2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

MM4MM: Map Matching Framework for Multi-Session Mapping in Ambiguous and Perceptually-Degraded Environments

Zhenyu Wu[1], Wei Wang[1], Chunyang Zhao[1], Yufeng Yue[2], Jun Zhang[1], Hongming Shen[1], and Danwei Wang[1], *Fellow, IEEE*

***Abstract*— Multi-session mapping serves as the pre-requisite for autonomous robots to fulfill various long-term tasks (*e.g.*, map updating, navigation, collaboration). However, it is challenging to implement multi-session mapping in enclosed or partially enclosed ambiguous environments (*e.g.*, long corridors, industrial warehouses). Existing solutions either depend heavily on the matching of elementary geometric features (*e.g.*, points, lines, and planes), which tends to fail in environments with ambiguous geometric features; or depend on the given guess of the initial transformation matrix of multiple single-session maps, which is not always obtainable and accurate enough. The ambient magnetic field has exhibited ubiquity and high distinctiveness at different location, which makes it suitable for estimating the initial transformation matrix. Thus, this paper proposes a novel probabilistic magnetic-aware Map Matching framework for Multi-session Mapping, namely MM4MM, to estimate the relative transformation of multiple single-session maps and to build the globally consistent maps in ambiguous and perceptually-degraded environments. The key novelties of this work are the designing of the hierarchical probabilistic map matching framework and the Particle Swarm Optimization strategy to associate the magnetic data of multiple sessions. Evaluations on both simulated and real world experiments demonstrate the greatly improved utility, accuracy, and robustness of multi-session mapping over the comparative methods.**

I. Introduction

In recent years, single-session simultaneous localization and mapping (SLAM) in ordinary environments with sufficient geometric features is considered as efficiently addressed with many existing solutions [1]–[6]. However, multi-session mapping for long-term and large-scale robotic operations remains as one of the key issues that are still unresolved [2], [7]–[10]. Multi-session mapping empowers the autonomous robots with the capabilities to perform long-term tasks (*e.g.*, place recognition [11], [12], collaboration [13]–[15]) in a robust and persistent way. In enclosed or partially enclosed ambiguous environments (*e.g.*, long corridors, car parks, industrial warehouses), the geometric features can be highly ambiguous [16] and the GNSS signals are considered as challenged or unreliable. Although partial maps can be built from each single-session data acquisitions in such scenarios, it is challenging to build the scalable and globally consistent map due to the perceptual degradation and data association ambiguities.

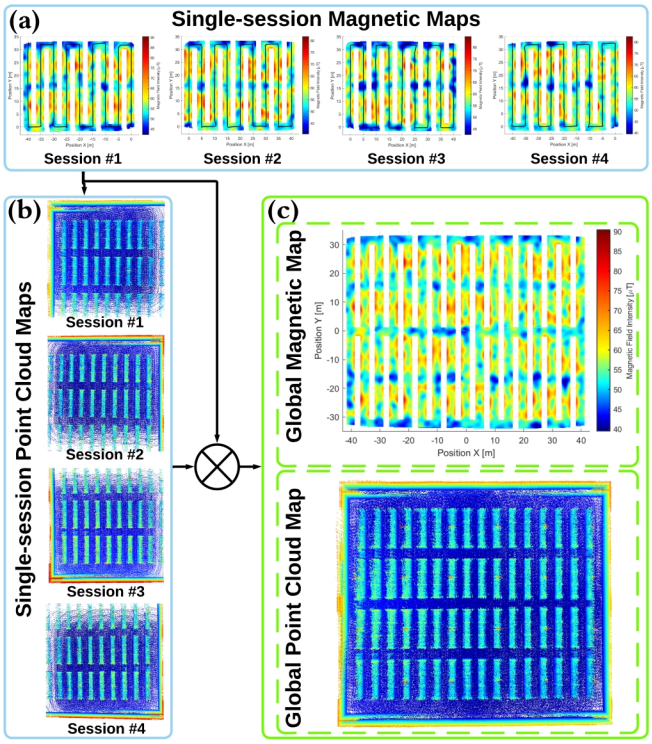


Fig. 1. The proposed MM4MM system working in a simulated industrial warehouse environment with ambiguous geometric features. (a) and (b): The multiple single-session MF and point cloud mappings, where colors of the ambient MF designate the MF integrated intensity and black lines in (a) denote the robot trajectories. (c) Utilizing the estimated relative transformation matrices, the globally consistent magnetic map and point cloud map can be obtained.

Specifically, multi-session mapping mainly involves dealing with the kidnapped robot or unknown initial positioning problem [8], where the robot is shut down after a single-session mapping operation and moved to another place without knowing *a priori* information about the position [7]. In previous studies, multi-session mapping algorithms mainly involve techniques such as map matching [2], [17], map merging [18], [19], and loop closure detection [7], [8], [10]. Considering the trade-off between utility and accuracy, map matching is considered as a promising solution to estimate

This research is supported by the National Research Foundation, Singapore under its Medium Sized Centre for Advanced Robotics Technology Innovation (CARTIN), the Agency for Science, Technology and Research (A*STAR) under its National Robotics Programme (Project No. M22NBK0109), and National Research Foundation, Singapore and Maritime and Port Authority of Singapore under its Maritime Transformation Programme (Project No. SMI-2022-MTP-04). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore and Maritime and Port Authority of Singapore.

[1]Z. Wu, W. Wang, C. Zhao, J. Zhang, H. Shen, and D. Wang are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. E-mail: zhenyu002@e.ntu.edu.sg

[2]Y. Yue is with the School of Automation, Beijing Institute of Technology, Beijing 100081, China.

the relative transformation matrix between multiple maps and construct globally consistent maps of the environment.

Existing research of map matching relies heavily on geometric features such as points [9] and lines [20] or visual image features [2], [17], which tend to fail in ambiguous structured environments where the geometric features are highly similar. Recently, the ambient magnetic field (MF)-based mapping and localization research has attracted dramatic attention in the robotics and indoor positioning community [21]–[33]. Due to its prominent location signatures and infrastructure-free non-intrusive characteristics, MF has became a viable alternative to GNSS signal in the local environments for autonomous robots localization [22]–[27], [29], [30], [33] and navigation [21], [31]. For the magnetic map building in single-session tasks, popular methods such as fingerprints (FP)-based [22], [28], linear interpolation (LI)-based [23], [25], [27] and Gaussian Process Regression (GPR)-based [26], [29]–[31] approaches are mostly utilized. For the multi-session magnetic mapping, the problem remains virtually unexplored. The magnetic maps built from different sessions provide useful high-level magnetic sequence-based constraints for map matching and are ideal to assist the long-term multi-session mapping tasks.

