2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Madrid, Spain, October 1-5, 2018

Indoor Mapping and Localization for Pedestrians using Opportunistic Sensing with Smartphones

Qing Liang[1], Lujia Wang[2], Youfu Li[3], and Ming Liu[1]

***Abstract*— Indoor localization for pedestrians has gained increasing popularity among the rich body of literature for the last decade. In this paper, a low-cost indoor mapping and localization solution is proposed using the opportunistic signals from ambient indoor environments with a smartphone. It is composed of GraphSLAM-based offline mapping and Bayesian filtering-based online localization using generated signal maps. The GraphSLAM front-end is constructed by motion constraints from pedestrian dead-reckoning (PDR), loop-closure constraints identified by magnetic sequence matching with WiFi signal similarity validation, and observation constraints from opportunistic magnetic headings after error rejection. Globally consistent trajectories are created by graph optimization, after which signal maps (e.g., WiFi, magnetic fields, lights) are generated by Gaussian Processes Regression (GPR) for later localization. We propose to use the pseudo-wall constraints from the GPR variance map of magnetic fields and the lights measurements as observations for particle filtering. The proposed method is evaluated on several datasets collected from both the in-compass office buildings and outside public areas. Real-time localization is demonstrated on a smartphone in an office building covering 2000 square meters with the 50- and 90-percentile accuracies being 2.30 m and 3.41 m, respectively.**

I. Introduction

To meet the ever-rising demand for location-based services (LBSs), low-cost solutions to indoor localization for pedestrians have attracted much attention in recent years, especially with the rapid development of sensory-rich and computational powerful smartphones [1]. In particular, opportunistic signals (e.g., WiFi, magnetic fields) have been successfully exploited for low-cost localization via fingerprinting [2]–[5], due to their pervasiveness and free usage in modern buildings. As for fingerprinting methods, a fingerprint database (a.k.a, signal maps) is required as a prior to archive online localization. Manual site surveys are conventional approaches to signal map generation. However, they are time-consuming and labor-intensive [6].

*This work was supported by Shenzhen Science, Technology and Innovation Commission (SZSTI) JCYJ20160428154842603, JCYJ20160401100022706, and JCYJ20170818153518789; Research Grant Council of Hong Kong SAR Government, China, under project No. 16212815, 21202816; partially supported by the HKUST Project IGN16EG12, awarded to Prof. Ming Liu; and was also supported by the National Natural Science Foundation of China No. 61603376, awarded to Dr. Lujia Wang.

[1]Qing Liang and Ming Liu are with Department of ECE, Hong Kong University of Science and Technology. dobe.liang@gmail.com eelium@ust.hk

[2]Lujia Wang is with Center for Cloud Computing, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. lj.wang1@siat.ac.cn

[3]Youfu Li is with Department of MBE, City University of Hong Kong. meyfli@cityu.edu.hk

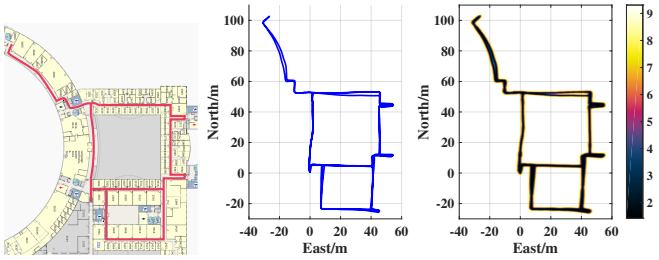


Fig. 1: From left to right: reference walking path labeled on a floorplan for a large area of the HKUST academic building; optimized trajectory by the proposed method and the generated GPR variance map of magnetic fields as a side product.

To improve survey efficiency, automatic site surveying approaches based on crowd-sourcing techniques [6], [7] have been proposed in the mobile computing community. Smartphone users are employed as free site-surveyors to crowd-source their local sensing data to a public location service provider for signal map generation. However, these methods suffer from the data-quality control problem caused by heterogeneous devices and unconstrained users [8]. In addition, some prior knowledge (e.g., a physical floorplan) of the survey area is required to provide ground truth locations to label signal fingerprints. Yet the floorplans may not be readily accessible to public users due to privacy concerns.

In the meantime, Simultaneous Localization and Mapping (SLAM) [9] has been exploited by some researchers for signal map generation (or site-survey) and pedestrian localization [4], [10]–[12]. Alternative to the robot SLAM solutions [13] that normally require information-rich perception data (e.g., 3D point clouds) [14] from expensive sensors (e.g., Lidar), signal SLAM techniques try to solve the mapping and localization problems with inbuilt sensors of mobile devices that feature low-quality sensing data. DPSLAM [4] was proposed to solve an online SLAM problem using opportunistic sensing on smartphones. However, the high computation demand renders it less effective in continuous localization on resource-constrained devices. Besides, GraphSLAM-based offline site-surveying methods try to simplify the manual-survey process, and accordingly, to reduce the labor and time taken by surveyors to reach a usable signal map.

WiFi GraphSLAM [10] first formulated the signal SLAM as a GraphSLAM problem. The authors came up with a WiFi measurement model assuming the interpretability of signal strength measurements at nearby locations. This assumption requires dense WiFi sampling in the testing area to ensure

reasonable interpretability at the cost of more surveying time and labor. Pedestrian Dead-reckoning (PDR) [15] is involved to provide odometry information. Generally, magnetic signals indoors are good location features benefiting from the location-specific and temporal-stable nature induced by anomalies from building construction materials. Gao et al. [11] proposed to use magnetic sequence matching for loop-closure detection within a GraphSLAM framework. Gaussian Processes Regression (GPR) [16] was used to generate probabilistic signal maps for later online localization. To reject false loop-closures, the authors used a strong assumption on the spatial distance between two matched sequences, which confined the searching space of each sequence to its spatial vicinity. This may be problematic in large-scale environments where PDR suffers from significant drifts over time. In addition to magnetic matching, authors from [12] incorporated the motion constraints from PDR into loop-closure detection to relax this assumption. Online localization was achieved by Kalman filtering using the generated magnetic maps.

