board, IEEE Control Systems Society from 1998 to 2005. He was a recipient of Alexander von Humboldt fellowship, Germany. His research interests include robotics, control theory and applications.

Yuanzhe Wang received the B.Eng. degree from the Southeast University, Nanjing, China, in 2010, the M.Eng. degree from the Beihang University, Beijing, China, in 2013, and the Ph.D. degree from the Nanyang Technological University, Singapore, in 2019. He has served as a Research Associate in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2018 to 2019, where he is currently a Research Fellow. His current research interests include robot motion planning and control, multi-robot coordination and human-robot interaction.