Our previous work [29], [32] explored the feasibility of utilizing the MF to localize a single robot in ambiguous environments. Besides, our previous research mainly focus on how to utilize the semantic information to enhance the accuracy of relative localization between collaborative robots in outdoor environments with rich geometric features [14], [34]. However, the main challenge of multi-session mapping in ambiguous and perceptually-degraded environments is to design a hierarchical probabilistic map matching framework that calculates accurate transformation matrix between multi-session maps and build globally consistent maps. Thus in this paper, we propose a novel framework named MM4MM, which bridges the gap between magnetic map matching algorithms that are yet to be utilized in multi-session mapping tasks and geometric map matching algorithms that only leverage elementary geometric features. A brief illustration of the proposed framework is shown in Fig. 1. The main contributions of this work are summarized as:

- A hierarchical probabilistic map matching framework is proposed, which addresses the problem of multi-session mapping in ambiguous and perceptually-degraded environments.
- A novel magnetic data association strategy is proposed using Particle Swarm Optimization to obtain the initial transformation between multiple single-session maps.
- The proposed framework has been validated in both simulated and real-world challenging ambiguous environments by building the globally consistent maps, thus showing the greatly improved utility and accuracy.

II. System Framework and Problem Formulation

Workflow of the proposed MM4MM system is presented in Fig. 2, which comprises two submodules: *single-session mapping*, and *map matching for multi-session mapping*.

Multi-Session Level Problem Formulation: *The initial transformation matrix* $\mathbf{T_B}$ *is introduced as the hidden state to enhance multi-session mapping performance, which can be derived by the magnetic map matching. Given a set of local magnetic maps* $\mathbf{B}_{(r,\mathbf{S}_r)}$ *and point cloud maps* $\mathbf{M}_{(r,\mathbf{S}_r)}$ *built from a set of single-session mapping processes* $\mathbf{S}_r$*, the objective is to estimate the relative transformation matrix* $\mathbf{T_M}$ *between different single-session maps.*

$$
\begin{aligned}
&p\left(\mathbf{T_M},\mathbf{T_B}|\mathbf{B}_r,\Phi(\mathbf{T_B}^{r_1,r_n},\mathbf{B}_{r_n}),\mathbf{M}_r,\Phi(\mathbf{T_M}^{r_1,r_n},\mathbf{M}_{r_n})\right)\\
&=\sum_{i=1}^{K}\underbrace{p(\mathbf{T_B}|\mathbf{B}_r,\Phi(\mathbf{T_B}^{r_1,r_n},\mathbf{B}_{r_n}))}_{Magnetic-based}\cdot\underbrace{p(\mathbf{T_M}|\mathbf{T_B},\mathbf{M}_r,\Phi(\mathbf{T_M}^{r_1,r_n},\mathbf{M}_{r_n}))}_{Geometric-based} \quad (1)
\end{aligned}
$$

Due to the limitation of computational capacity, the single-session mapping process r_1 receives the local magnetic and point cloud maps $\{B_{r_n},M_{r_n}\}$ from another single-session $r_n \in \mathbf{S}_r$. The $\Phi_{r_n\in\mathbf{S}_r}(T_B^{r_1,r_n},\mathbf{B}_{r_n})$ and $\Phi_{r_n\in\mathbf{S}_r}(T_M^{r_1,r_n},\mathbf{M}_{r_n})$ are the transformation function that transforms $\mathbf{B}_{r_n}$ and $\mathbf{M}_{r_n}$ to the coordinate frame of $\mathbf{B}_{r_1}$ and $\mathbf{M}_{r_1}$, respectively. Once the local maps are built and shared, it is necessary to establish the data association to estimate the relative transformation between two single-session maps. In our framework, the *Particle Swarm Optimization (PSO)* algorithm and *Expectation-Maximization (EM)* algorithm are implemented to utilize magnetic and geometric information, respectively, to calculate the transformation matrix in a hierarchical way (see Section III-B.1) and Section III-B.2)). Finally, we implement *map merging* to form globally consistent maps.

III. Magnetic-Aware Map Matching for Multi-Session Mapping

A. Single-Session Magnetic and Geometric Mapping

1) ***Single-Session Magnetic Field Mapping****:* The 3-D ambient MF $\mathbf{B} \in \mathbb{R}^3$ measured at time t can be modeled as $\mathbf{B}_t^l = [B_t^x, B_t^y, B_t^z]^T$ in the local coordinate frame [23], [29]. We transformed the MF data to the global coordinate frame, which can be denoted as a rotation-invariant vector as $\mathbf{B}_t^g = [B_t^h, B_t^v, B_t^F]^T$, where the horizontal component $B_t^h = \sqrt{(B_t^x)^2+(B_t^y)^2}$ points to the *Magnetic North*, the vertical component $B_t^v = B_t^z$ being parallel with a unit vector of the Earth gravity field, and B_t^F denotes the MF integrated intensity. We define the magnetic map $\mathbf{B}_{r_n}$ as a set of MF fingerprints on specified grid points in the environment. The environment is divided into a grid of square cells of length $5cm$ and the MF vectors of each cell are collected as the fingerprints. Based on the collected MF fingerprints, we implement LI-based algorithm [27] to build the magnetic map, where we refer the details of this process to paper [27].

2) ***Single-Session Point Cloud Mapping****:* For the single-session point cloud mapping part, the point cloud map $\mathbf{M}_r$ is chosen to represent the environment in an efficient and compact way. To generate the point cloud map, the 3-D SLAM algorithm FAST-LIO2 [35] is implemented.