In this paper, we propose a low-cost indoor mapping (site-survey) and localization system for pedestrians using opportunistic signals perceived by smartphone sensors. We build our system on the top of several previous works [10]–[12], and follow the generic framework of GraphSLAM-based offline mapping and Bayesian filtering-based online localization, yet with remarkable improvements. We use magnetic sequence matching to identify loop-closures between distinct steps as [11], [12]. The magnetic sequence is segmented into multiple sub-sequences at turning points which are salient motion patterns for human walking in man-made environments. In addition to the motion pattern, we propose to use sequence-wise WiFi signal similarities for loop-closure validation in order to reduce false positives. This is very important in large-scale environments where an increasing number of false positives will arrive due to mismatches of magnetic signals. Additionally, we propose to involve the opportunistic magnetic headings to the GraphSLAM front-end by carefully identifying magnetic anomalies and throwing away error-prone values. The benefits observed are two-fold: 1) global heading information is available for the optimized trajectory, and 2) map consistency is improved with extra measurement constraints.

GPR is used for signal map generation for WiFi signals and magnetic fields. WiFi maps provide a global location reference for the mapped area. By contrast, [12] only supported position tracking due to the lack of global references. Different from authors in [11] who declared the useless of magnetic maps in localization, we observe that the magnetic GPR variance map resembles well the physical floorplan in shape. We propose to use the generated magnetic fields variance map as a "pseudo floorplan" to provide wall-constraints for particle filtering to help the heading and step length estimation in PDR. Moreover, lights along the walking path are identified and then exploited to further constrain the step lengths. Localization performance will be evaluated on a smartphone in real-time.

We summarize our contributions as follows:

1) WiFi signal similarity validation in aid of magnetic matching-based loop-closure detection and opportunistic magnetic heading measurements for the GrapSLAM front-end construction.
2) Pseudo-wall constraints from generated GPR variance maps of magnetic fields and light measurements for particle filtering-based localization.
3) Implementation of the proposed mapping and localization system, through which globally consistent signal maps and real-time localization are achieved.

The remainder of this paper is organized as follows. The overall system architecture is introduced in Section II. We detail the GraphSLAM-based offline mapping in Section III. The particle filtering-based online localization using signal maps is proposed in Section IV. We present the evaluation results in Section V and conclude this paper in Section VI.

II. Overview

The proposed system (Fig. 2) consists of two core function blocks namely the offline mapping and online localization.

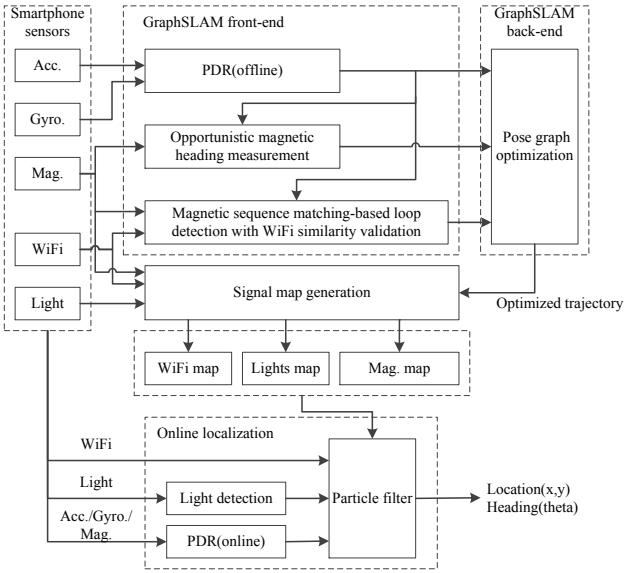


Fig. 2: Architecture overview of the proposed system.

Mapping follows the classical GraphSLAM framework with two main components, namely a front-end that populates an initial pose graph with constraints, and a back-end that optimizes the graph to be most consistent with these constraints. Given the optimized trajectory poses, location-labeled opportunistic signals(e.g., WiFi fingerprints, lights, magnetic field strength) are thus available for map generation, without the needs of a war-driven site survey. Localization follows the standard Bayesian filtering framework where PDR provides the motion model, and opportunistic signal measurements with smartphone sensors provide the observation model with the aid of generated signal maps in the mapping block. We do not desire an online SLAM running on smartphones, as SLAM algorithms are normally much more computational heavy than localization itself. Currently, we use a generic particle filter to implement

Bayesian filtering. Some advanced sensor fusion frameworks [17] can be exploited in the future.

III. GraphSLAM-based offline Mapping with Opportunistic Sensing

In this section, we introduce the GraphSLAM-based offline mapping system with an emphasis on the front-end creation. The whole system is composed of: 1) a GraphSLAM front-end composed of PDR, opportunistic magnetic heading measurements, and magnetic sequence matching-based loop-closure detection with WiFi similarity validation; 2) pose graph optimization; and 3) signal map generation.

A. PDR

PDR is normally composed of step detection, step length estimation, and heading estimation. As it is not the focus of our study, we choose a simple PDR algorithm based on zero-crossing detection using inbuilt inertial sensors on a smartphone. To avoid the side effects of magnetic anomalies, we use only the accelerometer and gyroscope readings for attitude and heading estimation. The step detection is achieved by monitoring the gait cycling pattern exhibited in vertical acceleration. An empirical step model [18] is used for step length estimation. In our context, PDR-derived poses are heavily used in constructing the GraphSLAM front-end as they offer informative motion patterns in aid of magnetic sequence segmentation, opportunistic magnetic heading measurements, and motion constraints for the pose graph generation.

