Infrastructure-Free Global Localization in Repetitive Environments: An Overview

Zhenyu Wu, Jun Zhang, Yufeng Yue, Mingxing Wen, Zichen Jiang, Haoyuan Zhang and Danwei Wang

Abstract— **Repetitive environment is a challenging scenario for mobile robot global localization due to its highly similar structures and lack of distinctive features. Existing solutions in such environments rely heavily on pre-installed infrastructures, which are neither flexible nor cost-effective. Besides, few of the previous research have been focused on the implementation of infrastructure-free localization approaches in repetitive scenarios. Thus, this paper serves as a survey to investigate the problem of infrastructure-free mobile robot global localization with low-cost and efficient sensors in repetitive environments. Three of the most popular infrastructure-free localization methods, namely LiDAR-based localization (LBL), vision-based localization (VBL), and magnetic field-based localization (MFL), are analyzed and evaluated. Extensive global localization experiments are conducted in real-world repetitive scenarios and the results demonstrate that VBL methods perform slightly better than LBL and MFL methods. The overall evaluations indicate that infrastructure-free global localization in repetitive environment is still a challenging problem which deserves more research efforts to develop new solutions.**

I. Introduction

Localization empowers the robots with the capabilities like navigation, security, rescue, and intelligent services. Over the past few decades, localization in indoor or GPS-denied environments has been recognized as a key research field in mobile robotics. Specifically, localization techniques fall into three basic categories: 1) Behavior-based approaches; 2) Infrastructure-based approaches; 3) Infrastructure-free approaches. Behavior-based approaches [1], [2] rely on the interaction of mobile robots with the environment/human to accomplish localization tasks. Although behavior-based approaches can be useful for certain tasks, their geometrical localization ability is limited due to their implicit navigation capability in the sensor/instruction history. Infrastructure-based methods [3], [4] rely on the recognition of infrastructures to accomplish the geometrical localization. Despite the relatively accurate localization that infrastructure-based approaches can achieve, they require both pre-installation and maintenance of the infrastructures in the environment. In contrast, infrastructure-free approaches [5]–[28] tend to utilize all available sensor information to update the robot pose. Most of these localization processes are done by matching sensor observations with a pre-built environment map. Thus, infrastructure-free approaches can take advantage of all available features to help localize the robot.

*This work was supported by the ST Engineering-NTU Corporate Laboratory through the NRF Corporate Laboratory@University Scheme.

The authors are with the ST Engineering-NTU Robotics ADVANCe Lab, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. E-mail: zhenyu002@e.ntu.edu.sg

Global localization in repetitive environments, such as corridors and warehouses, remains a challenging problem due to the lack of distinctive features. Existing solutions rely heavily on pre-installed infrastructures, such as radio beacons, QR-codes, and guiding tapes. The costs for pre-installing and maintaining such infrastructures are high, which is also not scalable [9], [25]. In contrast, infrastructure-free global localization approaches are cost-effective, and scalable to large-scale environments. Infrastructure-free localization approaches include different types of sensor modalities, such as LiDAR, camera, sonar, and magnetic sensor. We focus on the three most cost-efficient and accurate sensors, namely the 2-D LiDAR, camera, and magnetic sensor. LiDAR-based localization (LBL) approaches [5]–[13] are generally adopted for mobile robot localization due to their high accuracy as well as robustness to illumination and viewpoint changes. Among the LBL methods, the most widely utilized approaches assume Markov models and use Kalman filters or particle filters [9], [29]. Meanwhile, the vision-based localization (VBL) methods [15]–[20] are also popular due to their low cost and high efficiency. VBL methods are dependent on visual descriptors, where local feature descriptors such as the ORB [30] extract notable parts of the image, while global or whole-image descriptors like WI-SURF [31] process the image as a whole. Furthermore, the magnetic field-based localization (MFL) [22]–[26] is another viable alternative, where fingerprints are normally utilized to index locations. The pervasiveness and robustness of MF make it a feasible choice for localization.

In this paper, we investigate the problem of infrastructure-free mobile robot localization with cost-efficient sensors in repetitive environments, focusing on their global localization capability. The contributions of our work are as follows:

- Infrastructure-free localization methods with different sensor modalities are analyzed and evaluated, including the LBL, VBL, and MFL approaches. The overall evaluations show that VBL methods perform slightly better than LBL and MFL approaches.
- LBL methods are robust to lighting variations and have comparable accuracy with VBL methods, but they take more iterations and longer distances to be localized.
- The effectiveness of MFL approach has been proved in relatively large-scale repetitive environments. The evaluations demonstrate that MFL method has the most robust performance but relatively low accuracy.

The rest of this paper is organized as follows. Section II reviews the related work. Section III details different kinds

978-1-7281-5414-5/20/$31.00 ©2020 IEEE

of infrastructure-free global localization methods. Section IV presents experiments, results and evaluations. Eventually, Section V draws the conclusion.