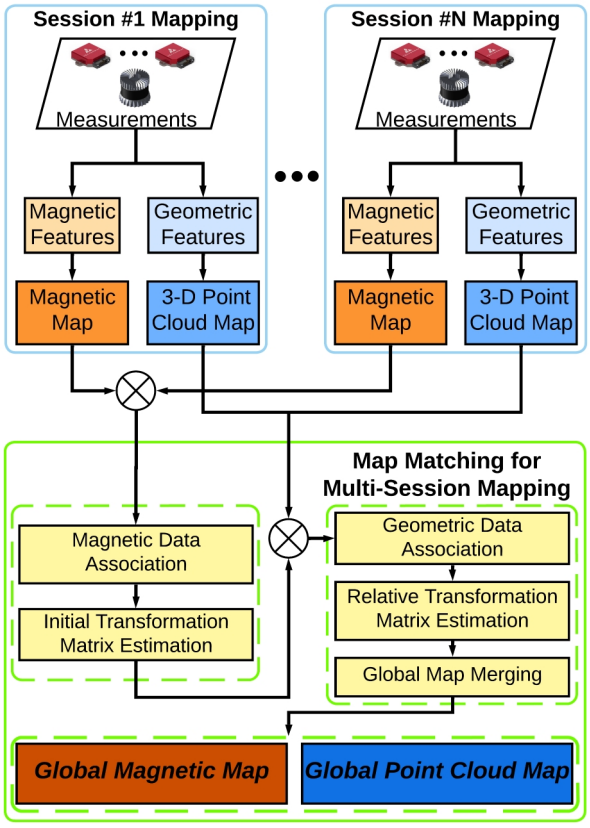


Fig. 2. Flowchart of the proposed MM4MM framework in ambiguous and perceptually-degraded environments. The local magnetic maps and point cloud maps produced by different single-session mapping processes are shared and matched, then the globally consistent maps can be built.

B. Multi-Session Magnetic and Geometric Map Matching

Refering to Eq. (1), we denote the inputs as the current single-session maps $\{\mathbf{B}_{r_1}, \mathbf{M}_{r_1}\}$ and the latest incoming map $\{\mathbf{B}_{r_n}, \mathbf{M}_{r_n}\}$ from another single-session process. We define each magnetic map and point cloud map as $\mathbf{B}_{r_n} = \{\mathbf{B}^i_{r_n}\}_{i=1}^{N_{\mathbf{B}_{rn}}}$ and $\mathbf{M}_{r_n} = \{\mathbf{M}^a_{r_n}\}_{a=1}^{N_{\mathbf{M}_{rn}}}$, where $N_{\mathbf{B}_{rn}}$ and $N_{\mathbf{M}_{rn}}$ are the number of MF fingerprints and geometric points, respectively.

*1) **Magnetic Map Matching**:* For the matching of two partial magnetic maps $\mathbf{B}_{r_1}$ and $\mathbf{B}_{r_n}$, the problem can be formulated as estimating the initial transformation matrix $\mathbf{T_B}$ with the highest probability, which is the *Magnetic-based* part in Eq. (1). We fix the $\mathbf{B}_{r_1}$ as the target map and utilize the other map $\mathbf{B}_{r_n}$ as the pairing map. The rigid transformation $\mathbf{T_B}$ transforms $\mathbf{B}_{r_n}$ to the coordinate of $\mathbf{B}_{r_1}$, where $\mathbf{T_B} \in SE(3)$ is a rigid transformation matrix. The *maximum a posterior (MAP)* estimate $\hat{\mathbf{T}}_\mathbf{B}$ of the matrix $\mathbf{T_B}$ is derived by applying Bayesian theorem as:

$$\begin{aligned}\hat{\mathbf{T}}_\mathbf{B} &= \underset{\mathbf{T_B}}{\operatorname{argmax}}\ \log\ p(\mathbf{T_B}|\mathbf{B}_{r_1}, \mathbf{B}_{r_n}) \\ &\propto \underset{\mathbf{T_B}}{\operatorname{argmax}}\ \log\Big(p(\mathbf{B}_{r_1}|\mathbf{T_B}, \mathbf{B}_{r_n}) \cdot p(\mathbf{T_B})\Big)\end{aligned} \tag{2}$$

where term $p(\mathbf{B}_{r_1}|\mathbf{T_B}, \mathbf{B}_{r_n})$ is a common *maximum likelihood estimation (MLE)* problem and $p(\mathbf{T_B})$ denotes the prior information. In this paper, we assume the prior transformation matrix is an uniform distribution and we focus on solving the *MLE* problem. To estimate $\hat{\mathbf{T}}_\mathbf{B}$, we need to establish the MF fingerprint-wise correspondences. With respect to the random variable $\mathbf{T_B}$, we formulate the estimation of $p(\mathbf{B}_{r_1}|\mathbf{T_B}, \mathbf{B}_{r_n})$ as a nonlinear optimization problem and denote it as:

$$\hat{\mathbf{T}}_\mathbf{B} = \underset{\mathbf{T_B}}{\operatorname{argmax}}\ \log\ p(\mathbf{B}_{r_1}|\mathbf{T_B}, \mathbf{B}_{r_n}) \tag{3}$$

Eq. (3) estimates the parameter $\hat{\mathbf{T}}_\mathbf{B}$ that maximizes the overlap between map $\mathbf{B}_{r_1}$ and map $\mathbf{B}_{r_n}$. However, Eq. (3) cannot be solved directly due to the reason that we must firstly estimate the data association between magnetic fingerprints.

a) Magnetic Data Association: Based on the well-known swarm intelligence [36], we propose a novel strategy to associate MF fingerprints from different mapping sessions utilizing the Particle Swarm Optimization (PSO) algorithm. In the PSO algorithm that we use, a swarm of particles n_L (*i.e.,* candidate solutions) iteratively move within a 3-D search region with bounds in order to jointly search for the best solution, where the 3-D search space comprises the translation in x-direction, translation in y-direction, and the rotation. In order to associate MF fingerprints from two single-session mapping processes, the MF errors at each corresponding pair should be as low as possible and the overlap region of the two magnetic map should be as large as possible. Thus, the fitness function D_B of the PSO algorithm in our framework is specifically designed as:

$$D_B(\mathbf{B}^i_{r_1}, \mathbf{B}^j_{r_n}) = \underset{i\in N_{\mathbf{B}_{r_1}}, j\in N_{\mathbf{B}_{rn}}}{\operatorname{argmin}} \left\{ \frac{\widehat{\operatorname{Var}}(||\mathbf{B}^i_{r_1} - \mathbf{B}^j_{r_n}||)}{\sqrt{P^n_1}} \right\} \tag{4}$$

$$P^n_1 = \frac{\sum_{i=1}^{N_{\mathbf{B}_{r_1}}} \sum_{j=1}^{N_{\mathbf{B}_{rn}}} As(\mathbf{B}^i_{r_1}, \mathbf{B}^j_{r_n})}{N_{\mathbf{B}_{r_1}} + N_{\mathbf{B}_{rn}}} \tag{5}$$

