This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

CrowdMagMap: Crowdsourcing based Magnetic Map Construction for Shopping Mall

Yan Wang, Jian Kuang, Tianyi Liu, Xiaoji Niu, and Jingnan Liu

***Abstract*—Indoor positioning is an important part of supporting the Internet of Things and location-based services. Crowdsourcing-based magnetic map construction is a key technology to realize wide-area consumer indoor positioning. However, current crowdsourcing-based magnetic map schemes are not suitable for typical indoor scenarios (e.g., shopping malls). The reason is that they ignore the characteristics of crowdsourced data, including short-term trajectory, various pedestrian motion patterns, large-scale dataset, and so on. In this paper, we propose a novel crowdsourcing-based magnetic map construction method. First, learning-based inertial odometry is used to recover precise user motion trajectories regardless of changes in motion patterns. Then, a keyframe-efficient association method of magnetic time-frequency features is proposed, which is suitable for short-term trajectories of various shapes. Finally, a two-step global estimation optimization is proposed to further eliminate false associations of keyframes and improve the robustness of the method. The feasibility of the proposed method is verified by using a multi-user dataset in a typical shopping mall scenario. The proposed method takes a total of 60.8 seconds to process a 12-hour dataset (sub-trajectories with a duration of 90s), and the average position error is 1.48m (with scale correction) and 2.53m (without scale correction). Compared with the existing crowdsourcing-based magnetic map scheme, the proposed method has been significantly improved in terms of feasibility, accuracy, and efficiency.**

***Index Terms*—Crowdsourcing, Magnetic Map, Keyframe Association, Pedestrain Navigation, Indoor Positioning**

I. Introduction

Indoor positioning and navigation is key technology in developing the Internet of Things (IoT) and location-based services (LBS) [1]. Various sensors, including Bluetooth low-energy (BLE) [2], WiFi [3] [4], magnetometer [5], and Inertial Measurement Unit (IMU) [6] [7] [8] have been adopted to provide accurate indoor positioning services for mass users. IMU-based methods [9] [10] is usually used as an auxiliary positioning technique because it can only provide accurate relative trajectories in a short period of time. The methods based on WIFI [11] [12], BLE [13], and magnetometer [14] [5] can achieve meter-level positioning relying on pre-built signal fingerprint maps (including radio frequency signal and magnetic field signal). Therefore, an indoor signal fingerprint map is the basic premise to ensure the availability of indoor positioning.

Following the principle of prioritizing signal fingerprint accuracy, traditional fingerprint mapping methods, including point-to-point and walking survey, require experts to use professional equipment (e.g., total station) to measure geographic coordinates and collect signal fingerprints according to pre-planned routes [15]. However, the traditional method is far from meeting the needs of large-scale indoor signal fingerprint construction due to time-consuming, laborious, and expensive [16], which is the main limitation for the promotion of wide-area indoor positioning services.

To reduce the cost of signal fingerprint mapping, crowdsourcing-based mapping methods have attracted the attention of many researchers [15] [17]. The core idea is to use the IMU-based methods to estimate the trajectories of each user, and use the similarity of signals or sensory landmarks to associate and fuse the trajectories of different users.

Walkie-Markie [18] correlates user trajectories by using unique landmarks formed at the moment of maximum RSSI, which can avoid the influence of multipath propagation of WiFi signals. Similarly, SoiCP [19] utilizes WiFi signals and building gates as landmarks for trajectory association. WiFi-RITA [16] proposes a robust iterative trace merging algorithm based on WiFi access points as signal markers, which can adapt to datasets with obvious noise, diverse motion patterns and short-period trajectories, and obtain significant efficiency improvements. [20] uses learning-based inertial odometry to offer relative trajectories and a WiFi bundle adjustment to recover trajectories. This approach is very resistant to different motion patterns and suitable for data collection by non-experts. However, due to the distribution density of the WiFi or Bluetooth access points (APs) is uncontrollable, the above methods will not work in areas with sparse distribution of wireless signals.

Infrastructure-free indoor positioning methods (e.g., magnetic matching) are the primary option for consumer-grade wide-area indoor positioning [14]. A crowdsourcing-based magnetic map generation approach independent of wireless signals is necessary. Compared with wireless signals, crowdsourcing map construction methods based on magnetic signals are more challenging because the signals do not have globally unique characteristics. Detail-wise, even if the magnetic field at a certain location in the two trajectories is almost identical, it is not guaranteed to be absolutely correct. [21] and [22] assume that the trajectory of a single user is composed of straight lines and corners, extract the magnetic field sequences corresponding to the corners as keyframes, and use the cluster-

This work was supported by the Fundamental Research Funds for the Central Universities (2042023kf0124). (Corresponding author: Jian Kuang)

X. Niu is with the GNSS Research Center, Wuhan University, and Hubei Luojia Laboratory, Wuhan, Hubei, CO 430072 PR China. (e-mail:xjniu@whu.edu.cn)

Jingnan Liu, J. Kuang, Y. Wang, and T. Liu are with the GNSS Research Center, Wuhan University, Wuhan, Hubei, CO 430072 PR China (e-mails: jnliu@whu.edu.cn;kuang@whu.edu.cn; wystephen@whu.edu.cn;liutianyi@whu.edu.cn)

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

ing method to associate them. Dynamic Time Warping (DTW) methods are used to measure the similarity of the keyframes in clustering methods. Based on the same assumptions and ideas, [23] proposes a step-based magnetic sequence similarity estimation method to replace DTW, which has significantly improved computational efficiency. However, these methods still have two issues for challenging application scenarios: 1) User trajectories in complex indoor environments (e.g., shopping malls) are irregular shapes instead of straight lines and corners, which violate the basic assumptions of these methods. 2) The computational efficiency is low, and it is not suitable for large-scale dataset in practical applications. The reason is that they use time-consuming clustering methods [24] [25] to associate user trajectories, including k-means, affinity propagation, and hidden Markov models.

To achieve wide-area indoor positioning for consumer applications, crowdsourcing-based magnetic field mapping schemes for typical indoor scenarios (e.g., shopping malls) need to be further explored. This paper summarizes the problems faced by crowdsourcing methods in typical public scenarios (e.g. shopping malls) in the II-A section, and proposes a novel crowdsourcing-based magnetic map construction solution. Compared with existing solutions, the major contributions of this paper are summarized as follows:

- We propose a novel crowdsourcing-based magnetic map construction method that can accommodate short-period trajectories with irregular shapes, poor sensor data (e.g. magnetometer with uncalibrated bias), and various pedestrian motion patterns.
- To deal with the huge time consumption of keyframe association, we propose a fast keyframe association method in which time consumption increases slowly with the size of the dataset. Benefiting from this design, the keyframe association obtained a 12x speed up in the shopping mall test scenario.
- To eliminates the influence of erroneously related keyframe pair, we design a two-step graph-optimization-based method for merging trajectories resistant to erroneous keyframe association. This approach eliminates erroneously related keyframe pairs by utilizing the magnetometer data, the magnetic field distribution, and the trajectories' geometry.

The remainder of the paper is organized as follows. Section II summarizes the problems faced by the crowdsourcing-based map construction methods and gives an overview of the proposed method. Section III describes the learning-based inertial odometry. Section IV describes the proposed keyframe association method. Section V provides a detailed description of the graph-optimization-based trajectory merge method. Section VI uses field tests to prove the feasibility and evaluate the performance of the proposed method.

II. Problem Statement and System Overview

A. Problem Statement of crowdsourcing-based magnetic map construction

The challenges and requirements for crowdsourcing-based magnetic map construction in harsh scenarios are listed as follows.

Short-period Trajectory: Regarding real-world crowdsourcing data collection using smartphone apps, one major limitation is the device's battery capacity. As a result, the app cannot run in the background for extended periods. The data collected through crowdsourcing is restricted to trajectories that only last 1-3 minutes. Moreover, magnetic signals have only three-dimensional independent components. Thus, the reliability of sub-trajectories association methods based on magnetic signal will be seriously degraded.

Various Motion Patterns: Since the crowdsourced data is generated when the user uses the mobile phone in a natural state, the data contains various motion patterns. However, conventional IMU-based pedestrian dead reckoning (PDR) cannot accurately determine the walking trajectory of pedestrians when they use smartphones in various motion patterns, such as texting, calling, and swaying [26] [27]. Although it is feasible to filter out simple motion pattern sensor data to generate a magnetic map, the data utilization rate is extremely low, and the cycle of providing positioning services will be significantly prolonged. Therefore, an accurate and robust trajectory estimation method that can adapt to complex pedestrian motion patterns is an essential property of crowdsourcing-based magnetic map generation methods.

Uncalibrated Sensors: The smartphone's built-in sensors are of low quality, so the sensor measurements usually have large bias errors [28]. In particular, the magnetometer bias will change significantly under the influence of the electromagnetic effect caused by the change in the working state of the smartphone. Then, because the magnetometer observations cannot reflect the real environmental magnetic field, the user trajectory association methods based on the assumption of a consistent environmental magnetic field at the same location will become unavailable. Meanwhile, it is unrealistic to assume that the user will correctly calibrate the magnetometer biases in the user-insensitive mode. Therefore, the proposed method should be able to somehow calibrate the sensor online using the acquired data.

Irregular Trajectory: Restricted by the indoor space structure, the shape of the user's motion trajectory is complex and diverse, including straight lines, right-angle turns, irregular arcs, and so on. The existing crowdsourcing-based magnetic map generation methods almost all assume that the user's trajectory consists of straight lines and right-angle turns, which limits its application in complex indoor environments. Therefore, it is unreasonable to make too many assumptions about the shape of the user's walking trajectory for dealing with realistic user motion habits.