II. Related Work

A. LiDAR-based Localization

In general, the LBL problem is modeled as a Bayesian filtering problem by matching LiDAR measurements with a globally consistent environment map. Probabilistic-based algorithms including multiple hypothesis tracking [5], [6] and particle filters [7], [9] are among the most popular methods to solve this problem. Besides, the integrated system presented in [9] combines the particle filters with scan matching to improve the localization accuracy. Moreover, laser-based deep learning framework [11] was proposed to solve the loop closure detection problem.

B. Vision-based Localization

VBL approaches are feasible choices for mobile robot localization, where visual place recognition plays a critical role [18]. The survey by [16] demonstrated that the appearance-based place recognition method, which is image-to-image matching, performs better than image-to-map or map-to-map approaches. Moreover, a robust system was presented in [15] to integrate Monte Carlo Localization with an image retrieval system, which is based on image features that are translation-invariant and scale-invariant. More recently, the *ORB-SLAM2* approach [19] developed upon ORB descriptors allows loop closure and relocalization, which leverages matches to visual odometry and matches to pre-built feature map.

C. Magnetic Field-based Localization

Magnetic field (MF) has noticeable signatures for different places. Researchers have proved the stability of the disturbed MF over a long period of time and demonstrated the feasibility of using the MF alone for indoor localization [22], [24]. Besides, MF-based localization is also utilized for challenging applications, where [23] demonstrated the capability of perching aerial robots on power lines.

In summary, the performance of existing solutions for localization in repetitive environments is not satisfying. The other aforementioned LBL and VBL approaches operate well in regular environments with sufficient distinctive features, but their performance can deteriorate or even fail in repetitive environments. Besides, the effectiveness of MFL is yet to be proved in repetitive environments. Hence, this paper aims to investigate the above mentioned problems.

III. Infrastructure-free Global Localization Methods

This paper focuses on the 2-D global localization problem, meaning that the robot pose $\mathbf{x}_t$ at time step $t \in \mathbb{Z}_{\geq 0}$ is a random variable which satisfies $\mathbf{x}_t = [x, y, \theta]^T \in \mathbb{R}^2 \times \mathbb{S}$. In general, mobile robots can be accurately localized in regular environments with sufficient unique features. However, repetitive environments contain highly similar structures with few distinctive features. Thus, infrastructure-free localization methods can be inaccurate or even fail in such environments. Details of the evaluated methods are presented below.

A. LiDAR-based Localization

Most of the LBL methods assume underlying Markov models. In probabilistic terms, localization can be considered as the process of estimating the current state of the robot by maintaining a posterior distribution $p(\mathbf{x}_t|\mathbf{z}_t, \mathbf{u}_t, \mathbf{m})$ of the robot pose, which is updated by the recursive equation:

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}, \mathbf{m}) = \eta\, p(\mathbf{z}_t|\mathbf{x}_t, \mathbf{m}) \int p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t)\, p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t-1}, \mathbf{m}) d\mathbf{x}_{t-1} \quad (1)$$

where η is a normalization constant, $\mathbf{z}_t$ is the sensor observation, $\mathbf{u}_t$ is the control input, and $\mathbf{m}$ denotes the map.

In Monte Carlo Localization, multi-modal distributions can be approximated. On the other hand, scan matching leverages a simple Gaussian distribution. But interestingly, both of the two methods utilize the same general technique which can proceed in the following two phases:

Prediction Phase In this phase, a motion model is utilized to predict the new robot pose $\mathbf{x}_t$ from the previous pose $\mathbf{x}_{t-1}$. This process can be represented as:

$$p(\mathbf{x}_t) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t)\, p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t-1}, \mathbf{m}) d\mathbf{x}_{t-1} \quad (2)$$

where $p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t)$ is the state transition probability (i.e. motion model) which represents the probability of robot pose being at $\mathbf{x}_t$ given it was at $\mathbf{x}_{t-1}$ and executed control $\mathbf{u}_t$.

Update Phase In this phase, observation model is incorporated to update the $\mathbf{x}_t$ according to the Bayes theorem:

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}, \mathbf{m}) = \eta\, p(\mathbf{z}_t|\mathbf{x}_t, \mathbf{m}) p(\mathbf{x}_t) \quad (3)$$

where $p(\mathbf{z}_t|\mathbf{x}_t, \mathbf{m})$ (i.e. observation model) denotes the likelihood of perceiving observation $\mathbf{z}_t$ given that the robot is at pose $\mathbf{x}_t$. By this way, the posterior distribution $p(\mathbf{x}_t|\mathbf{z}_t, \mathbf{u}_t, \mathbf{m})$ is maintained and updated. Moreover, the initial state of the robot $\mathbf{x}_0$ is assumed to be available as $p(\mathbf{x}_0)$, which is supposed to be a uniform distribution in the global localization case.