where $\widehat{\operatorname{Var}}(\cdot)$ denotes the variance of the magnetic value differences calculated by Euclidean distance between any two corresponding MF fingerprints. P^n_1 depicts the percentage of overlapping fingerprints over the total number of fingerprints of the two magnetic maps. $As(\mathbf{B}^i_{r_1}, \mathbf{B}^j_{r_n})$ in Eq. (5) is 1 when $||\mathbf{B}^i_{r_1}|| = ||\mathbf{B}^j_{r_n}|| \pm 1.0\mu$T and 0 otherwise. The fitness function measures the similarities between the matched magnetic maps, where higher similarities can be obtained with lower MF error variances and larger overlap region. Thereafter, the probability of $p(\mathbf{B}_{r_1}|\mathbf{T_B}, \mathbf{B}_{r_n})$ can be denoted as:

$$p(\mathbf{B}_{r_1}|\mathbf{T_B}, \mathbf{B}_{r_n}) \propto \prod_{i\in N_{\mathbf{B}_r}} \frac{1}{\sigma_B\sqrt{2\pi}} \exp\Big(-\frac{1}{2}\frac{D^2_B(\mathbf{B}^i_{r_1}, \mathbf{B}^j_{r_n})}{\sigma^2_B}\Big) \tag{6}$$

where σ_B denotes the uncertainty caused by magnetic map building error and sensor noises. Thus, Eq. (3) can be calculated and $\mathbf{T_B}$ can be estimated correspondingly.

*2) **Geometric Map Matching**:* Firstly, the overlapping area can be delimited by the aforementioned magnetic data association. For the matching of two partial point cloud maps $\mathbf{M}_{r_1}$ and $\mathbf{M}_{r_n}$, the problem can be formulated as estimating the relative transformation matrix $\mathbf{T_M}$ with the highest probability, which is the *Geometric-based* part in Eq. (1). Similarly, the rigid transformation $\mathbf{T_M} \in SE(3)$ transforms $\mathbf{M}_{r_n}$ to the coordinate of $\mathbf{M}_{r_1}$. The *MAP* estimate $\hat{\mathbf{T}}_\mathbf{M}$ of matrix $\mathbf{T_M}$ is derived as:

$$\begin{aligned}\hat{\mathbf{T}}_\mathbf{M} &= \underset{\mathbf{T_M}}{\operatorname{argmax}}\ \log\ p(\mathbf{T_M}|\mathbf{T_B},\mathbf{M}_{r_1},\mathbf{M}_{r_n}) \\ &\propto \underset{\mathbf{T_M}}{\operatorname{argmax}}\ \log\Big(p(\mathbf{M}_{r_1}|\mathbf{T_M},\mathbf{M}_{r_n})\cdot p(\mathbf{T_M}|\mathbf{T_B})\Big) \end{aligned} \tag{7}$$

where the term $p(\mathbf{M}_{r_1}|\mathbf{T_M},\mathbf{M}_{r_n})$ is also a *MLE* problem. $p(\mathbf{T_M}|\mathbf{T_B})$ denotes the prior estimation of the transformation matrix from previous magnetic map matching. To estimate $\hat{\mathbf{T}}_\mathbf{M}$, we need to establish the geometric point-wise correspondences. We define each point $\mathbf{M}_r^a$ in 3-D point cloud map as $\mathbf{M}_{r_n}^a = \{m_{r_n}^a\} \in \mathbb{R}^3$, where $m_{r_n}^a$ denotes the 3-D coordinates. Similarly, the estimation of $p(\mathbf{M}_{r_1}|\mathbf{T_M},\mathbf{M}_{r_n})$ can be formulated as a nonlinear optimization problem as:

$$\hat{\mathbf{T}}_\mathbf{M} = \underset{\mathbf{T_M}}{\operatorname{argmax}}\ \log\ p(\mathbf{M}_{r_1}|\mathbf{T_M},\mathbf{M}_{r_n}) \tag{8}$$

Eq. (8) estimates the parameter $\hat{\mathbf{T}}_\mathbf{M}$ that maximizes the overlap between map $\mathbf{M}_{r_1}$ and map $\mathbf{M}_{r_n}$. Similarly, the data association between geometric points must be estimated first then Eq. (8) can be solved.

Since the geometric point-wise correspondence is not *a priori* information, it is necessary to infer the hidden data association. We utilize $\mathbf{D}_g = \{d_{a,b}\}_{a=1}^{N_{\mathbf{M}_{r_1}}}$ to represent all the correspondence sets between map $\mathbf{M}_{r_1}$ and $\mathbf{M}_{r_n}$. Let $d_{a,b}$ denotes the correspondence relationship between point $\mathbf{M}_{r_1}^a$ and point $\mathbf{M}_{r_n}^b$, where $d_{a,b}=1$ signifies that point $\mathbf{M}_{r_1}^a$ corresponds to point $\mathbf{M}_{r_n}^b$. Then, Eq. (8) can be rewritten into the format of Expectation-Maximization (EM) problem in Eq. (9) given that the geometric data are associated on the point level, where $d_{a,b}$ is simplified as d_a and $a \in N_{\mathbf{M}_{r_1}}$:

$$\begin{aligned}\hat{\mathbf{T}}_\mathbf{M} &= \underset{\mathbf{T_M}}{\operatorname{argmax}}\ \mathbb{E}_{\mathbf{D}_g}\left[p(\mathbf{M}_{r_1},\mathbf{D}_g|\mathbf{T_M},\mathbf{M}_{r_n})\right] \\ &= \underset{\mathbf{T_M}}{\operatorname{argmax}} \prod_a \sum_{d_a\in\mathbf{D}_g} \underbrace{p(d_a|\mathbf{M}_{r_1}^a,\mathbf{M}_{r_n}^b,\mathbf{T_M})}_{\text{Expectation}} \cdot \underbrace{p(\mathbf{M}_{r_1}^a|d_a,\mathbf{T_M},\mathbf{M}_{r_n}^b)}_{\text{Maximization}} \end{aligned} \tag{9}$$

In general, the EM problem consists of the Expectation (E)-step and Maximization (M)-step. Firstly, the expectations of the hidden data associations are estimated by the E-step. Then, given the hidden variables estimated by E-step, nonlinear optimization methods can be implemented to solve the transformation matrix in the M-step. We iteratively perform the E-step and M-step until the convergence requirement is satisfied. We refer the details of EM derivation to paper [37].