Large-scale Dataset: The necessity to manage massive datasets is a requirement of crowdsourced mapping for the reasons listed below. Firstly, a short-term trajectory carries less information, so the probability that any two trajectories can be uniquely associated is low, and the area covered by accurately associated two trajectories is small. Secondly, it is difficult to ensure that the gathered trajectories are spread evenly in space. To guarantee that certain locations that are infrequently visited by people may also be effectively recreated, we must collect a huge quantity of data to ensure that these regions also have

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

TABLE I
PERFORMANCE COMPARISON OF RELATED WORKS.

	Information Source	Short-period Trajectory	Various Motion Patterns	Uncalibrated Sensors	Irregular Trajectory	Large-scale Dataset
WiFi-RITA [16]	WiFi, IMU	Very Good	Good	Very Good	Very Good	Good
[20]	WiFi, IMU	Not Good	Very Good	Very Good	Very Good	Good
[21]	IMU, Mag	Not Good	Not Good	Not Good	Not Good	Not Good
[22]	IMU, Mag	Very Good	Not Good	Unknown	Not Good	Not Good
[23]	IMU, Mag	Good	Not Good	Unknown	Not Good	Not Good
CrowdMagMap (proposed)	IMU, Mag	Very Good	Very Good	Very Good	Very Good	Good

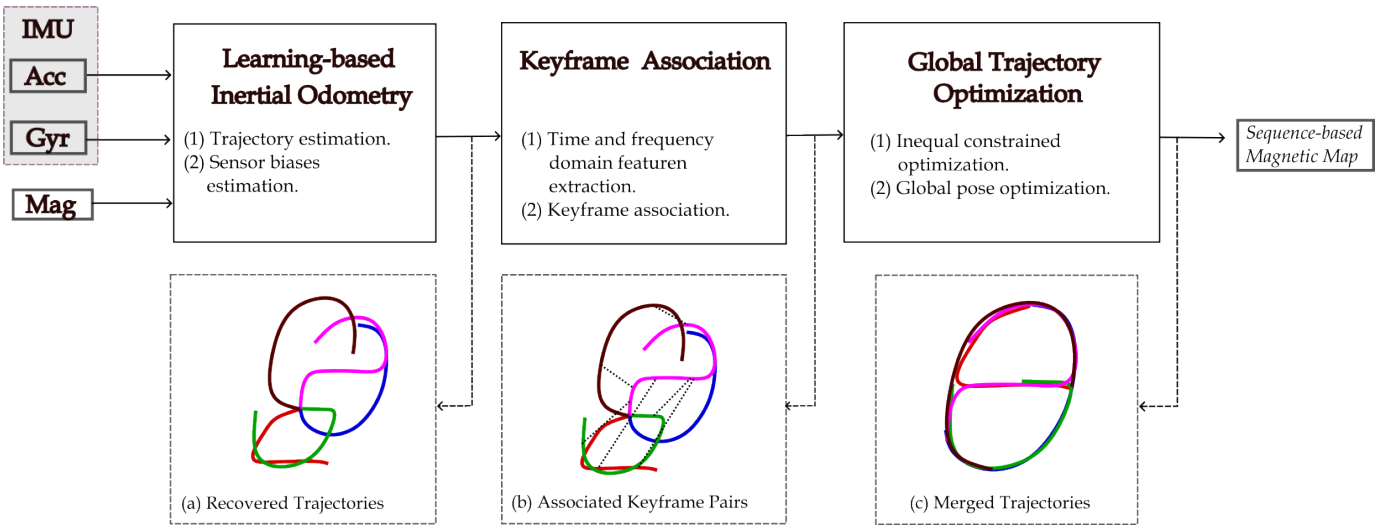

Fig. 1. Data process flow of the proposed solution.

enough data for generating a magnetic map.

B. System Overview

The algorithm flow of the proposed method is shown in Figure 1. The crowdsourcing-based magnetic map construction method can be divided into three stages: learning-based inertial odometry (LIO), keyframe association, and global trajectory optimization.

1) Learning-based Inertial Odometry: The magnetometer-enhanced learning-based inertial odometry method proposed by our previous work is employed to reconstruct each short-period trajectory by using IMU and magnetometer measurements. Compared with the traditional PDR method, LIO can estimate complex pedestrian motion trajectories more accurately. To obtain accurate environmental magnetic observations, the magnetometer bias is modeled and estimated concurrently with the smartphone's pose.

2) Keyframe Association: We extract keyframes from each trajectory according to predetermined distance interval and use the similarity of the magnetic field feature to associate keyframes. Compared to keyframes based on trajectory shape (such as corner and corridor), the proposed method has the benefit of being free from specific trajectory shapes. However, the calculation process of magnetic similarity is very time-consuming, especially when dealing with large crowdsourced datasets. To overcome this drawback, we present a method for efficient keyframe association, which gives the whole system a high level of efficiency under good adaptability.

3) Global Trajectory Optimization: We propose a two-step optimization method to estimate the positions of all keyframes. In the first step, the optimization problem is defined as a nonlinear least square problem with inequality constraints. The relative poses of keyframes within the same trajectory are fixed, and the optimal parameters are the 2D location and direction of the initial keyframe of each trajectory. Moreover, the heading of the first keyframe is constrained by an inequality equation. This stage establishes a starting location for each keyframe to expedite further optimization and eliminate certain falsely related keyframe pairs. In the second step, a comprehensive optimization of the pose graph is implemented.

After the above three stages, a density-based technique is employed to exclude outlier trajectories. Then, the pose of the estimated trajectory and the corresponding magnetic field features with bias compensation and coordinate transform are stored to generate a magnetic field map for sequence-based matching positioning.

In conclusion, the proposed method not relies on assumptions about the trajectory's shape and can build a map using short trajectories. The proposed method adapts learning-based inertial odometry to various pedestrian motion patterns and uncalibrated sensors. Then, a keyframe association method combining time-frequency features is proposed, which can achieve high-efficiency trajectory association of large-scale datasets even with irregular trajectories. Finally, the global graph optimization method is used to exploit the spatial constraints between trajectories to improve the phenomenon

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

that short-period trajectory reduces the accuracy of trajectory association.

III. Learning-based Inertial Odometry

The graph optimization enhanced LIO proposed by our previous work [28] [10] can adapt to various motion modes and is an ideal trajectory estimation method. Unlike the setting suggested by [28], this paper adopts global optimization instead of sliding window optimization because the data post-processing mode is available. To reduce the time consumption, we reduce the number of iterations by providing optimal initial values for graph optimization. The initial value consists of two parts, the magnetometer biases and the orientation of each instant.

The initial value of magnetometer bias is obtained by using the constraint that the magnetic vector is constantly combined with the relative rotation from integrating gyroscope observations. Assuming that the magnetic field vector remains unchanged in a short period of time (e.g., 2 seconds), the relationship between magnetometer observations, magnetometer bias, and relative rotation can be described as follows:

$$\Delta \boldsymbol{R}_{ij}(\boldsymbol{m}_j - \boldsymbol{b}_m) = (\boldsymbol{m}_i - \boldsymbol{b}_m) \tag{1}$$

where $\Delta \boldsymbol{R}_{ij}$ represents rotation from moment j to i calculated by integrating gyroscope measurements, $\boldsymbol{m}_i$ and $\boldsymbol{m}_j$ are magnetometer measurements at i and j, respectively, $\boldsymbol{b}_m$ is magnetometer bias. To ensure that Eq.(1) is solvable, we select data of short periods with significant rotation around x, y, and z-axis of the magnetometer, respectively.

Then, a typical attitude and heading system (AHRS) approach is used to estimate the initial value of orientation at each instant based on the constraints of the gravity vector and magnetic field vector. To align the headings of all trajectories, a graph-optimization problem is defined. The system state of trajectory k is defined as follows:

$$\boldsymbol{X}_k = [\boldsymbol{R}_{nb_0} ... \boldsymbol{R}_{nb_i}, \boldsymbol{b}_m^k] \tag{2}$$

where $\boldsymbol{R}_{nb_i}$ represents rotation from i-th body frame to the navigation frame. $\boldsymbol{b}_m^k$ is the magnetometer biases of trajectory k. Due to the length of each trajectory being short (no more than 2 minutes), it is feasible to assume that the magnetometer biases are constant.

The maximum posterior estimation of all system states is obtained by minimizing the following cost function:

$$\begin{aligned} \{\boldsymbol{X}_0, ... \boldsymbol{X}_k, \boldsymbol{m}^n\} = \\ \underset{\{\boldsymbol{X}_0, ... \boldsymbol{X}_k, \boldsymbol{m}^n\}}{\arg\min} \{ \|\boldsymbol{r}_{prior}\|^2 + \sum \|\boldsymbol{r}_{mag}\|^2_{\Sigma_{mag}} \\ \sum \|\boldsymbol{r}_{gyr}\|^2_{\Sigma_{gyr}} + \sum \|\boldsymbol{r}_{gra}\|^2_{\Sigma_{gra}} \} \end{aligned} \tag{3}$$

where $\boldsymbol{m}^n$ is the magnetic field vector in the navigation frame. $\boldsymbol{r}_{prior}$ represents the prior constraint to $\boldsymbol{m}^n$. Because we assume the x-axis and y-axis of the navigation frame are orientated to the magnetic north and east, the y-axis of the $\boldsymbol{m}^n$ should be close to zero. $\boldsymbol{r}_{gyr}$ is the relative rotation constraint between adjacent instants provided by the gyroscope measurements. $\boldsymbol{r}_{gra}$ is the gravity orientation constraint as described in [28]. $\boldsymbol{r}_{mag}$ is the magnetic field constraint which is defined as:

$$\boldsymbol{r}_{mag} = \boldsymbol{R}_{nb_i}(\boldsymbol{m}_i^k - \boldsymbol{b}_m^k) - \boldsymbol{m}^n \tag{4}$$

where $\boldsymbol{m}_i^k$ is magnetometer measurements in trajectory k at i-th moment.