B. Magnetic Sequence Matching-based Loop-closure Detection with WiFi Similarity Validation

We first perform turns detection on the PDR derived trajectory and then split the whole magnetic sequence into multiple segments. Loop-closures are identified by magnetic sequence matching and then validated by the WiFi signal similarity checking to reduce false positives.

1) Turns detection and magnetic sequence segmentation: We assume a space-constraint indoor environment that features straight corridors interconnected by salient turns (e.g., left/right/U-turn). This provides geometry constraints to the user's walking trajectory, which are crucial to successful magnetic matching due to the limited coherence distance of magnetic signals. Turn-taking is assumed as a salient motion pattern in such scenarios. We thus exploit these turns to segment the whole magnetic sequence recorded during the site survey into sub-sequences for later loop detection.

Normally, turns generate salient variations (e.g., peaks and valleys) on the curvature of the PDR-derived trajectory, which may be trivially identified by peak detection, as shown in Fig. 3a. Some user actions (e.g., side movement during door-opening) may induce false positives. In such cases, we try to merge two consecutive segments if they are checked to be in a line. The magnetic field sequence is sliced into multiple segments as per the detected turning points. To reduce mismatches, only those informative segments are selected, with sufficient traveling distances and significant signal variations.

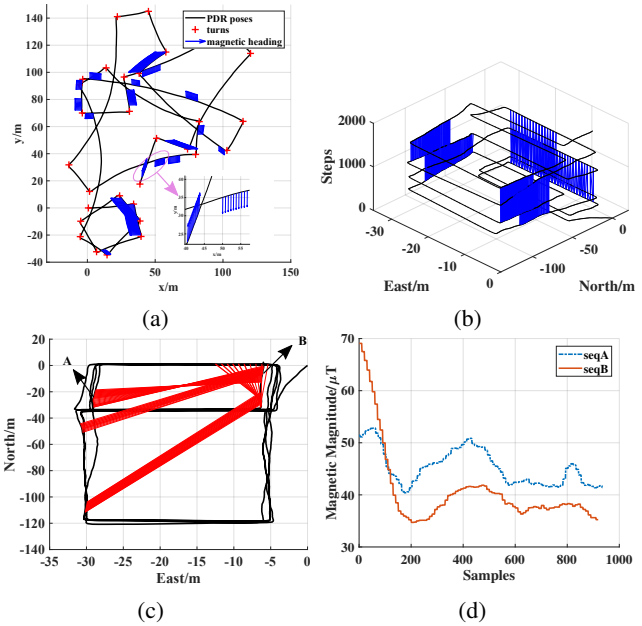


Fig. 3: (a) Turns (red plus signs) and magnetic headings (blue arrows) marked on the PDR trajectory. The magnified view is for better visualization of headings. (b) true positive loop closures shown on the final optimized trajectory stretched along z-axis for better visulization. (c) false positives shown on the 2D trajectory. (d) Mismatched magnetic sequences at location A and B.

2) Loop detection using magnetic matching: The magnetic signals suffer from the spatial sampling density variation problem [3] due to varying walking speeds and sampling frequencies. To counter this problem, we use dynamic time warping (DTW) to match these magnetic sequences. It is proven to be an effective tool in measuring similarities between time-varying signal series. As the user may revisit the same space in opposite directions, we compare each pair of magnetic sequences both chronologically and reverse-chronologically. Given the alignment between two magnetic sequences, the step-to-step correspondences (say loop-closures) between the walking trajectories can be easily determined with timestamps.

Fig. 3a illustrates the correctly detected loop-closures. The pair-wise loop-closures are shown as blue line segments connecting distinct poses on the optimized trajectory from one of the collected datasets. These loop-closures exhibit considerable spatial consistency between overlapped walking trajectories. On the other hand, an increasing number of false positives may occur in large-scale environments (see Fig. 3b), since magnetic sequence matching is error-prone due to inevitable magnetic ambiguities (see Fig. 3c). For instance, loop-closures between trajectory segments A and B are incorrect as they are indeed on two opposite corridors. The magnetic field sequences are shown in Fig. 3d. The two sequences are very similar to each other in shape, despite the fairly long samples in use. We will address this problem in the next section.

3) Loop-closure validation using WiFi similarity: Magnetic signals with anomalies are inherently local features

lacking global uniqueness. Magnetic matching could suffer from an increasing number of false positives in large-scale environments. Too many false positives can corrupt the optimization result. It is important to reduce false positives while maintaining as many true positives as possible.

WiFi signals at faraway locations exhibit poor similarities in signal strengths due to the significant signal attenuation in complex indoor environments. The spatial uniqueness of WiFi measurements improves with the dense deployment of WiFi APs. Inspired by the complementary nature of the two signal modalities, we propose to use the WiFi signal similarity as an auxiliary tool to help identify incorrect loop-closures from magnetic matching.

WiFi scans collected along the walking trajectory form a sequence of measurements indexed by timestamps. To compare the similarity between two given WiFi sequences, a sequence-wise WiFi signal similarity metric, as illustrated in Fig. 4a, is proposed. A similarity matrix between the two sequences is created by computing point-wise signal similarities for each pair of samples[1]. If the two sequences match chronologically as reported by the loop detection module, the average similarity of the diagonal elements in the similarity matrix is adopted as the sequence-wise similarity. Otherwise, the back-diagonal elements are used, as shown in Fig. 4b. Two examples of the similarity matrix are illustrated in Fig. 4c and Fig. 4d, which are created by two pairs of WiFi measurement sequences with false-positive and true-positive loop-closures, respectively.