a) Geometric Data Association: We introduce the traditional search of Euclidean distance based nearest neighbor as our geometric map matching approach. Thus, the E-step can be derivated into Eq. (10). To be specific, firstly the data association $\hat{\mathbf{D}}_g^{k+1}$ is estimated based on the transformation matrix $\hat{\mathbf{T}}_\mathbf{M}^k$ of the last timestep k. Then, the estimation of $\hat{\mathbf{T}}_\mathbf{M}^{k+1}$ is updated based on $\hat{\mathbf{D}}_g^{k+1}$.

$$\hat{\mathbf{D}}_g^{k+1} = \underset{d_a}{\operatorname{argmax}} \prod_{a\in N_{\mathbf{M}_{r_1}}} \sum_{d_a\in\mathbf{D}_g} \underbrace{p(d_a|m_{r_1}^a,m_{r_n}^b,\hat{\mathbf{T}}_\mathbf{M}^k)}_{\text{Geometric Association}} \tag{10}$$

To measure the weight of the corresponding geometric pairs, we introduce $\omega_g = e^{-||m_{r_1}^a - f(\mathbf{T_M},m_{r_n}^b)||_{\Sigma_g}^2}$, which is closely related to the distance metric between $m_{r_1}^a$ and $m_{r_n}^b$.

b) Nonlinear Optimization: Based on the previously-established geometric data association, the nonlinear optimization problem can be solved by minimizing the error metric as:

$$\hat{\mathbf{T}}_\mathbf{M}^{k+1} = \underset{\mathbf{T_M}}{\operatorname{argmin}} \sum_a \sum_{d_a\in\mathbf{D}} \omega_g \cdot ||m_{r_1}^a - f(\mathbf{T_M},m_{r_n}^b,\hat{d}_a^{k+1})||_{\Sigma_g}^2 \tag{11}$$

where $||\cdot||_\Sigma^2$ denotes the Mahalanobis distance and Σ_g is the covariance of geometric features. $f(\mathbf{T_M},m_{r_n}^b,\hat{d}_a^{k+1})$ denotes the transformation function that transforms each point $m_{r_n}^b$, where $f(\mathbf{T_M},m_{r_n}^b,\hat{d}_a^{k+1}) = \mathbf{T_M}\cdot m_{r_n}^b = \mathbf{R_M}\cdot m_{r_n}^b + \mathbf{t_M}$. Thereafter, the nonlinear optimization problem of Eq. (11) is calculated by utilizing the Levenberg-Marquardt solver.

C. Global Map Merging

In the multi-session mapping framework, each single-session mapping process maintains and updates its corresponding map. After the estimation of the relative transformation matrix, next we fuse local magnetic maps $\mathbf{B}_{r_1}$ and $\mathbf{B}_{r_n}$ into a global magnetic map $\mathbb{B}_r$, and fuse local point cloud maps $\mathbf{M}_{r_1}$ and $\mathbf{M}_{r_n}$ into a global point cloud map $\mathbb{M}_r$. The latest incoming single-session magnetic map and the current magnetic map is denoted as $\mathbf{B}_{r_n}$ and $\mathbf{B}_r$, respectively. Given the magnetic data association $d_{i,j}$ and corresponding MF fingerprint pairs b_r^i, $b_{r_n}^j$, the corresponding pairs can be fused into the global MF fingerprints $\mathbb{B}_r^i$, which can be denoted as:

$$p(\mathbb{B}_r|\mathbf{D}_m,\mathbf{B}_r,\mathbf{B}_{r_n}) = \prod_{i\in N_{\mathbb{B}_r}} \sum_{d_i\in\mathbf{D}_m} p(\mathbb{B}_r^i|d_{i,j},b_r^i,b_{r_n}^j) \tag{12}$$

Similarly for the geometric data, we can obtain:

$$p(\mathbb{M}_r|\mathbf{D}_g,\mathbf{M}_r,\mathbf{M}_{r_n}) = \prod_{a\in N_{\mathbb{M}_r}} \sum_{d_a\in\mathbf{D}_g} p(\mathbb{M}_r^a|d_{a,b},m_r^a,m_{r_n}^b) \tag{13}$$

Considering that the distribution of each geometric point can be assumed as Gaussian, the objective of geometric point-wise fusion is to integrate two Gaussian distributions. We refer the readers to our previous work [38] for the details.

IV. Experimental Validations and Discussions

The robot platforms are shown in Fig. 4. All the experiments are executed on a laptop with an Intel i7-10875H CPU @2.3GHz with 32GB RAM. FAST-LIO2 [35] is applied for the single-session robot pose estimation. Since no available open source dataset which contains both the dense magnetic

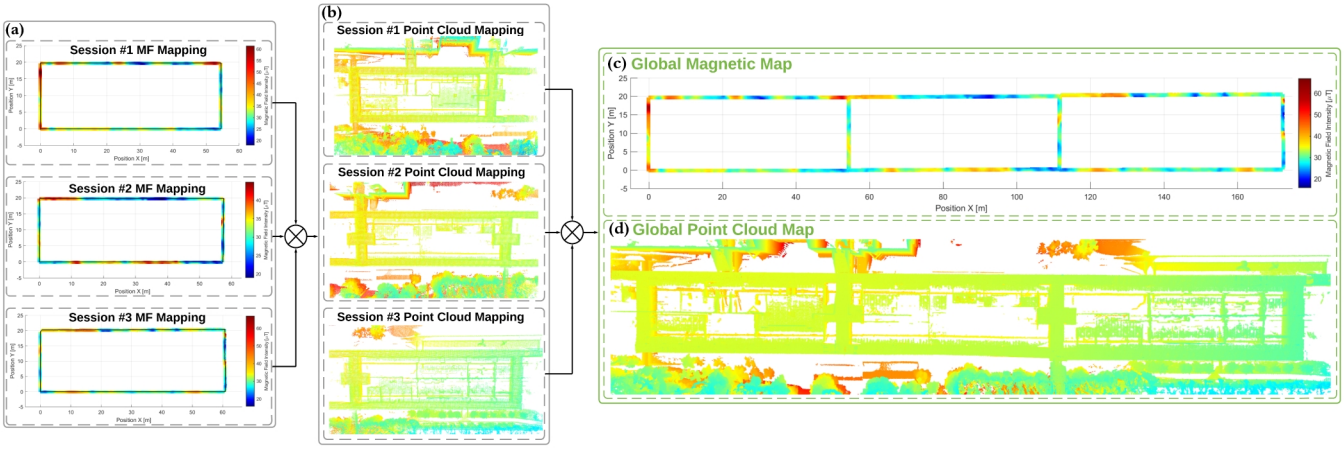
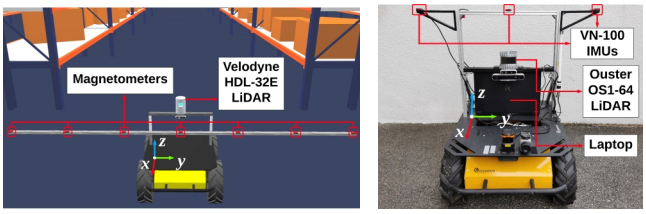