Solving Eq. (3), we obtain the orientation of each instant and magnetometer biases of each trajectory. Benefiting from these global consistency and accuracy rotations of each moment, we estimate the shape of each trajectory represented in the navigation frame based on the neural network discussed in [10]. This method is robust to various motion patterns and can provide a consistency scale that is superior to conventional PDR methods. These advantages play a key role in the following processing.

IV. Keyframe Association

This section introduces the proposed fast keyframe association method. The feature extraction and similarity calculation method are described in Section IV-A. Section IV-B describes a two-step keyframe association method, which uses the frequency-domain feature to achieve fast candidate keyframe pair searching and confirm the candidate keyframe pair based on the time-domain feature and geometric consistency.

A. Feature Extraction and Feature Similarity Calculation

1) Preprocess of magnetic field sequence: The availability of magnetic field sequence-based keyframe association identification relies on the fact that the magnetic field vector at a certain point is constant across time but fluctuates with the position. To obtain the characteristics of the magnetic field feature changing with the spatial position, we need to perform data preprocessing on the original observations of the magnetometer, including bias compensation, coordinate transformation and resampling according to the spatial distance.

The raw magnetometer measurements are corrected by estimated magnetometer biases of Section III and then converted from the sensor frame to the navigation frame using the following equation:

$$\boldsymbol{m}_i^n = \boldsymbol{R}_{nb_i}(\boldsymbol{m}_i^k - \boldsymbol{b}_m^k) \tag{5}$$

Then, the magnetic field sequence can be converted to a function of moving distance using the estimated trajectory by LIO as discussed in Section III. Benefiting from this conversion, the moving speed no longer influences the magnetic field sequence. Thus, we can avoid using DTW-like methods to compare two magnetic field sequences.

This step outputs the magnetic field vector sequence denoted as $\boldsymbol{m}_i^k = [\boldsymbol{m}_{xi}^k, \boldsymbol{m}_{yi}^k, \boldsymbol{m}_{zi}^k]^T$, which is 256×3 matrix. $\boldsymbol{m}_i^k$ obtained by equidistant sampling on the i-th keyframe (the magnetic field sequence within a distance of 10 meters is 1 keyframe). $\boldsymbol{m}_{xi}^k$, $\boldsymbol{m}_{yi}^k$, and $\boldsymbol{m}_{zi}^k$ are three-axis components of $\boldsymbol{m}_i^k$ respectively.

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

2) Frequency-domain Feature: The frequency-domain feature can be obtained by using the Fast Fourier Transform (FFT) for each axis sequence of $\boldsymbol{m}_i^k$. Details of the frequency-domain feature for magnetic field sequence can be found in [29]. It is defined as:

$$^F\mathcal{F}_i^k = [FFT(\boldsymbol{m}_{yi}^k)^T, FFT(\boldsymbol{m}_{zi}^k)^T, FFT(\boldsymbol{m}_{xi}^k)^T]^T \quad (6)$$

where $^F\mathcal{F}_i^k$ is frequency-domain feature of keyframe i in the trajectory k. $FFT(\cdot)$ represents the fast Fourier transform function, and the result is mapping the input sequence in the frequency domain. We only save the first eight components of the $FFT(\cdot)$ to speed up distance calculation. This feature vector definition causes significant information loss [29] but is helpful for fast feature similarity comparison.

The distance of two frequency-domain features is defined as:

$$^F\mathcal{D}(^F\mathcal{F}_i, ^F\mathcal{F}_j) = \|^F\mathcal{F}_i - ^F\mathcal{F}_j\|_2 \quad (7)$$

where $\|\cdot\|_2$ is $l2$-norm. Since the distance of two frequency-domain features is defined as $l2$-norm, the k-dimensional tree (KD Tree)-based method can be adopted for fast similar feature finding.

3) Time-domain Feature: The time-domain feature is defined as follows:

$$^T\mathcal{F}_i^k = \boldsymbol{m}_i^k \quad (8)$$

The main advantage of this feature is without additional transform, indicating that this processing is without information loss. The disadvantage of this feature is that the distance metric is hard to define as $l2$-norm. Because the small displacement of two sub-trajectories may cause large Euclidean distances in feature space which are not expected.

To avoid the disadvantage above, we defined the distance between two raw features as:

$$^T\mathcal{D}(^T\mathcal{F}_i, ^T\mathcal{F}_j)) = \frac{1}{N}\mathcal{S}(^T\mathcal{F}_i, ^T\mathcal{F}_j) \quad (9)$$

where $\mathcal{S}(\cdot)$ is a function that uses the maximum value of three axes to align two features and estimate distance. Three potential shift distances can be obtained by using the maximum value of three sequences (x, y, and z-axis) to align two time-domain features. Furthermore, we should reverse the sequences if users pass the same route with opposing directions. Figure 2 shows the six possible alignment ways. For example, considering the x-axis can obtain two possible alignments way, the 3-axis of $^T\mathcal{F}_j$ and flipped $^T\mathcal{F}_j$ are shifted to align the maximum value of the x-axis with $^T\mathcal{F}_i$'s. The feature distance of each alignment way is calculated based on the overlap parts. The final feature distance between two features is the minimum one in these six possible pairs of matched sequences. This distance metric considers the minor displacement and opposite movement direction of two trajectories with no loss of information. Using the premise that the moving distance of each trajectory is precise, it achieves more time efficiency than DTW-based sequence distance. This hypothesis is fulfilled by adopting the LIO.

In summary, the proposed keyframe association method achieves a compromise between efficiency and accuracy. First, a rapid candidate keyframe association is conducted by using low-accuracy and low-efficiency techniques. Next, an accurate but computationally expensive method is employed to identify relevant keyframe pairs in this candidate set, which achieves a significant reduction in computational demands compared to the entire dataset.

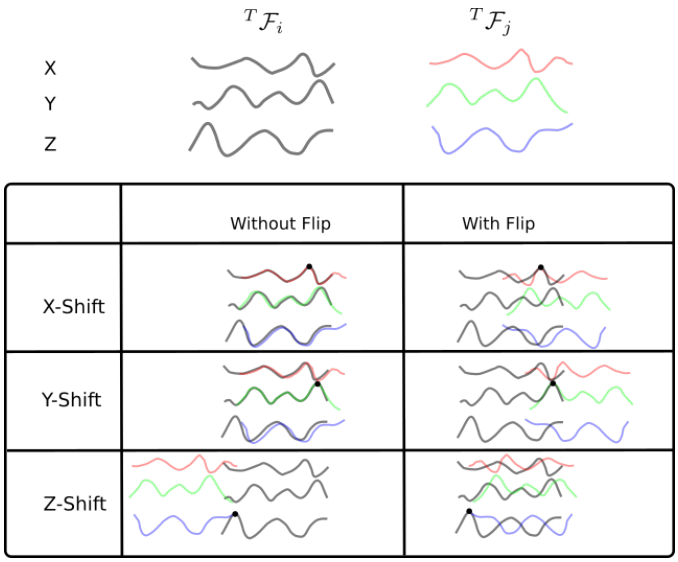

Fig. 2. 6 possible alignments for $^T\mathcal{D}(^T\mathcal{F}_i, ^T\mathcal{F}_j)$ calculation. Black points represent points used to alignment two sequences.

B. Keyframe Pair Determination

The conventional magnetic sequences-based keyframe association method needs to calculate the similarity between each keyframe pair. For N keyframes, the conventional method needs to calculate similarity $N * N/2$ times. This time complexity is not acceptable when processing large-scale datasets. To overcome this problem, the proposed method uses a k-d tree based on the frequency-domain feature to find similar magnetic sequences roughly. Then it uses a time-domain feature to determine a few candidate keyframe pairs accurately. Since the k-d tree-based roughly searching avoids calculating most of the keyframe pair similarity and its computation complexity is significantly lower than brute force searching on large-scale datasets, this operation can significantly save time. It is important to note that, as mentioned before, in order to utilize the k-d tree for faster keyframe association, the selected feature must be capable of calculating similarity using the $l2$-norm, and also be resilient to misalignment when using the $l2$-norm. In detail, when two trajectories pass the same path, the similarity metric based on $l2$-norm must not significantly affect by the slight difference in the sequence start point. The frequency-domain feature fulfills this condition, but the time-domain cannot. The time-domain feature must align two sequences first and then evaluate the similarity using $l2$-norm between the two aligned sequences.

The time complexity of building the k-d tree corresponding to the frequency-domain feature of all keyframes is $\mathcal{O}(nlog(n))$, n is the number of samples in the dataset. Based on the k-d tree data structure, the candidate keyframe pairs of any keyframe are quickly determined by using the

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

criterion that the frequency feature distance is less than a preset threshold $\mathcal{T}_F$.