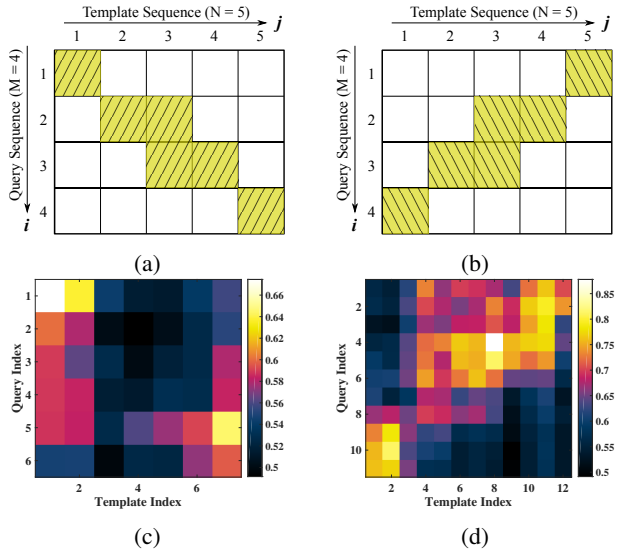


Fig. 4: (a) and (b) show the proposed similarity metric between two WiFi sequences, where we have $M = 4$ samples in the query sequence and $N = 5$ samples in the template sequence. (c) shows the case of mismatching and (d) correct matching.

As for the case in Fig. 3b, the mismatched loop-closures can be identified by the proposed sequence-wise WiFi similarity. Admittedly, it is not the case for false loop-closures created between two poses at nearby locations due to WiFi's poor spatial resolution. However, the proposed strategy works reasonably well in rejecting large amounts of incorrect loop-closures on real-world datasets.

[1]We use the Gaussian kernel distance to measure the pair-wise signal similarity between two WiFi scans.

C. Opportunistic Magnetic Heading Measurements

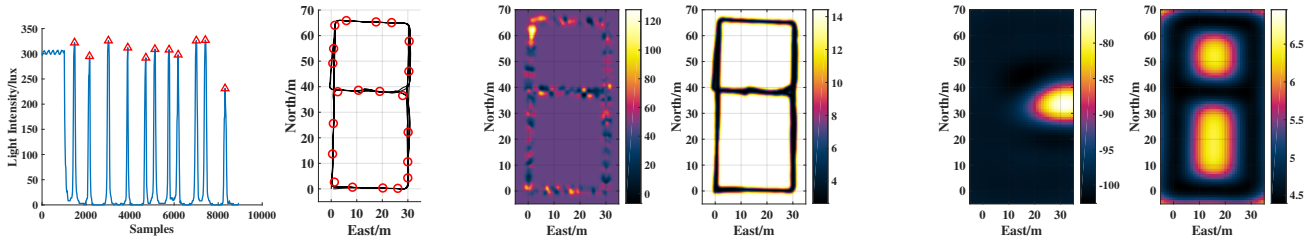
In this section, we consider another case where geomagnetic fields dominate ambient magnetic signals. Note that the typical coherence distance of magnetic fields is around 30 cm [11]. This reveals that the magnetic field from given magnet sources varies fast across space. We observe that magnetic signals measured by a smartphone held by a walking user are fairly stable and uncorrupted in open spaces and wide corridors. This is because in such circumstances smartphone sensors are far away from the steel building materials, e.g., in the supporting pillars. Magnetic signals with rich local anomalies are good candidates for loop-closure detection as they are more spatially unique and informative. On the contrary, undisturbed magnetic signals can provide global heading measurements relative to Earth. To our knowledge, these measurements have rarely been used in previous works to create the GraphSLAM front-end.

Algorithm 1 Opportunistic magnetic heading measurements.

Input:
Raw readings: acc. $\{\mathbf{a}_{t_n}\}_{n=1}^{N}$, gyro. $\{\mathbf{w}_{t_n}\}_{n=1}^{N}$, and mag. $\{\mathbf{m}_{t_n}\}_{n=1}^{N}$, number of sensor samples N;
PDR odometry: $\{(\delta\theta_{t_k}, \ell_{t_k})\}_{k=1}^{K}$, number of steps K;
Output:
Opportunistic magnetic headings Θ at certain steps S;

compute orientations $\{\hat{\theta}_{t_n}^{mag}\}_{n=1}^{N}$ relative to Magnetic North using $\mathbf{a}_{t_n}$, $\mathbf{w}_{t_n}$, and $\mathbf{m}_{t_n}$;
for each $n \in [1,\, N]$ **do**
  $sin\theta_n \leftarrow \sin\left(\hat{\theta}_{t_n}^{mag}\right)$, $cos\theta_n \leftarrow \cos\left(\hat{\theta}_{t_n}^{mag}\right)$;
end for
$\Theta \leftarrow \emptyset$, $S \leftarrow \emptyset$, $winsize = 7$
for each $k \in [1,\, K - winsize + 1]$ **do**
  $\langle rot\_angle, arc\_len\rangle \leftarrow getCurvature(\{(\delta\theta_{t_i}, \ell_{t_i})\}_{i=k}^{k+winsize-1})$;
  $std\_sin\theta \leftarrow std\left(\{sin\theta_n\}_{t_k \le t_n \le t_{k+winsize-1}}\right)$;
  $std\_cos\theta \leftarrow std\left(\{cos\theta_n\}_{t_k \le t_n \le t_{k+winsize-1}}\right)$;
  $\bar{\theta} \leftarrow averageAngle\left(\{\hat{\theta}_{t_n}^{mag}\}_{t_k \le t_n \le t_{k+winsize-1}}\right)$;
  if rot_angle, arc_len, $std\_sin\theta$, and $std\_cos\theta$ meet certain thresholding conditions; **then**
    $S \leftarrow S \cup \{k\}$, $\Theta \leftarrow \Theta \cup \{\bar{\theta}\}$;
  end if
end for
return S, Θ;

We opportunistically derive the magnetic headings when we identify that the local magnetic disturbance is neglectable. The proposed algorithm 1 is straight-forward yet effective in practice. The intuition here is that if we keep walking along a straight line (say corridors), the magnetic heading should be constant in addition to bounded variations due to a human's walking pattern. This is the case when geomagnetic fields dominate the instant magnetic signals. We may safely assume



(a) Detected lights (b) Magnetic GPR maps (c) WiFi GPR maps

Fig. 5: Samples of generated signal maps. (a) detected lights marked with red triangles in the light intensity signal and clustered locations shown with red circles on the optimized trajectory; units of colorbar: μT for magnetic maps (b) and dBm for WiFi maps (c).

that the magnetic signals are undisturbed if the magnetic heading measurements are stable over a period of time during a straight line walking.