Fig. 3. Demonstration of the multi-session mapping for a real ambiguous corridor environment. (a) Single-session magnetic maps; (b) Single-session point cloud maps; (c) The global magnetic map built by multi-session mapping; (d) The globally consistent point cloud map built by multi-session mapping.

field and point cloud information, we conduct the experiments in two different ambiguous and perceptually-degraded environments as:

- **Simulated warehouse environment**: the operation area of the warehouse is $85m \times 75m$ in Gazebo simulator [39], which contains ambiguous goods shelves and steel concrete pillars, as shown in Fig. 1. The 3-D MF fingerprints collected in an open space area around the same size serves as the basis for the simulated MF data. Then, local MF variations such as steel concrete pillars and steel goods shelves were modeled according to [40] and were overlaid on the basic MF data.
- **Real corridor environment**: the operation area of the corridor is $173m \times 20m$, which contains polygonal pillars and ambiguous straight walls, as shown in Fig. 3. To collect MF data as efficient as possible, the robot was manually operated to run on pre-defined trajectories to cover traversable area of the corridor.

Comparison Baseline: The closely-related work is single robot SLAM or geometric map matching algorithms, where single robot SLAM aims to estimate the local odometry in a single-session process while map matching is focused on calculating the relative transformation matrix between different single-session maps. Due to the data heterogeneity, it is not reasonable and feasible to directly compare these two methods. As no available open source work has addressed the problem of magnetic map matching, the proposed framework is compared with two latest geometric map matching algorithms MFS-Match [38] and Octo-Match [41]. We also compare our framework with the well-recognized point cloud registration algorithm GICP [42] and NDT [43].



(a) Simulated robot platform (b) Real robot platform

Fig. 4. (a) The simulated Husky robot is equipped with seven magnetometers and a Velodyne HDL-32E LiDAR; (b) The real Husky A200 robot is equipped with three VN-100 IMUs and a Ouster OS1-64 LiDAR. The lateral distance between any two neighboring magnetometers/IMUs are empirically set at 0.5m to balance between the mapping accuracy and efficency [29].

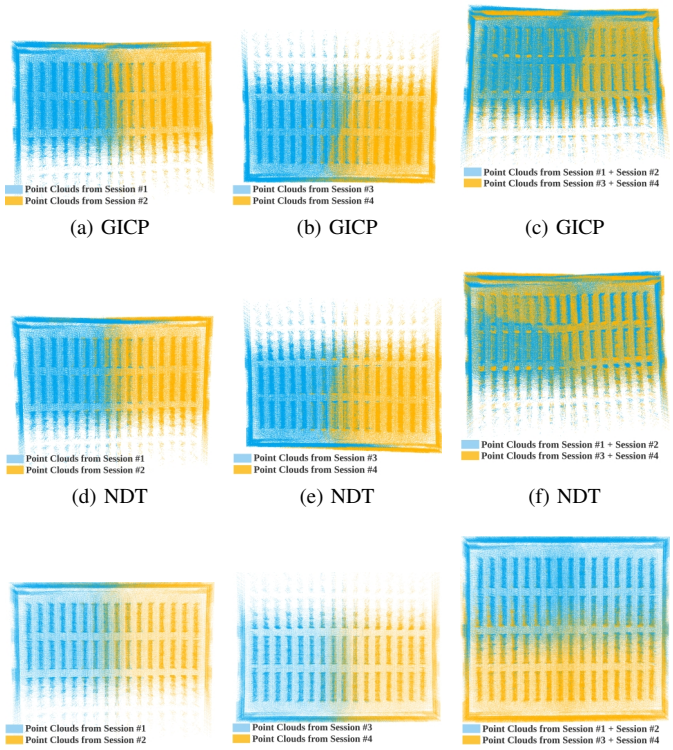
Evaluation Metrics: We compare the estimated transformation T_e to the ground-truth T_g. The error $\triangle T$ is calculated as $\triangle T = T_e \cdot T_g^{-1} = \{\triangle R, \triangle t\}$, where the 3-D rotation error $e_r = ||\triangle R|| = \arctan \triangle R$ and 3-D translation error $e_t = ||\triangle t||$ are the Euler norm and Euclidean norm of the difference between the output and the ground-truth. Moreover, it is worth pointing out that since the multi-session mapping is an offline process and the partial maps can be frequently updated. Thus, there is no related evaluations on the computational efficiency.

A. Mapping Results Visualization

To demonstrate the accuracy and utility of the proposed MM4MM system, the multi-session mapping results are verified by visualizing the final obtained global map. The experiments shown in Fig. 1 were done at a simulated enclosed industrial warehouse, where GNSS signal is not available. The four mapping sessions share the perceptions in some overlapping areas. The initial transformation among the multiple single-session maps is about $0.2m$ translation in x axis and 180° rotation in *yaw* angle. The experiments shown in Fig. 3 were done at a real long corridor, where GNSS signal is severely challenged. The three mapping sessions share the perceptions in some overlapping areas. The initial transformation among the multiple single-session maps is about $55m$ translation in x axis and 0° rotation in *yaw* angle. In both experiments, MM4MM is capable to infer the correct transformation information. As shown in Fig. 1(c), Fig. 3(c) and Fig. 3(d), the global magnetic map and point cloud map are both accurate and consistent, which demonstrate that the map matching process is successful.