Due to the loss of information in the frequency-domain features, there are many misjudgments in the candidate keyframe pairs association, so the time-domain features and geometric consistency are used to further verify the correctness of the candidate keyframes. Firstly, the candidate keyframe pair whose ${}^T\mathcal{D}$ is less than threshold $\mathcal{T}_T$ is confirmed. The displacement and whether the sequences are reversed are also noted. In particular, we may determine the one-to-one correspondence of each sample point of the magnetic field sequence based on the information on this displacement and whether or not to reverse the sequence. After determining the one-to-one correspondence of the magnetic field sequence, the one-to-one correspondence of the two sub-trajectories may be determined. The one-to-one relationship between the sub-trajectories' points can be utilized for further geometric consistency checks.

According to the correspondence provided before, the metric of geometric consistency can be defined as:

$$\mathcal{D}_{geo} = \sum_n \|\boldsymbol{p}_n^i - \boldsymbol{T}_{ij}\boldsymbol{p}_n^j\|_2 \tag{10}$$

$\boldsymbol{p}_n^i$ and $\boldsymbol{p}_n^j$ are correspondence 2D points in trajectory i and j, respectively. $\boldsymbol{T}_{ij}$ is a 2D transformation matrix that fulfills the following equation:

$$\begin{aligned} \boldsymbol{T}_{ij} = \quad & \underset{\boldsymbol{T}_{ij}}{\arg\min} \sum_n \|\boldsymbol{p}_n^i - \boldsymbol{T}_{ij}\boldsymbol{p}_n^j\|_2 \\ s.t. \quad & \boldsymbol{\theta}_{lowbound}^{\boldsymbol{T}_{ij}} < \boldsymbol{\theta}^{\boldsymbol{T}_{ij}} < \boldsymbol{\theta}_{upperbound}^{\boldsymbol{T}_{ij}} \end{aligned} \tag{11}$$

where $\boldsymbol{\theta}^{\boldsymbol{T}_{ij}}$ is rotation component of $\boldsymbol{T}_{ij}$. The associated keyframe pair, which fulfills that $\mathcal{D}_{geo}$ less than threshold $\mathcal{T}_{geo}$ passes the geometric consistency confirmation.

This step produces a pair of associated keyframes, which shows that corresponding keyframes travel through the same location and have a similar trajectory shape. The result of the keyframe pair association is then applied to merge all trajectories.

V. Global Trajectory Optimization

Global trajectory optimization exploits the relative pose constraints between adjacent keyframes in the same trajectory and the same-position constraints of associated keyframe pairs to merge, optimize, and filter global trajectories. In this paper, we proposed a two-step graph optimization method including the inequality constraint optimization and the pose graph optimization. Section V-A describes inequality constraints for preliminary trajectory merging and higher-level geometric information to verify the correctness of the associated keyframe pair. Section V-B describes the merging and shape optimization of trajectories through global bundle adjustment.

A. Inequality Constraint Optimization

Assuming that the relative poses of all keyframes in the same trajectory are accurate, then the errors of all keyframes are consistent. The optimization parameters of one sub-trajectory can be simplified as position $\boldsymbol{p}^k = [x^k, y^k]^T$ and heading $\boldsymbol{\theta}^k$. This operation helps for adopting the sub-trajectory shape to determine incorrect keyframe association. To avoid the phenomenon of small-area circles in a single sub-trajectory destroying the assumption that the magnetic heading disturbance conforms to a zero-mean Gaussian distribution [28], the assumption that the difference in direction between the mean magnetic field vector of sub-trajectories and the global mean magnetic field vector is smaller than a specific threshold is used to construct inequality constraints. This operation helps for adopting the sub-trajectory heading to determine incorrect keyframe association. Due to the robust kernel function being adopted to the distance constraint, the effect of the distance constraints of incorrectly associated keyframe pairs is limited in the optimization problem. Benefiting from properly using the LIO outputs, including the sub-trajectory shape and heading, and the robust kernel function for distance constraints, the inequality constraint optimization is robust to incorrect keyframe association. According to the optimized trajectory poses, the incorrectly associated keyframes can be determined and removed.

The inequality-constrained trajectory merging problem can be defined as:

$$\begin{aligned} \{\boldsymbol{p}^k, \boldsymbol{\theta}^k\} \quad & = \underset{\{\boldsymbol{p}^k, \boldsymbol{\theta}^k\}}{\arg\min} \sum \boldsymbol{\rho}(\|\boldsymbol{r}_{dis}\|^2_{\Sigma_{dis}}) \\ s.t. \quad & \boldsymbol{\theta}_{lowbound}^k < \boldsymbol{\theta}^k < \boldsymbol{\theta}_{upperbound}^k \end{aligned} \tag{12}$$

where

$$\boldsymbol{\rho}(s) = \begin{cases} 1 & s \geq 1 \\ 2\sqrt{s} - 1 & s < 1 \end{cases} \tag{13}$$

$$\boldsymbol{r}_{dis} = (\boldsymbol{R}(\boldsymbol{\theta}^k)[\boldsymbol{t}_{l_k b_j}]_{xy} + \boldsymbol{p}^k) - (\boldsymbol{R}(\boldsymbol{\theta}^q)[\boldsymbol{t}_{l_q b_i}]_{xy} + \boldsymbol{p}^q) \tag{14}$$

$$\boldsymbol{R}(\boldsymbol{\theta}) = \begin{bmatrix} cos(\boldsymbol{\theta}) & -sin(\boldsymbol{\theta}) \\ sin(\boldsymbol{\theta}) & cos(\boldsymbol{\theta}) \end{bmatrix} \tag{15}$$

$\|\cdot\|^2_\Sigma$ represents the Mahalanobia norm, $\boldsymbol{\rho}(\cdot)$ represents the Huber norm [30], $\boldsymbol{r}_{dis}$ represents the distance between the corresponding keyframes, $[\boldsymbol{t}_{l_k b_j}]_{xy}$ represents x and y components of $\boldsymbol{t}_{l_k b_j}$ which is position of keyframe j in the coordinate l_k, $\boldsymbol{R}(\boldsymbol{\theta})$ is the 2D rotation matrix of $\boldsymbol{\theta}$. $\theta_{lowbound}^k$ and $\theta_{upperbound}^k$ are set as -20 and 20 degree separately based on empirical. Eq. (12) defines a nonlinear least-squares problem with inequality constraint. We utilize ceres solver [31] to solve this problem. This problem can converge quickly because the state space is small, $3K$-dimensional (K is the number of trajectories).

The optimized pose of each sub-trajectory is used to transform all keyframes in the same coordinate frame (denoted as n-frame). The pose of keyframe i in the trajectory k denoted as $\{R_{nb_i}^k, t_{nb_i}^k\}$. Based on the estimated poses, we remove the false keyframe pairs that the distance between a keyframe pair is larger than 10 meters.

B. Pose Graph Optimization

This part adopts the global pose optimization to fuse all the constraints obtained in the previous part, so as to obtain the

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

optimal pose of each keyframe. The problem can be defined as:

$$\begin{aligned}\{\boldsymbol{R}^k_{nb_i}, \boldsymbol{t}^k_{nb_i}\} = \\ \underset{\{\boldsymbol{R}^k_{nb_i}, \boldsymbol{t}^k_{nb_i}\}}{\arg\min} \{ \sum \boldsymbol{\rho}(\|\boldsymbol{r}_{odo}\|^2_{\Sigma_{odo}}) + \sum \boldsymbol{\rho}(\|\boldsymbol{r}_{dis}\|^2_{\Sigma_{dis}}) \\ + \sum \|\boldsymbol{r}_{gra}\|^2_{\Sigma_{gra}} + \sum \|\boldsymbol{r}_{\overline{mag}}\|^2_{\Sigma_{\overline{mag}}}\}\end{aligned} \tag{16}$$

$\boldsymbol{r}_{odo}$ represents the relative pose residual between two adjacent keyframes,

$$\begin{aligned}\boldsymbol{r}_{odo}(\{\boldsymbol{R}_{nb_i}, \boldsymbol{t}_{nb_i}\}, \{\boldsymbol{R}_{nb_{i+1}}, \boldsymbol{t}_{nb_{i+1}}\}) = \\ \begin{bmatrix} \boldsymbol{R}^T_{nb_i}(t_{nb_{i+1}} - t_{nb_i}) - \Delta\boldsymbol{t} \\ Log_{SO(3)}(\Delta\boldsymbol{R}^T(\boldsymbol{R}^T_{nb_i}\boldsymbol{R}_{nb_{i+1}})) \end{bmatrix}\end{aligned} \tag{17}$$

$\Delta\boldsymbol{R}$ and $\Delta\boldsymbol{t}$ represents the relative rotation and translation provided by Sec III, $Log_{SO(3)}(\cdot)$ represents the logarithm function for $SO(3)$.

$\boldsymbol{r}_{dis}$ represents the relative distance residual between two keyframes in different sub-trajectories.

$$\boldsymbol{r}_{dis}(\boldsymbol{t}^k_{nb_i}\boldsymbol{t}^q_{nb_j}) = [\boldsymbol{t}^k_{nb_i} - \boldsymbol{t}^q_{nb_j}]_{xy} \tag{18}$$

$[\cdot]_{xy}$ represents x and y components of the 3D vector. i and j represent the keyframe index, k and q represent the sub-trajectory index.