We adopt Madgwick's AHRS (a.k.a, attitude and heading reference system) solver [19] to compute magnetic headings incorporating magnetometer readings. The derived magnetic headings are error-prone due to local anomalies. To reject errors, we exploit the motion patterns derived by PDR. We iterate over the derived poses and compute the instant curvature of the walking trajectory within a sliding window along with the variations of magnetic headings. We assert a straight walking pattern if the curvature is fairly small. Furthermore, the magnetic heading is assumed to be correct with fair confidence once we identify bounded heading variations. To account for the angle singularity problem (transition from $-\pi$ to π and vice versa), we use the sin and cos values for variation calculation instead. An example is shown in Fig. 3a with opportunistic headings marked with blue arrows.

D. Pose Graph Optimization

The initial pose graph is populated with motion constraints from PDR, loop-closure constraints from magnetic sequence matching, and global heading constraints by opportunistic magnetic heading measurements. Then a pose graph optimization back-end is applied to optimize the poses to be most consistent with the constraints. We use g2o [20] to implement the graph optimization. There are still false positive loop-closures even after the global constraints validation in large-scale environments. The robust kernels are exploited to counter this problem in g2o.

E. Singal Map Generation

The optimized trajectory from the GraphSLAM back-end provides the "ground truth" locations for opportunistic signals collected during walking. The location-labeled signal fingerprints are then used for signal map generation. An example of signal maps is illustrated in Fig. 5. As for WiFi and magnetic fields, GPR proves to be effective in signal map generation. The generated GPR maps include both the mean and variance maps. We do not use the magnetic field map directly for localization since it is difficult to achieve a comprehensive magnetic map using only the measurements collected along the walking traces [11]. However, the variance map for magnetic signals approximately matches the physical floorplan. Considering that wall-constrained particle filters show great effectiveness in reducing PDR drifts, we propose to use the generated magnetic variance map as a pseudo floorplan to help with localization. The lights are first identified by peak detection (Fig. 5a). We assume that each light is revisited a few times, and then cluster these initially detected lights according to their spatial locations.

IV. Localization with Particle Filter using Signal Map

In this section, an approach to using the generated signal map for real-time localization is described.

A. Pseudo-wall and Lights Constraints

We do not assume the availability of floorplans. Instead, the magnetic variance map generated previously by GPR is utilized to provide "pseudo-wall constraints" to help with the heading and step length estimation. As is illustrated by 5b, the variance map of magnetic fields indicates approximately the traversable area. The regions with small and large variances resemble the mapped corridors and unmapped areas, respectively. Aside from the pseudo wall constraints from the magnetic variance map, the lights along the corridor provide more constraints on the step lengths of particles and help speed up the convergence of particle filter.

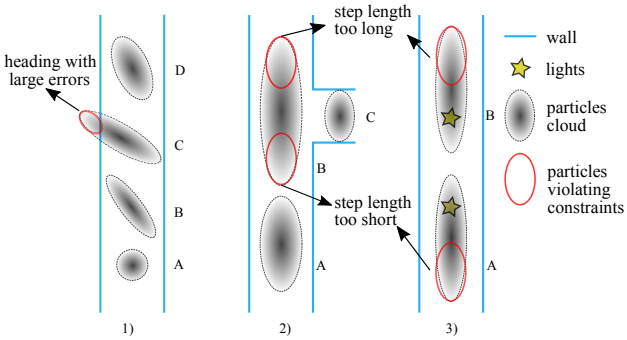


Fig. 6: 1) particles with large heading errors hit the wall and perish while evolving from A to D; 2) particles with either too long or too short step lengths hit the wall when taking a turn; 3) particles violating light constraints get killed at each light detection event.

Fig. 6 demonstrates how the wall constraints and light constraints work on heading and step length estimation. Intuitively, particles with correct headings will survive after a period of straight walking, while those with large heading

errors are most likely to be killed in importance sampling as they violate the wall constraints(Fig. 6-1). The particle cloud is shrunk in the lateral direction (say perpendicular to the walking direction), but elongated due to multi-hypothesis of step lengths. Particles with longer step lengths move faster than those with shorter step lengths. When the user takes a turn, particles will violate the wall constraints and get killed, whose step lengths are either too long or too short(Fig. 6-2). Once a light is detected (Fig. 6-3), particles are killed beyond the lights' coverage vicinity. Similar to the effect of wall constraint during taking turns, only those particles with moderate step lengths are likely to survive.

B. Particle Filtering

Each particle maintains a hypothesis of the state $X = (x, y, \theta, \ell)$ that involves the user's 2D position (x, y), heading θ, and step length ℓ. The particle state gets updated at each step k taken by the user incorporating the motion measurements from PDR, namely step length estimate $\hat{\ell}_k$ and heading change estimate $\delta\hat{\theta}_k$. When new observations(e.g., WiFi scan, lights detection) arrive, weights for each particle are updated and a new set of particles will be generated from the previous through importance sampling if necessary. The state estimation is achieved by taking the centroid of current particle clouds at each step. We accommodate the basic particle filtering procedure.