B. Map Matching Illustrations

The point cloud map matching results of GICP, NDT, and MM4MM in the simulated warehouse and real corridor

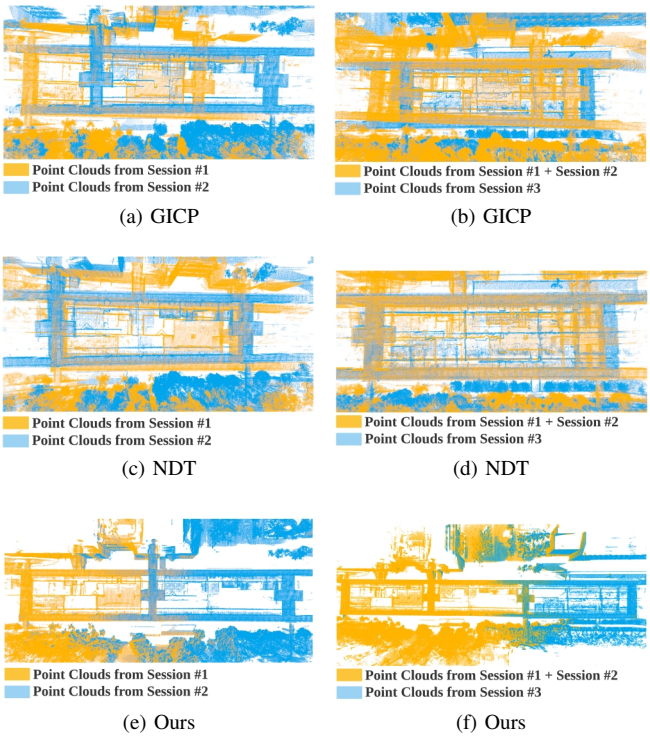


(a) GICP (b) GICP (c) GICP

(d) NDT (e) NDT (f) NDT

(g) MM4MM (h) MM4MM (i) MM4MM

Fig. 5. The point cloud map matching results of different methods in the simulated warehouse environment. (a), (d), and (g): session #1 matching with session #2; (b), (e), and (h): session #3 matching with session #4; (c), (f), and (i): session #1 + #2 matching with session #3 + #4.



(a) GICP (b) GICP

(c) NDT (d) NDT

(e) Ours (f) Ours

Fig. 6. The point cloud map matching results of different methods in the real corridor environment. (a), (c), and (e): session #1 matching with session #2; (b), (d), and (f): session #1 + #2 matching with session #3.

environments are shown in Fig. 5 and Fig. 6, respectively. It is clear that our proposed MM4MM system can successfully and accurately match and merge the maps from different sessions then finally obtain a globally consistent map, while GICP and NDT methods tend to mismatch them and lead to a mess of the merged map due to the ambiguous geometric data association.

C. Map Matching Accuracy

The average translation error and rotation error of the map matching experiments in simulated warehouse are summarized in Table. I. It is clear that Octo-Match algorithm has the worst performance due to that it only utilizes geometric coordinates as features. By only taking the ambiguous geometric features into consideration, it can be very hard to establish the correct correspondences in such challenging ambiguous environments. Thus for GICP, NDT, and MFS-Match algorithms, the estimation errors are large and all of them fail to converge to the global minimum. It is easy to see that the proposed MM4MM system has superior performance over the comparative methods. Overall, MM4MM outperforms the other comparative algorithms due to the full usage of magnetic information to establish the correct correspondences, which demonstrates the greatly improved matching accuracy and robustness.

TABLE I

ESTIMATED MEAN ERROR OF TRANSLATION AND ROTATION OF THE SIMULATED WAREHOUSE ENVIRONMENT

	Session # 1 + 2		Session # 3 + 4		Session # 1 + 2 + 3 + 4	
	Trans (m)	Rot (°)	Trans (m)	Rot (°)	Trans (m)	Rot (°)
GICP	1.513	0.595	0.793	0.300	2.769	182.610
NDT	**1.146**	0.647	0.231	0.253	2.713	180.350
MFS-Match	1.599	2.928	1.112	3.804	10.439	180.965
Octo-Match	11.340	1.744	11.599	2.796	9.184	179.743
MM4MM	1.162	**0.472**	**0.177**	**0.192**	**0.150**	**0.378**

V. CONCLUSION

This paper has presented a novel map matching framework named MM4MM, for multi-session mapping in challenging ambiguous and perceptually-degraded environments. The proposed MM4MM framework calculates relative transformation matrix of multiple single-session maps and builds globally consistent maps of the environment. Thanks to the introduction of the distinct magnetic field (MF) information, the proposed algorithm is still able to establish correct point-wise correspondences in the ambiguous environments. Extensive comparisons have been conducted on both simulated experiments and real practical missions, and the results have shown the improved accuracy and robustness over the comparative methods. In the future, we plan to further integrate the MF information into the autonomous robots active tasks, such as navigation and collaborations.

REFERENCES

[1] C. Cadena *et al.*, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, 2016.

[2] R. Giubilato *et al.*, "Relocalization with submaps: Multi-session mapping for planetary rovers equipped with stereo cameras," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 580–587, 2020.

[3] T.-M. Nguyen *et al.*, "VIRAL-Fusion: A visual-inertial-ranging-lidar sensor fusion approach," *IEEE Trans. Robot.*, vol. 38, no. 2, pp. 958–977, 2021.

[4] J. Zhang, H. Zhuge, Z. Wu, G. Peng, M. Wen, Y. Liu, and D. Wang, "4DRadarSLAM: A 4d imaging radar slam system for large-scale environments based on pose graph optimization," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*. IEEE, 2023, pp. 8333–8340.

[5] H. Shen, Q. Zong, B. Tian, X. Zhang, and H. Lu, "PGO-LIOM: Tightly coupled lidar-inertial odometry and mapping via parallel and gradient-free optimization," *IEEE Trans. Ind. Electron. (TIE)*, vol. 70, no. 11, pp. 11 453–11 463, 2023.

[6] J. Zhang *et al.*, "NTU4DRadLM: 4d radar-centric multi-modal dataset for localization and mapping," in *Proc. IEEE Int. Conf. Intell. Transp. Syst. (ITSC)*. IEEE, 2023, pp. 4291–4296.

[7] J. McDonald *et al.*, "Real-time 6-DOF multi-session visual SLAM over large-scale environments," *Robot. Autonom. Syst.*, vol. 61, no. 10, pp. 1144–1158, 2013.

[8] M. Labbe and F. Michaud, "Online global loop closure detection for large-scale multi-session graph-based SLAM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*. IEEE, 2014, pp. 2661–2666.

[9] S. Saeedi, M. Trentini, M. Seto, and H. Li, "Multiple-robot simultaneous localization and mapping: A review," *J. Field Robot.*, vol. 33, no. 1, pp. 3–46, 2016.

[10] H. Do, S. Hong, and J. Kim, "Robust loop closure method for multi-robot map fusion by integration of consistency and data similarity," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 5701–5708, 2020.