$\boldsymbol{r}_{gra}$ represents the constraint of gravity orientation. It enforces the estimated map alignment to the gravity vector.

$$\boldsymbol{r}_{gra}(\boldsymbol{R}_{nb_i}) = \boldsymbol{R}_{nb_i}\bar{\boldsymbol{g}}^{b_i} - \boldsymbol{g}^n \tag{19}$$

$\bar{\boldsymbol{g}}^{b_i}$ and $\boldsymbol{g}^n$ represent gravity vectors in the local frame and the navigation frame, respectively. Since the pose of smartphones related to the pedestrian may vary during data collection, we adopt the average accelerometer measurement in a short period (1 second) to represent the $\boldsymbol{g}^{b_i}$. $\bar{\boldsymbol{g}}^{b_i}$ is defined as:

$$\bar{\boldsymbol{g}}^{b_i} = \frac{1}{N}\sum_s^N \boldsymbol{R}_{b_i b_{i+s}}(\boldsymbol{a}_{i+s} - \boldsymbol{b}_a) \tag{20}$$

$$\boldsymbol{R}_{b_i b_{i+s}} = \prod_s^N Exp_{SO(3)}((\boldsymbol{w}_{i+s} - \boldsymbol{b}_g)\boldsymbol{\delta t}) \tag{21}$$

where, $\boldsymbol{a}_{i+s}$ and $\boldsymbol{w_{i+s}}$ represent accelerometer and gyroscope measurement, $\boldsymbol{b}_a$ and $\boldsymbol{b}_g$ represent biases of accelerometer and gyroscope estimated by LIO, $\boldsymbol{R}_{b_i b_{i+s}}$ represent rotation of the body frame from moment $i+s$ to moment i, $Exp_{SO(3)}(\cdot)$ taking an axis-angle vector to its rotation matrix in $SO(3)$.

$\boldsymbol{r}_{\overline{mag}}$ represents the constraint of the global average magnetic field vector. To reduce the effect of local magnetic field perturbation, the average magnetometer measurements are adopted.

$$\boldsymbol{r}_{\overline{mag}}(\boldsymbol{R}_{nb_i}, \boldsymbol{b}_m, \boldsymbol{m}^n) = \boldsymbol{R}_{nb_i}\bar{\boldsymbol{m}}^{b_i} - \boldsymbol{m}^n \tag{22}$$

$$\bar{\boldsymbol{m}}^{b_i} = \frac{1}{N}\sum_s^N \boldsymbol{R}_{b_i b_{i+s}}(\boldsymbol{m}_{i+s} - \boldsymbol{b}_m) \tag{23}$$

where $\boldsymbol{b}_m$ is magnetometer biases and is different for each trajectory. To achieve higher efficiency, we ignore the error of $\boldsymbol{R}_{b_i b_{i+s}}$ caused by integrated angular velocity within 1 second. Because, benefiting from the gyroscope biases estimated by LIO, the maximum error of $\boldsymbol{R}_{b_i b_{i+s}}$ is less than 1 degree within 1 second.

The cost function (16) defines a nonlinear least-squares problem that can be solved by iteratively linearizing (16) and finding a solution to the linearized problem. We use ceres- solver [31] to solve the problem. We select Levenberg-Marquardt as the non-linear least squares solver and Conjugate Gradients solver on the normal equations (CGNR) as the linear solver.

After pose graph optimization, there are still some outlier sub-trajectories that cannot be processed normally. Therefore, we use the following criteria to identify and remove outlier sub-trajectories: select a keyframe as the center, count the number of keyframes within a preset distance radius, and the keyframe is available when the number of frames is greater than the preset threshold. Then, sub-trajectories are judged as outliers when the number of valid keyframes is less than 90%. Finally, the final merged trajectory is obtained.

VI. Experiments and Results

A. Experiments Setup

We conduct a comprehensive experiment to evaluate the system performance of the proposed method in a typical indoor shopping mall environment. Figure 3 (a) shows the colorized point cloud of the shopping mall, as the floor plan. Figure 3 (b) and (c) show two images captured in the shopping mall. Compared with the typical office building scene, the magnetic field features in the shopping mall are less distinguishable due to the open environment. Moreover, there are almost no strict right-angle turns in the spatial structure of the shopping mall, which makes the traditional crowdsourcing-based construction method of magnetic field maps unusable.

In the experiment, six volunteers (4 males and 2 females) participated in the data collection and used five smartphones (Samsung S10, Samsung S20, Mi 10, iPhone 13 Pro, and iPad Pro). To reflect the real activity habits of the public users in the actual indoor scene, we do not restrict the behavior of the test users during the data collection process, including pedestrian movement patterns, movement speeds, trajectory shapes, and mobile phone holding methods. The acquired sensor data includes a gyroscope, accelerometer, and magnetometer at 100 Hz. The dataset includes 12 hours of data without ground truth and a subset (about 40 minutes) with ground truth. The ground truth is calculated by using depth-assisted visual bundle adjustment.

To improve data collection efficiency, the duration of a single test for a single user lasts 10 to 20 minutes. This obviously does not conform to the objective law of the real activity duration of users. Therefore, we must divide the trajectory according to the specified parameters such that the length of each resulting trajectory is smaller than the specified threshold $\boldsymbol{l}_{traj}$. Here, $\boldsymbol{l}_{traj}$ is the time length of the trajectory. Notably, because the magnetic field vector of a single point lacks any degree of discrimination, the matching between keyframes can only depend on a magnetic field sequence comparison. Therefore, the magnetic field sequences cannot be connected before and after the split point. This differs from

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

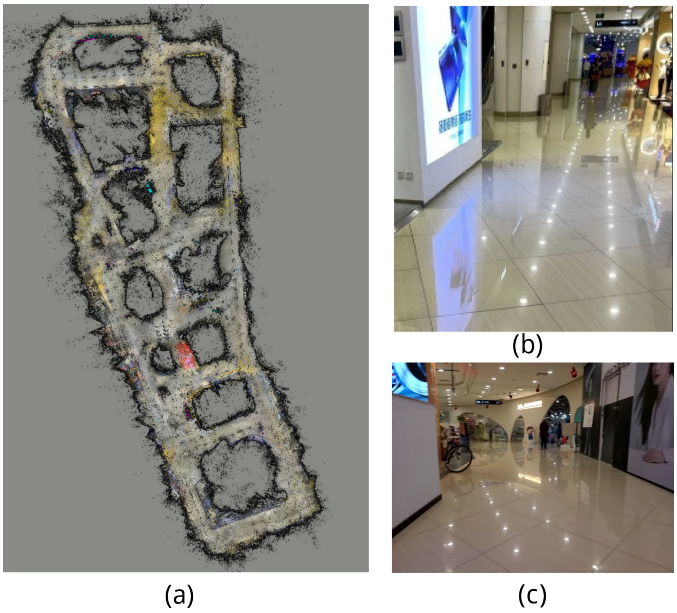

Fig. 3. The spatial structure of the shopping mall. (a) Point cloud map generation based on post-processing visual SLAM. (b) and (c) The image captured in the shopping mall.

WiFi and visual, which can do keyframe association using all signals observed in one epoch.

B. Evaluation Metric

We divide the test data with reference truth into two parts: the first part (about 12 minutes) is used to evaluate the accuracy of the crowdsourcing-based construction of the magnetic field map, and the second part (about 28 minutes) is used to evaluate the magnetic positioning accuracy using the built magnetic map.

After the magnetic map construction, the estimated positions of the first part of the trajectories are denoted as $\{\hat{\boldsymbol{t}}_0, ..., \hat{\boldsymbol{t}}_N\}$ and the corresponding reference positions are denoted as $\{\boldsymbol{t}_0, ..., \boldsymbol{t}_N\}$. Three metrics, including E_m, E'_m, and E_p, are adopted to evaluate the proposed method.

The aligned horizontal position residuals E_m is defined as:

$$E_m = \frac{1}{N} \sum_{i \in [0,N]} \|\boldsymbol{R}_a(\hat{\boldsymbol{t}}_i - \boldsymbol{t}_a) - \boldsymbol{t}_i\| \tag{24}$$

where,

$$\{\boldsymbol{R}_a, \boldsymbol{t}_a\} = \underset{\{\boldsymbol{R}_a, \boldsymbol{t}_a\}}{\arg\min} \sum_{i \in [0,N]} \|\boldsymbol{R}_a(\hat{\boldsymbol{t}}_i - \boldsymbol{t}_a) - \boldsymbol{t}_i\| \tag{25}$$

$\|\cdot\|$ is $l2$-norm. $\boldsymbol{R}_a$ and $\boldsymbol{t}_a$ are 2D rotation and translation, respectively.

$E_{m'}$ is the horizontal position error while accounting for the effect of scale, defined as:

$$E_{m'} = \frac{1}{N} \sum_{i \in [0,N]} \|\boldsymbol{s} \cdot \boldsymbol{R}'_a(\hat{\boldsymbol{t}}_i - \boldsymbol{t}'_a) - \boldsymbol{t}_i\| \tag{26}$$

where

$$\{\boldsymbol{R}'_a, \boldsymbol{t}'_a, \boldsymbol{s}\} = \underset{\{\boldsymbol{R}'_a, \boldsymbol{t}'_a, \boldsymbol{s}\}}{\arg\min} \sum_{i \in [0,N]} \|\boldsymbol{s} \cdot \boldsymbol{R}'_a(\hat{\boldsymbol{t}}_i - \boldsymbol{t}'_a) - \boldsymbol{t}_i\| \tag{27}$$

$\boldsymbol{R}_a$, $\boldsymbol{t}_a$, and $\boldsymbol{s}$ are 2D rotation, translation and scale, respectively.