Initialization: To enable global localization, we use the Maximum Likelihood Estimation (MLE) with WiFi fingerprinting to provide the initial location guess, after which particle filtering is used for position tracking. Even though it is possible to rely on particle filtering without point-mass initialization to achieve global localization, the convergence speed is not guaranteed especially when running on a resource-constrained smartphone. To further improve the filtering convergence, the noisy magnetic heading is used to narrow down the initial heading guess.

Weight Update: The overall weights w_k^i are updated by a few separate importance weights indicating how likely it is that the measurements are consistent with the current particle states. The generic importance weight for the ith particle is $\kappa_i \propto P\left(z_k^i | \hat{X}_k^i\right)$. The weights for the WiFi observation, lights observation, and pseudo wall constraints are shown in the following:

$$\kappa_i^{\text{WiFi}} \propto \exp\left\{ \sum_{j=1}^{N^{\text{WiFi}}} \frac{(\hat{z}_j^{\text{WiFi}} - z_j^{\text{WiFi}})^2}{-2(\hat{\sigma}_j^{\text{WiFi}})^2} \right\}, \tag{1}$$

where $\hat{z}_j^{\text{WiFi}}$ and z_j^{WiFi} are the predicted signal measurement and the observed for the jth WiFi access point (AP), respectively, N^{WiFi} is the number of observed APs in common, and $(\hat{\sigma}_j^{\text{WiFi}})^2$ is the predicted variance from the WiFi variance maps;

$$\kappa_i^{\text{light}} \propto \exp\left\{ \frac{\min\limits_{1 \le j \le N^{\text{light}}} \hat{d}_j^2}{-2\sigma_{\text{light}}^2} \right\}, \tag{2}$$

where $\hat{d}_j$ is the predicted distance of the current particle from the jth light in the lights map, N^{light} is the number of lights, and σ_{light}^2 indicates the coverage of each light;

$$\kappa_i^{\text{wall}} \propto \exp\left\{ \frac{\hat{\sigma}_{\text{magn}}^2}{-2\sigma_{\text{wall}}^2} \right\}, \tag{3}$$

where $\hat{\sigma}_{\text{magn}}^2$ is the predicted magnetic signal variance at the current particle location, and σ_{wall}^2 indicates the penalty on particles that violate the pseudo wall constraints. Intuitively, the penalty is in negative correlation with σ_{wall}^2.

These weights are normalized to preserve a valid probability distribution over all the particles. If some measurements are unavailable for the current step, the corresponding weights are uniformly assigned. For instance, the lights-based weight update is only triggered each time a light spot is detected. In the end, the total weights are updated as per $w_k^i = w_{k-1}^i \times \eta \kappa_i^{\text{WiFi}} \kappa_i^{\text{light}} \kappa_i^{\text{wall}}$, where η is a normalization factor. In our implementation, we choose $\sigma_{\text{light}} = 2\,\text{m}$, $\sigma_{\text{wall}} = 5\,\mu\text{T}$.

V. Evaluation

Six datasets are collected from various indoor environments including office buildings in the HKUST campus and a shopping mall outside, with walking traces lasting from 10 min to 20 min and coverage areas from $2000\,\text{m}^2$ to $4000\,\text{m}^2$ approximately. An Android smartphone (model Samsung Galaxy S5) is used both for data collection and localization evaluation throughout this paper. The surveyor is asked to simply hold the phone and walk.

A. Mapping Results

1) Loop detection with validation: The statistics on loop-closure correctness is summarized in TABLE I as per the total number of True Positives (TPs) against False Positives (FPs) for each dataset using the final optimized trajectories as ground truth. False positives are reduced to the minimum after the validation which proves the effectiveness of the proposed WiFi similarity-based validation method. We choose a strict similarity checking threshold that successfully rejected all the FPs yet at the cost of throwing away a certain portion of TPs. A less strict threshold could be feasible if a more robust back-end is used.

TABLE I: Statistics on TPs/FPs of the detected loop-closures before and after WiFi similarity validation

Dataset	CYT	AC3-1	AC3-2	AC4-1	AC4-2	SOGO
Before	$\frac{2156}{289}$	$\frac{1038}{873}$	$\frac{322}{215}$	$\frac{1162}{2430}$	$\frac{610}{80}$	$\frac{223}{157}$
After	$\frac{1366}{0}$	$\frac{916}{0}$	$\frac{322}{0}$	$\frac{591}{0}$	$\frac{610}{0}$	$\frac{113}{0}$

2) Trajectory estimation: Fig. 7 only show the results for three more datasets in addition to those already shown in Fig. 1 and Fig. 5 due to limited space. The raw PDR paths reveal the drift-prone nature of PDR, especially with the low-cost smartphone sensors. However, the relative motion patterns are correctly measured such as the long straight walking and left/right/U-turns. As shown in Fig. 7, optimized trajectories are not aligned correctly to the Earth frame without magnetic