[11] S. Jin, Z. Wu, C. Zhao, J. Zhang, G. Peng, and D. Wang, "SectionKey: 3-D semantic point cloud descriptor for place recognition," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*. IEEE, 2022, pp. 9905–9910.

[12] G. Peng, Y. Yue, J. Zhang, Z. Wu, X. Tang, and D. Wang, "Semantic reinforced attention learning for visual place recognition," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*. IEEE, 2021, pp. 13 415–13 422.

[13] Y. Yue, C. Zhao, Z. Wu, C. Yang, Y. Wang, and D. Wang, "Collaborative semantic understanding and mapping framework for autonomous systems," *IEEE/ASME Trans. Mechatron.*, vol. 26, no. 2, pp. 978–989, 2020.

[14] Y. Yue, C. Zhao, M. Wen, Z. Wu, and D. Wang, "Collaborative semantic perception and relative localization based on map matching," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*. IEEE, 2020, pp. 6188–6193.

[15] H. Shen, Q. Zong, B. Tian, and H. Lu, "Voxel-based localization and mapping for multirobot system in gps-denied environments," *IEEE Trans. Ind. Electron. (TIE)*, vol. 69, no. 10, pp. 10 333–10 342, 2022.

[16] Z. Wu, M. Wen, G. Peng, X. Tang, and D. Wang, "Infrastructure-free global localization in repetitive environments: An overview," in *Proc. IECON 2020 - 46th Annu. Conf. of the IEEE Ind. Electron. Soc.* IEEE, 2020, pp. 626–631.

[17] X. Ding *et al.*, "Multi-session map construction in outdoor dynamic environment," in *Proc. IEEE Int. Conf. Real-time Comput. Robot. (RCAR)*. IEEE, 2018, pp. 384–389.

[18] T. M. Bonanni *et al.*, "3-D map merging on pose graphs," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1031–1038, April 2017.

[19] A. Petitti *et al.*, "A distributed map building approach for mobile robotic networks," in *Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE)*. IEEE, 2018, pp. 116–121.

[20] Y. Yue *et al.*, "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.

[21] B. Gozick *et al.*, "Magnetic maps for indoor navigation," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 12, pp. 3883–3891, 2011.

[22] B. Li *et al.*, "How feasible is the use of magnetic field alone for indoor positioning?" in *Proc. Int. Conf. Ind. Posi. Ind. Navi. (IPIN)*. IEEE, 2012, pp. 1–9.

[23] M. Angermann *et al.*, "Characterization of the indoor magnetic field for applications in localization and mapping," in *Proc. Int. Conf. Ind. Posi. Ind. Navi. (IPIN)*. IEEE, 2012.

[24] M. Frassl *et al.*, "Magnetic maps of indoor environments for precise localization of legged and non-legged locomotion," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2013, pp. 913–920.

[25] J. Jung *et al.*, "Indoor mobile robot localization and mapping based on ambient magnetic fields and aiding radio sources," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 7, pp. 1922–1934, 2015.

[26] N. Akai *et al.*, "Gaussian processes for magnetic map-based localization in large-scale indoor environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2015, pp. 4459–4464.

[27] Z. Wu *et al.*, "Magnetic-assisted initialization for infrastructure-free mobile robot localization," in *Proc. 2019 IEEE Int. Conf. on Cyb. and Intelli. Syst. (CIS) and IEEE Conf. on Robot., Autom. and Mechatron. (RAM)*. IEEE, 2019, pp. 518–523.

[28] S.-C. Yeh *et al.*, "Study on an indoor positioning system using earth's magnetic field," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 3, pp. 865–872, 2019.

[29] Z. Wu, Y. Yue, M. Wen, J. Zhang, J. Yi, and D. Wang, "Infrastructure-free hierarchical mobile robot global localization in repetitive environments," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[30] Z. Wu, Y. Yue, M. Wen, J. Zhang, G. Peng, and D. Wang, "MSTSL: Multi-sensor based two-step localization in geometrically symmetric environments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*. IEEE, 2021, pp. 5245–5251.

[31] T. Takebayashi *et al.*, "Development of magnetic-based navigation by constructing maps using machine learning for autonomous mobile robots in real environments," *Sensors*, vol. 21, no. 12, p. 3972, 2021.

[32] Z. Wu, W. Wang, J. Zhang, Q. Lyu, H. Zhang, and D. Wang, "Global localization in repetitive and ambiguous environments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*. IEEE, 2023, pp. 12 374–12 380.

[33] Z. Wu, W. Wang, H. Zhang, J. Zhang, Q. Lyu, and D. Wang, "Magnetic field-aided global localization in repetitive environments," in *Proc. 2023 IEEE Int. Conf. on Cyb. and Intelli. Syst. (CIS) and IEEE Conf. on Robot., Autom. and Mechatron. (RAM)*. IEEE, 2023, pp. 1–6.

[34] Y. Yue *et al.*, "COSEM: Collaborative semantic map matching framework for autonomous robots," *IEEE Trans. Ind. Electron.*, vol. 69, no. 4, pp. 3843–3853, 2021.

[35] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lio2: Fast direct lidar-inertial odometry," *IEEE Trans. Robot.*, vol. 38, no. 4, pp. 2053–2073, 2022.

[36] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Netw. (ICNN)*, vol. 4. IEEE, 1995, pp. 1942–1948.

[37] F. Dellaert, "The expectation maximization algorithm," Georgia Institute of Technology, Tech. Rep., 2002.

[38] Y. Yue *et al.*, "A multilevel fusion system for multirobot 3-D mapping using heterogeneous sensors," *IEEE Syst. J.*, vol. 14, no. 1, pp. 1341–1352, 2019.

[39] N. Koenig and A. Howard, "Design and use paradigms for Gazebo, an open-source multi-robot simulator," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*. IEEE, 2004, pp. 2149–2154.

[40] K. P. Subbu *et al.*, "LocateMe: Magnetic-fields-based indoor localization using smartphones," *ACM Trans. Intell. Syst. and Tech.*, vol. 4, no. 4, pp. 1–27, 2013.

[41] J. Jessup *et al.*, "Robust and efficient multirobot 3-D mapping merging with octree-based occupancy grids," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1723–1732, 2017.

[42] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proc. Robotics: Science and Systems (RSS)*, vol. 2, no. 4, 2009, p. 435.

[43] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3D-NDT," *J. Field Robot.*, vol. 24, no. 10, pp. 803–827, 2007.