E_p is the horizontal position error using the built magnetic map that is alignment with the reference framework using (27). The positioning method is a brute force sequence-matching-based method [32]. The matching is based on a magnetic field sequence withing a distance of 10 meters of the keyframe. The E_p is defined as:

$$E_p = \frac{1}{N} \sum_{i \in [0,N]} \|\hat{\boldsymbol{t}}^l_i - \boldsymbol{t}_i\| \tag{28}$$

where, $\hat{\boldsymbol{t}}^l_i$ is localization result based on built magnetic map. $\boldsymbol{t}_i$ is correspondence reference position.

C. System Performance

This section evaluates the accuracy of the magnetic field maps generated by the proposed method using a real dataset. To obtain more believable conclusions, we evaluate the system performance under different parameters, including dataset size of 3, 6, 9, and 12 hours, and sub-trajectory time length of 60, 75, and 90 seconds. Next, we analyze and discuss the results with the data size of 12 hours and a sub-trajectory time length of 90 seconds.

Figure 4 (a) shows the crowdsourcing trajectories reconstructed by the learning-based inertial odometry as described in Section III. Since there are no presumptions regarding the beginning point of each trajectory, the first frame's coordinates are used as the origin. We can find that the heading of all trajectories is approximately aligned. This benefits from the assumption that the y-axis of the magnetic field for each trajectory is zero. Figure 4 (b) depicts the outcome of the pose graph optimization (Section V-B), whereby the majority of trajectories have been combined. Some trajectories are affected by magnetometer bias or heading estimation error and cannot be accurately correlated with other trajectories. However, this phenomenon accounts for a small proportion of the dataset and does not affect the accuracy of the correct correlation trajectory with a large proportion. At this time, the density of a small number of wrong keyframe pair is obviously lower than that of the correctly associated keyframe at this position, and we can easily eliminate them by density clustering. Furthermore, Figure 4 (c) shows the local area of the recovered trajectories, clearly showing that the shape and path of each trajectory used to construct the map are irregular and not designed.

Figure 5 (a) and Figure 5 (b) show the results of aligning the estimated trajectories with the reference trajectories using two different alignment methods (with or without scale correction), Figure 5 (c) shows the localization results using the reconstructed map that aligned with the reference framework using (27). The semi-transparent gray lines represent the reference trajectory, all the colored lines represent the estimated

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

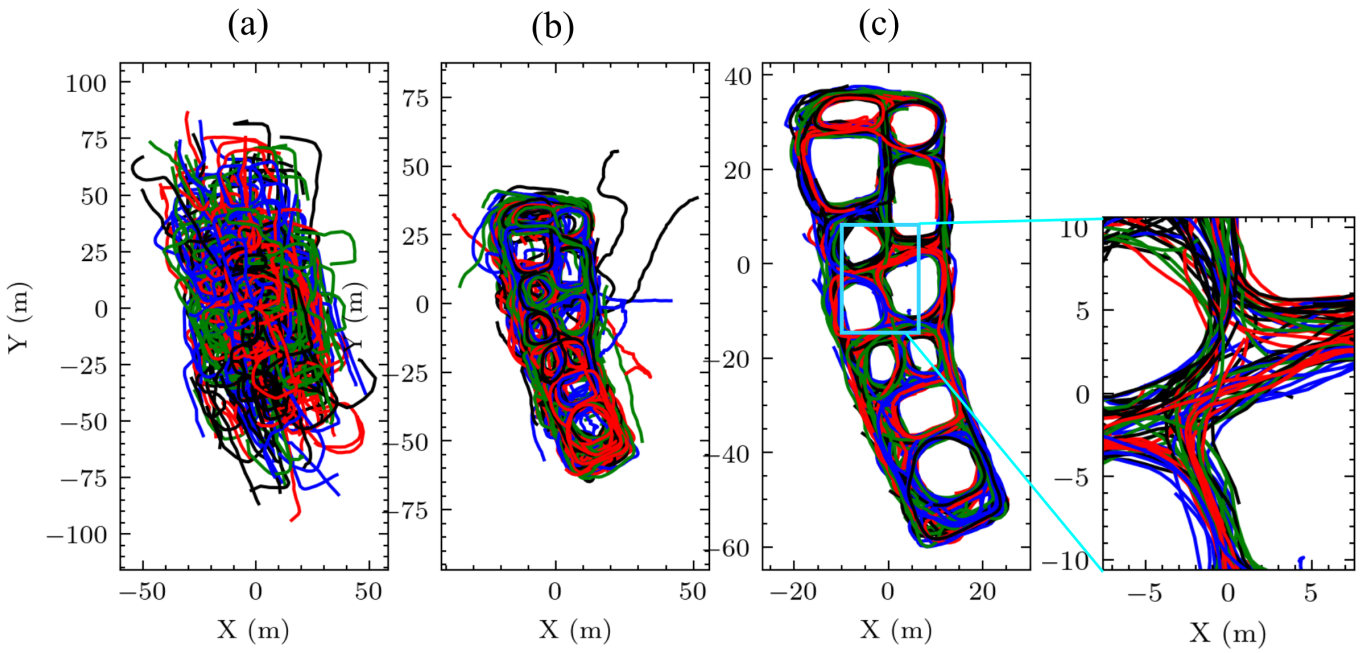

Fig. 4. The recovered trajectories of the proposed crowdsourcing-based magnetic map construction method. (a) Raw trajectories estimated using learning-based inertial odometry. (b) Trajectories recovered using two-step graph-pose-optimization. (c) Final selected trajectories through outlier rejection.

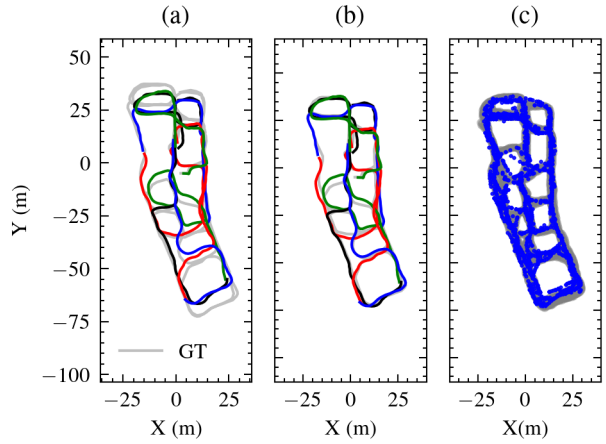

Fig. 5. Results on the test dataset estimated using the proposed method. The gray line represents reference positions, the colored solid lines represent the recovered trajectories using the first part of the validation dataset, and the blue dots represent the positioning results using the second part of the validation dataset. (a) The recovered trajectories with scale correction. (b) The recovered trajectories without scale correction. (c) Positioning results are based on the magnetic map generated by the proposed method.

TABLE II
ACCURACY OF THE PROPOSED METHOD WITH DIFFERENT DATASET SIZE AND TRAJECTORY LENGTH

Trajectory Length (Second)		Dataset Size (Hour) 12	9	6	3
60	E_m (m)	4.26	**4.06**	4.67	8.25
	$E_{m'}$ (m)	2.37	**2.17**	3.37	9.88
	E_p (m)	**3.20**	3.36	3.60	9.39
75	E_m (m)	**3.66**	3.67	4.02	5.43
	$E_{m'}$ (m)	1.91	**1.89**	2.47	4.38
	E_p (m)	**2.60**	2.77	3.41	4.78
90	E_m (m)	3.18	3.15	**3.00**	3.94
	$E_{m'}$ (m)	1.48	1.46	**1.36**	2.84
	E_p (m)	**2.53**	2.83	2.68	3.91

trajectories, and the blue dots represent the results of sequence-matching-based localization. Figure 5 (a)shows the proposed method can accurately recover the relative spatial relationship between crowdsourced trajectories but suffers from significant scale error. The reason is that there is a scale error in the user trajectory estimated by LIO, which is characteristic of almost all traditional PDR algorithms and deep learning-based inertial odometry. Nevertheless, the effect of scale bias can be eliminated by global correction, as shown in Figure 5(b). At the same time, based on the scale-corrected magnetic field map, most of the positions estimated based on magnetic field sequence matching have good overlapping consistency with the reference position, as shown in Figure 6(c). The average horizontal error of magnetic map and localization are 1.48m ($E_{m'}$) and 2.53m (E_p), respectively. In general, the above results can prove the feasibility and effectiveness of the proposed method to a certain extent.