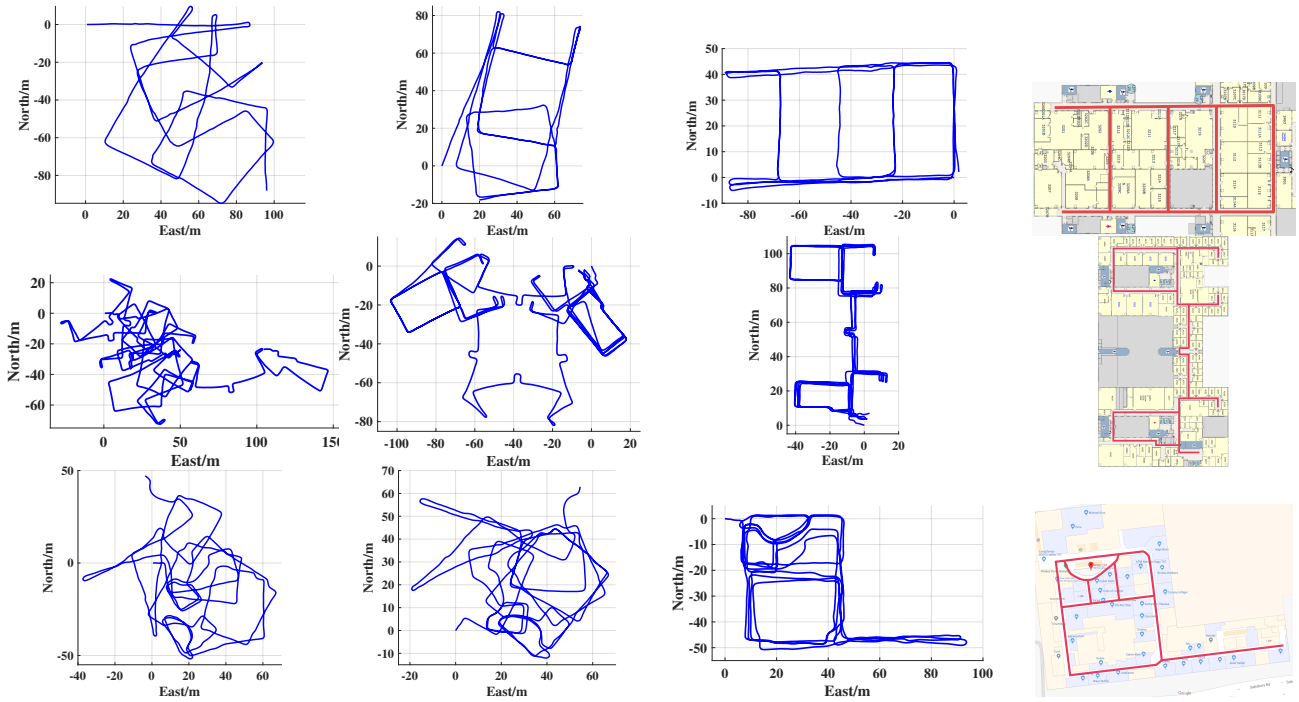


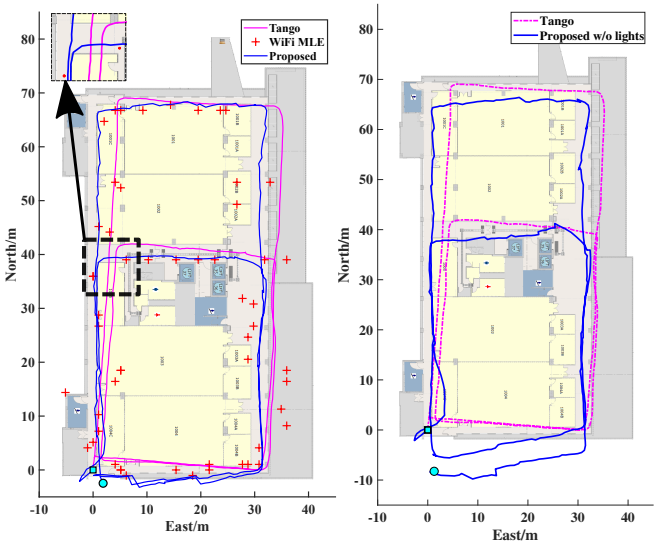
Fig. 7: From top to bottom: **AC3-2**, **AC4-1**, and **SOGO**. From left to right: 1) raw PDR-derived trajectories before optimization, 2) optimized trajectories without incorporating opportunistic magnetic heading measurements, 3) the final optimized trajectories by the proposed method, and 4) reference paths on the floorplan. The floorplan is used only for visuallization throughout this paper.

headings measurements. In some cases, they are not even globally consistent rendering them useless for signal map generation. However, the final optimized trajectories show considerable consistency to physical floorplans.

TABLE II: Statistics on the ROE errors

Dataset	CYT	AC3-1	AC3-2	AC4-1	AC4-2	SOGO
Error/m	0.87	2.44	2.99	8.10	6.67	22.32

Due to the lack of ground truths, we characterize "how correctly the user returns back to the origin" as Return-to-Origin Error (ROE). The statistics are shown in TABLE II. The ROE error 0.87 m is considerably small for **CYT**. We credit this to the rich magnetic anomalies induced by the construction materials in the building, working electronic equipment from inside labs, and narrow corridors that confine the walking trajectory taken each time. The ROE errors for other office areas range from 2.44 m to 8.10 m mainly due to varying coverage area and amounts of magnetic anomalies. The ROE error 22.32 m for **SOGO** is significantly large due to poor magnetic anomalies found in this area with wide corridors and inconsistent walking trajectories caused by the random interferences from the crowded people around.



(a) (b)

Fig. 8: Illustration of the localization results with (a) and without (b) lights measurements. The testing walk originates from the left-bottom corner with the square marking the start point and the circle marking the endpoint.

B. Localization Results

Localization is tested with the generated signal maps from the dataset **CYT**, and a real-time localization demo is shown in the supplementary video. 1000 particles are used throughout this experiment. The results are highlighted in Fig. 8a by the blue lines which show sound consistency to the reference path. The WiFi MLE results are shown as red plus signs that are scattered through the region due to lack of motion and wall constraints. The trajectory from a Project Tango device is also plotted against the floorplan for visual comparison. We observe clear drifts in Tango's path as is highlighted by the magnified view in Fig. 8a while

the results of the proposed method are constrained well in the corridor region. This reveals that the proposed particle filtering method with signal maps could achieve consistent localization over time. Localization without light constraints is shown in Fig. 8b. The performance degradation is due to the lack of enough constraints on step lengths.