To further verify the performance of the proposed methods, an ablation test is provided to illustrate the effect of the time length of trajectories and the size of the dataset. As illustrated in Table II, we investigate E_m, $E_{m'}$, and E_p of different time lengths and dataset sizes. We can conclude that larger dataset sizes and longer trajectory lengths result in more precise map construction and positioning. Longer trajectory length improves the mapping accuracy because a longer trajectory length can increase the probability of generating correct associated keyframes with other trajectories. Meanwhile, it is helpful to eliminate incorrectly associated frames. The benefit of a larger dataset is that more sub-trajectory shapes are

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

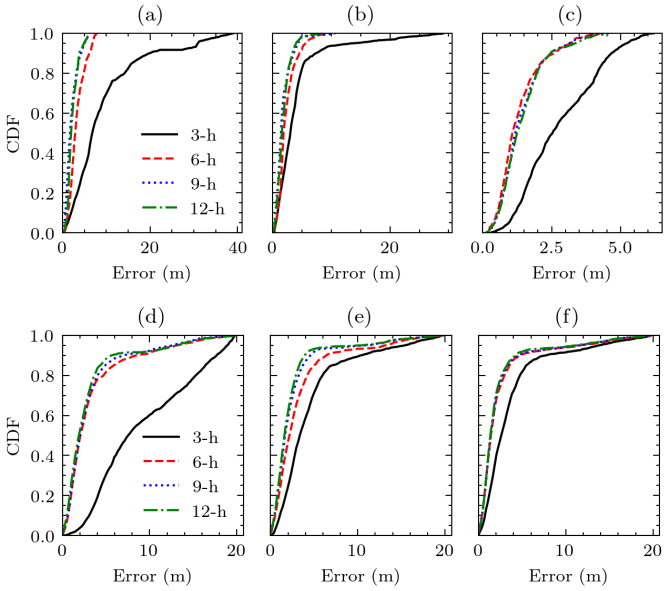

Fig. 6. CDF of trajectory error ($E_{m'}$) and positioning error (E_p) using various dataset sizes and trajectory length. (a) and (d) trajectory length is 60 seconds, (b) and (e) trajectory length is 75 seconds, (c) and (f) trajectory length is 90 seconds.

TABLE III
TIME CONSUMPTION OF THE PROPOSED METHOD (12-HOUR DATASET).

	LIO[1]	KA[2]		GTO[3]	Total Time
		FE[4]	KPD[5]		
Time (second)	43.0	0.9	9.8	7.1	60.8
Ratio (%)	70.7	1.5	16.1	11.7	

[1] Learning-based Inertial Odometry (Section III)
[2] Keyframe Association (Section IV)
[3] Global Trajectory Optimization (Section V)
[4] Feature Extraction (Section IV-A)
[5] Keyframe Pair Determination (Section IV-B)

mutually constrained to reduce incorrectly associated keyframe pairs.

Figure 6 shows the cumulative distribution function (CDF) of the proposed method using different dataset sizes and trajectory lengths. Figure 6 (a), (b), and (c) show the CDF of trajectory error ($E_{m'}$) using various dataset sizes to build the magnetic map. Figure 6 (d), (e), and (f) show the CDF of positioning error (E_p) to show the effect of dataset size on the positioning performance of the built magnetic map. The proposed method using a small-size dataset (3 hours) shows significantly lower accuracy in both $E_{m'}$ and E_p. Thus, the proposed method needs enough data to achieve the magnetic map construction when the trajectory length is limited. In other words, a large dataset is necessary for magnetic map building via crowdsourcing using short trajectory. Nonetheless, the spatial distribution of the data collected in this experiment is uniform, which may not reflect reality. Therefore, the total amount of data required for a region of the same size in a real-world scenario may be greater than that used in experiments.

TABLE IV
TIME CONSUMPTION OF KEYFRAME ASSOCIATION OF RELATED CROWDSOURCING-BASED METHODS.

Method	Data Source	Dataset Size	Time Cost
WiFi-RITA [16]	IMU,WiFi	10 hours	4520.3 seconds
[23]	IMU,Mag	15 minutes	37.3 seconds
Direct-Method	IMU,Mag	12 hours	127.7 seconds
Fast-Method (proposed)	IMU,Mag	12 hours	9.8 seconds

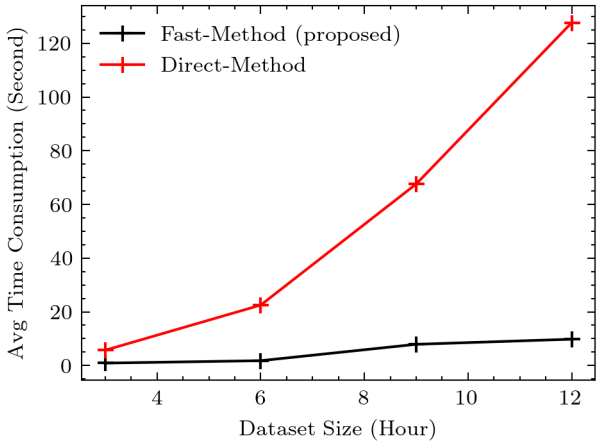

Fig. 7. Time consumption of different keyframe association methods.

D. Efficiency of the proposed keyframe association method

Since large-scale data sets are the basic guarantee for the feasibility of crowdsourcing-based solutions, the efficiency of data processing is a very important evaluation indicator. Table III gives the specific time consumption of each algorithm module when using the proposed method to process the 12-hour data set. The computing platform is a laptop, the main parameters include a Ryzen R7-5800H 8-core CPU and an NVIDIA GTX 3060 GPU. The proposed algorithm takes a total of 60.8 seconds to process the 12-hour data set, and the fast key frame association method takes 10.7 seconds, accounting for less than 20%. We can learn that the proposed fast keyframe association method is efficient and will no longer be the largest time-consuming module of crowdsourcing-based solutions. At the same time, we find that LIO is the most time-consuming algorithm module, taking 43 seconds, accounting for 70%. This is because of the poor performance of the GPU used in this experiment, and using a higher-performance GPU is an effective way to improve computational efficiency further. The global trajectory optimization causes only 7.1 seconds. It benefits from the optimization being split into two-stage. The first stage described in Sec V-A converges quickly since the state space is small and can provide a good initial value for the second global pose graph optimization to improve the convergence speed of the second stage.

Figure 7 shows the time consumption of direct keyframe association (denoted as Direct-Method) and the proposed fast keyframe association method (denoted as Fast-Method) for different sizes of datasets. As the dataset size increases, the time consumption of the Direct-Method increases linearly, while the time consumption of the Fast-Method increases

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

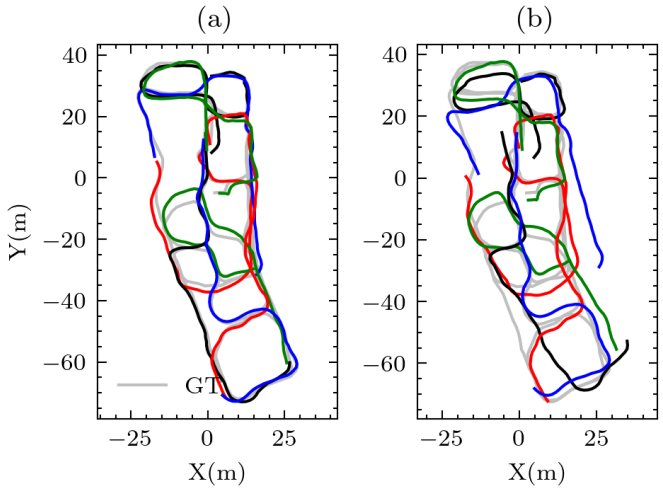

Fig. 8. Recovered trajectories with and without inequality constraint optimization. (a) With inequality constraint. (b) Without inequality constraint.

slowly. In particular, when processing a 12-hour data set, the efficiency of the Fast-Method is more than 90% higher than that of Direct-Method, showing that the proposed method is feasible and the efficiency improvement is very obvious.

Tabel IV summarizes the time consumption of keyframe association for state-of-the-art crowdsourcing-based schemes. WiFi-RITA [16] considers many interference factors of real-world environments and is the most advanced WiFi-based crowdsourcing method in the previous literature. The associated keyframe determination of this method is efficiency (takes only 12.3 seconds). But it has a time-consuming preprocessing stage (WiFi mark searching) which costs 4508 seconds. [23], the magnetic field-based crowdsourcing solution which notices to reduce time consumption, takes 37.3 seconds for keyframe association when processing a 15-minute dataset. This solution is inefficient because it uses a hidden Markov model to do keyframe association. The fact that the time consumption (37.3 seconds) of [23] processing 15 minutes dataset exceeds the time consumption of the proposed method processing 12 hours dataset indicates the proposed method with significantly better efficiency. Furthermore, it is noticed that [23] relies on detecting corners and only associating corners to reduce time consumption. Thus, [23] is hard to adapt to complex buildings that not only consist of hallways and corners.

E. Effect of the proposed two-step optimization

This section mainly explains that two-step graph optimization can more effectively resist the destructive effects of wrong keyframe associations than one-step graph optimization. Figure 8 shows the comparison of mapping results using the global optimization with or without inequality constraint optimization. The optimized trajectories without inequality constraints (Figure 9 (b)) show significantly worse relative pose estimation performance compared to the trajectories in Figure 9 (a). $E_{m'}$ of graph optimization method with or without inequality constraint optimization are 1.48m and 3.60m, respectively. The reason is incorrectly associated keyframe pairs resulting in incorrect subtrack headings and shapes. In that case, the inequality constraint optimization is immune to the effect of this incorrectly associated keyframe pair. The good initial value helps the two-step pose graph optimization get rid of being badly affected by the false keyframe association.

VII. Conclusion and Future Work

Aiming at the problem that existing methods are not suitable for typical indoor environments (e.g., shopping malls) due to too many restrictions on crowdsourced data, this paper proposes a novel crowdsourcing-based magnetic map construction method. The proposed method consists of three stages, including learning-based inertial odometry (LIO), keyframe association, and global trajectory optimization. We adopt the LIO to estimate accurate trajectory and sensor biases regardless of various motion modes. Benefiting from properly using the information of the time-domain feature and frequency-domain feature, the keyframe association method can achieve efficient keyframe association using short-period trajectories with irregular shapes. The global trajectory optimization is proposed to remove incorrectly associated keyframe pairs and improve robustness.