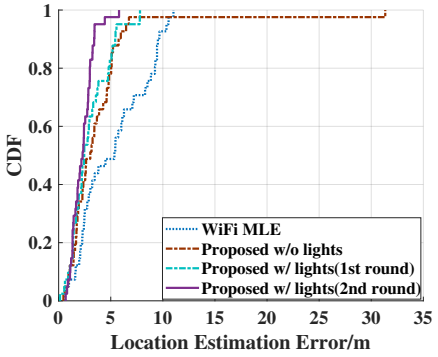


Fig. 9: Localization error CDF.

We placed 25 check-points along the corridor and derived their ground truth locations using a floorplan. The cumulative distribution function plot (CDF) is then generated for error analysis, as shown in Fig. 9. The 50-percentile and 90-percentile accuracies are summarized in TABLE III for three localization algorithms namely WiFi MLE, particle filtering without and with light measurements. We observe significant large errors (32 m, see the dash-dot line in Fig. 9) for the proposed method with light detection disabled. This is due to the recursive nature of particle filtering where real-time localization results normally improve with time. The initial error is likely to be unacceptable but will be reduced to a bounded value once the particle filter coverages. We take two consecutive rounds for the final proposed method. Particle filtering with the generated WiFi maps and the "pseudo floorplan" shows significantly better performance than WiFi MLE. The performance is further improved with additional constraints from the lights map on step lengths.

TABLE III: Localization accuracy statistics

Algo.	MLE	PF-w/o	PF-w/ (1st)	PF-w/ (2nd)
50-ile Error/m	5.35	3.04	2.45	2.30
90-ile Error/m	9.52	5.97	5.46	3.41

VI. Conclusions

In this paper, a low-cost indoor mapping and localization solution was proposed using the opportunistic signals from ambient environments with a smartphone. The system was split into two blocks namely offline mapping within the classical GraphSLAM framework and online localization with Bayesian filtering using the generated signal maps. The proposed WiFi similarity validation for loop-closure detection showed great effectiveness in our implementation where false positives were significantly rejected. The incorporation of opportunistic magnetic heading measurements improved the GraphSLAM-based trajectory optimization and helped generate more consistent maps. Real-time localization on smartphones was achieved through particle filtering with the generated signal maps. The pseudo-wall constraints from the magnetic fields GPR variance map and the lights constraints worked well in confining the particle clouds. The proposed system was thoroughly evaluated on several datasets collected from both the in-compass office buildings and off-compass public areas where globally consistent trajectories were reached. Real-time localization performance was evaluated in an office building with a coverage area of $2000\,\mathrm{m}^2$. The 50- and 90-percentile accuracies achieved on a smartphone were 2.30 m and 3.41 m respectively.

References

[1] L. Pei, M. Zhang, D. Zou, R. Chen, and Y. Chen, "A survey of crowd sensing opportunistic signals for indoor localization," *Mobile Information Systems*, vol. 2016, 2016.

[2] Y. Sun, M. Liu, and M. Q.-H. Meng, "WiFi signal strength-based robot indoor localization," in *Proc. ICIA*. IEEE, 2014, pp. 250–256.

[3] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magicol: Indoor localization using pervasive magnetic field and opportunistic WiFi sensing," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1443–1457, 2015.

[4] R. Faragher, C. Sarno, and M. Newman, "Opportunistic radio SLAM for indoor navigation using smartphone sensors," in *Proc. PLANS*. IEEE, 2012, pp. 120–128.

[5] Q. Liang and M. Liu, "Plugo: A VLC systematic perspective of large-scale indoor localization," *arXiv preprint arXiv:1709.06926*, 2017.

[6] C. Wu, Z. Yang, and Y. Liu, "Smartphones based crowdsourcing for indoor localization," *IEEE Trans. Mobile Comput.*, vol. 14, no. 2, pp. 444–457, 2015.

[7] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Proc. Mobicom*. ACM, 2012, pp. 293–304.

[8] C. Gao and R. K. Harle, "Semi-automated signal surveying using smartphones and floorplans," *IEEE Trans. Mobile Comput.*, 2017.

[9] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "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.

[10] J. Huang, D. Millman, M. Quigley, D. Stavens, S. Thrun, and A. Aggarwal, "Efficient, generalized indoor WiFi graphSLAM," in *Proc. ICRA*. IEEE, 2011, pp. 1038–1043.

[11] C. Gao and R. Harle, "Sequence-based magnetic loop closures for automated signal surveying," in *Proc. IPIN*. IEEE, 2015, pp. 1–12.

[12] S. Wang, H. Wen, R. Clark, and N. Trigoni, "Keyframe based large-scale indoor localisation using geomagnetic field and motion pattern," in *Proc. IROS*. IEEE, 2016, pp. 1910–1917.

[13] I. Deutsch, M. Liu, and R. Siegwart, "A framework for multi-robot pose graph SLAM," in *Proc. RCAR*. IEEE, 2016, pp. 567–572.

[14] M. Liu, F. Pomerleau, F. Colas, and R. Siegwart, "Normal estimation for pointcloud using GPU based sparse tensor voting," in *Proc. ROBIO*. IEEE, 2013, pp. 91–96.

[15] R. Harle, "A survey of indoor inertial positioning systems for pedestrians," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1281–1293, 2013.

[16] K. Qiu, F. Zhang, and M. Liu, "Visible light communication-based indoor localization using Gaussian process," in *Proc. IROS*. IEEE, 2015, pp. 3125–3130.

[17] M. Liu, L. Wang, and R. Siegwart, "DP-Fusion: A generic framework for online multi sensor recognition," in *Proc. MFI*, 2012, pp. 7–12.

[18] H. Weinberg, "Using the ADXL202 in pedometer and personal navigation applications," *Analog Devices AN-602 application note*, vol. 2, no. 2, pp. 1–6, 2002.

[19] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. ICORR*. IEEE, 2011, pp. 1–7.

[20] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A general framework for graph optimization," in *Proc. ICRA*. IEEE, 2011, pp. 3607–3613.