The performance of the proposed method is verified by a field test conducted in a shopping mall. The proposed method processes a dataset (dataset size is 12 hours, trajectory length is 90 seconds) in 60.8 seconds and provides merged trajectories with an average position error of 1.48 meters (with scale correction). The ablation study demonstrates the effect of dataset size and trajectory length and proves the necessity of large-scale datasets. Furthermore, we verify the efficiency of the proposed keyframe association method and the robustness of the proposed two-step graph optimization method.

The proposed method relies on recovered user trajectories, however, LIO is a data-driven approach, and its performance on general public user datasets is still an open issue. Therefore, in future work, we intend to enhance the LIO method and provide reliability metrics for LIO using real datasets from general public users. Moreover, consideration should be given to the method for creating maps in more complex scenarios, such as areas with changing local magnetic fields (e.g., underground garages) and multi-story buildings.

Acknowledgment

References

[1] N. El-Sheimy and Y. Li, "Indoor navigation: State of the art and future trends," *Satellite Navigation*, vol. 2, no. 1, pp. 1–23, 2021.
[2] A. Mackey, P. Spachos, L. Song, and K. N. Plataniotis, "Improving ble beacon proximity estimation accuracy through bayesian filtering," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3160–3169, 2020.
[3] L. Zhang, Z. Chen, W. Cui, B. Li, C. Chen, Z. Cao, and K. Gao, "Wifi-based indoor robot positioning using deep fuzzy forests," *IEEE Internet Things J.*, vol. 7, no. 11, pp. 10 773–10 781, 2020.
[4] F. Liu, J. Liu, Y. Yin, W. Wang, D. Hu, P. Chen, and Q. Niu, "Survey on wifi-based indoor positioning techniques," *IET Communications*, vol. 14, no. 9, pp. 1372–1383, 2020.
[5] J. Kuang, X. Niu, P. Zhang, and X. Chen, "Indoor positioning based on pedestrian dead reckoning and magnetic field matching for smartphones," *Sensors*, vol. 18, no. 12, p. 4142, 2018.
[6] C. Chen, X. Lu, A. Markham, and N. Trigoni, "Ionet: Learning to cure the curse of drift in inertial odometry," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3305840

[7] W. Liu, D. Caruso, E. Ilg, J. Dong, A. I. Mourikis, K. Daniilidis, V. Kumar, and J. Engel, "Tlio: Tight learned inertial odometry," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4.
[8] J. Kuang, X. Niu, and X. Chen, "Robust pedestrian dead reckoning based on mems-imu for smartphones," *Sensors*, vol. 18, no. 5, p. 1391, 2018.
[9] S. Herath, H. Yan, and Y. Furukawa, "Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, & new methods," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 3146–3152.
[10] Y. Wang, J. Kuang, X. Niu, and J. Liu, "Llio: Lightweight learned inertial odometer," *IEEE Internet of Things Journal*, 2022.
[11] H. Cao, Y. Wang, J. Bi, S. Xu, M. Si, and H. Qi, "Indoor positioning method using wifi rtt based on los identification and range calibration," *ISPRS International Journal of Geo-Information*, vol. 9, no. 11, p. 627, 2020.
[12] Y. Zhu, X. Luo, S. Guan, and Z. Wang, "Indoor positioning method based on wifi/bluetooth and pdr fusion positioning," in *2021 13th International Conference on Advanced Computational Intelligence (ICACI)*. IEEE, 2021, pp. 233–238.
[13] L. Hua and J. Yang, "Smartfps: Neural network based wireless-inertial fusion positioning system," *arXiv preprint arXiv:2209.13261*, 2022.
[14] J. Kuang, T. Li, Q. Chen, B. Zhou, and X. Niu, "Consumer-grade inertial measurement units enhanced indoor magnetic field matching positioning scheme," *IEEE Transactions on Instrumentation and Measurement*, 2022.
[15] B. Zhou, W. Ma, Q. Li, N. El-Sheimy, Q. Mao, Y. Li, F. Gu, L. Huang, and J. Zhu, "Crowdsourcing-based indoor mapping using smartphones: A survey," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 177, pp. 131–146, 2021.
[16] Z. Li, X. Zhao, Z. Zhao, and T. Braun, "Wifi-rita positioning: Enhanced crowdsourcing positioning based on massive noisy user traces," *IEEE transactions on wireless communications*, vol. 20, no. 6, pp. 3785–3799, 2021.
[17] B. Lashkari, J. Rezazadeh, R. Farahbakhsh, and K. Sandrasegaran, "Crowdsourcing and sensing for indoor localization in iot: A review," *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2408–2434, 2018.
[18] G. Shen, Z. Chen, P. Zhang, T. Moscibroda, and Y. Zhang, "{Walkie-Markie}: Indoor pathway mapping made easy," in *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, 2013, pp. 85–98.
[19] Z. Li, X. Zhao, F. Hu, Z. Zhao, J. L. C. Villacrés, and T. Braun, "Soicp: A seamless outdoor–indoor crowdsensing positioning system," *IEEE internet of things journal*, vol. 6, no. 5, pp. 8626–8644, 2019.
[20] Y. Qian, H. Yan, S. Herath, P. Kim, and Y. Furukawa, "Single user wifi structure from motion in the wild," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 2157–2163.
[21] H. Luo, F. Zhao, M. Jiang, H. Ma, and Y. Zhang, "Constructing an indoor floor plan using crowdsourcing based on magnetic fingerprinting," *Sensors*, vol. 17, no. 11, p. 2678, 2017.
[22] A. Ayanoglu, D. M. Schneider, and B. Eitel, "Crowdsourcing-based magnetic map generation for indoor localization," in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2018, pp. 1–8.
[23] M. Kwak, C. Hamm, S. Park, and T. T. Kwon, "Magnetic field based indoor localization system: A crowdsourcing approach," in *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2019, pp. 1–8.
[24] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.
[25] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *science*, vol. 315, no. 5814, pp. 972–976, 2007.
[26] X. Hou and J. Bergmann, "Pedestrian dead reckoning with wearable sensors: A systematic review," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 143–152, 2020.
[27] S. Sun, D. Melamed, and K. Kitani, "Idol: Inertial deep orientation-estimation and localization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 7, 2021, pp. 6128–6137.
[28] Y. Wang, J. Kuang, Y. Li, and X. Niu, "Magnetic field-enhanced learning-based inertial odometry for indoor pedestrian," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–13, 2022.
[29] Y. Wang, X. Li, and J. Zou, "A foot-mounted inertial measurement unit (imu) positioning algorithm based on magnetic constraint," *Sensors*, vol. 18, no. 3, p. 741, 2018.
[30] P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in statistics*. Springer, 1992, pp. 492–518.
[31] S. Agarwal, K. Mierle, and T. C. S. Team, "Ceres Solver," 3 2022. [Online]. Available: https://github.com/ceres-solver/ceres-solver
[32] Y. Li, X. Niu, P. Zhang, H. Lan, Y. Zhuang, and N. El-Sheimy, "Smartphone-based indoor navigation using pdr and magnetic matching," in *Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*, 2015, pp. 2060–2066.

Yan Wang received the B.Eng. degree in Chemical Engineering and Technology from China University of Mining and Technology, Xuzhou, China, in 2016. And he received an M.S. degree in Computer Applied Technology from China University of Mining and Technology, Xuzhou, China, in 2019. He is currently pursuing a Ph.D. degree in GNSS Research Center, Wuhan University, Wuhan, China. His research interests focus on indoor navigation, sensor fusion algorithm, and computer vision.

Jian Kuang received the B.Eng. degree and Ph.D. degree in Geodesy and Survey Engineering from Wuhan University, Wuhan, China, in 2013 and 2019, respectively. He is currently an associate research fellow at the GNSS Research Center of Wuhan University, China. His research interests focus on inertial navigation, magnetic field positioning, and smartphone-based vehicle and pedestrian navigation.

Tianyi Liu received the B.S. degree in electronic information engineering from Wuhan University, Wuhan, Hubei, in 2015 and the M.S. degree in signal and information processing from Chinese Academy of Science, Beijing, in 2018. He is currently pursuing the Ph.D. degree in communication and information system at Wuhan University, Wuhan, Hubei, China. His research interest includes deep learning, LiDAR SLAM and integrated navigation.

Xiaoji Niu received the B.Eng. degree (with honors) in Mechanical and Electrical Engineering and the Ph.D. from Tsinghua University, Beijing, China. He was a Postdoctoral Fellow with the Mobile MultiSensor Systems (MFMSS) Research Group, Department of Geomatics Engineering, and the University of Calgary. He was a Senior Scientist with SiRF Technology, Inc. At present, he is a Professor of the GNSS Research Center, Hubei Luojia Laboratory, at Wuhan University, Wuhan, China. His research interests focus on INS, GNSS/INS integration for vehicle, robot, and pedestrian navigation.

Jingnan Liu is a member of the Chinese Academy of Engineering. He was born in 1943 and graduated from the Wuhan Institute of Surveying and Mapping with a bachelor's degree in astronomical geodesy. He obtained his master's degree in engineering in 1982. Professor Liu has long been engaged in research and teaching in geodetic surveying. He is considered a pioneer in the application of GNSS technology. Professor Liu has taken part in a number of research programs to promote the application of satellite positioning systems in China.

Authorized licensed use limited to: Wuhan University. Downloaded on September 28,2023 at 01:24:51 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.