*1) **Adaptive Monte Carlo Localization**:* In general, Adaptive Monte Carlo Localization (*AMCL*) [5], [7] leverages laser range as input. It is a method based on sampling and importance-based resampling of the particles. Particle filters represent the posterior distribution $p(\mathbf{x}_t|\mathbf{z}_t, \mathbf{u}_t, \mathbf{m})$ by a recursive set S_t of n weighted samples drawn from it:

$$S_t = \{s_t^i \,|\, i = 1, ..., n\} = \{\langle \mathbf{x}_t^{(i)}, w_t^{(i)} \rangle \,|\, i = 1, ..., n\} \quad (4)$$

where each sample $\mathbf{x}_t^{(i)} = (x_t^{(i)}, y_t^{(i)}, \theta_t^{(i)})^T$ represents a possible state of the robot and $w_t^{(i)}$ denotes the corresponding numerical factor called importance weights. In the initialization phase, the samples are uniformly distributed over the possible positions of the 2-D space. Meanwhile, $\sum_{i=1}^{n} w_t^{(i)} = 1$ is assumed for consistency.

AMCL can be generalized into three phases, namely prediction, update, and adaptive sample set sizes.

a) Prediction Phase: In the first phase, the motion model is applied to each sample s_{t-1}^i by drawing one sample ${s'}_t^i$ from the distribution $p(\mathbf{x}_t|s_{t-1}^i, \mathbf{u}_t)$. Then a new sample set ${S'}_t$ is generated which approximates random samples from the predictive distribution $p(\mathbf{x}_t)$.

b) Update Phase: In this phase, the observation model is incorporated to re-weight the samples in ${S'}_t$ by the weight $w_t^{(i)} = p(\mathbf{z}_t|{s'}_t^i, \mathbf{m})$. Then S_t can be obtained by resampling the samples with higher probability from the re-weighted set.

c) Adaptive Sample Set Sizes: To achieve a desired level of accuracy, the number of required samples varies in different scenarios. In the practical process of *AMCL*, the sample set size is determined adaptively based on an error estimate utilizing the Kullback–Leibler divergence (KLD)-sampling [7]. If the location predicted by odometry matches well with the sensor observations, the sample density is focused on a small part of the state space. Each individual weight $w_t^{(i)}$ is large and the sample set will be small, as is typically the pose tracking case. However, if the sensor observations carries large amount of noises, the state uncertainty will be high. Thus, the individual weight $w_t^{(i)}$ is small and the sample set will be large, as is typically the global localization case or when the robot lost track of its location.

*2) **Scan Matching**:* Apart from *AMCL*, the scan matching-based method is another category of the LBL methods [8]–[11]. The key process of scan matching is to translate or rotate a laser scan in the way that maximizes the overlap between the laser sensor observations and a prior map. Normally, scan matching is implemented with a pre-obtained 2-D occupancy grid map (OGM), which is built by 2-D LiDAR-based Simultaneous Localization and Mapping (SLAM). In the scan matching process, an initial estimation of the robot pose is essential. Single Gaussian distribution is utilized to model the robot pose and its update from scan matching. With the implementation of the efficient Kalman filtering (KF) to calculate the update phase, robot pose can be derived with high accuracy.

In general, Extended Kalman Filter (EKF) is adopted in scan matching-based approaches. For each time step t, the robot pose and the error covariance matrix can be depicted as $\mathbf{x}_t = (x_t, y_t, \theta_t)^T$ and $\Sigma_{\mathbf{x}_t}$, respectively. $\mathbf{u}_t = (\Delta d_t, \Delta\theta_t)^T$ represents the translational and rotational movement measured by encoders from state $\mathbf{x}_{t-1}$ to state $\mathbf{x}_t$. Then the robot pose and error covariance matrix are predicted according to:

$$\mathbf{x}_t{}' = \langle \mathbf{x}_{t-1}, \mathbf{u}_t \rangle = \begin{pmatrix} x_{t-1} \\ y_{t-1} \\ \theta_{t-1} \end{pmatrix} + \begin{pmatrix} \Delta d_t \cos\theta_{t-1} \\ \Delta d_t \sin\theta_{t-1} \\ \Delta\theta_t \end{pmatrix} \quad (5)$$

$$\Sigma'_{\mathbf{x}_t} = \nabla\mathbf{x}_{t-1} \cdot \Sigma_{\mathbf{x}_{t-1}} \cdot \nabla\mathbf{x}_{t-1}^T + \nabla\mathbf{u}_t \cdot \Sigma_{\mathbf{u}_t} \cdot \nabla\mathbf{u}_t^T \quad (6)$$

Subsequently, a pose update $\mathbf{x}_s$ together with an error covariance matrix Σ_s are obtained from scan matching. Then the robot pose and covariance matrix are updated by:

$$\mathbf{x}_t = (\Sigma_{\mathbf{x}_t}^{'-1} + \Sigma_s^{-1})^{-1} \cdot (\Sigma_{\mathbf{x}_t}^{'-1} \cdot \mathbf{x}_t{}' + \Sigma_s^{-1} \cdot \mathbf{x}_s) \quad (7)$$

$$\Sigma_{\mathbf{x}_t} = (\Sigma_{\mathbf{x}_t}^{'-1} + \Sigma_s^{-1})^{-1} \quad (8)$$

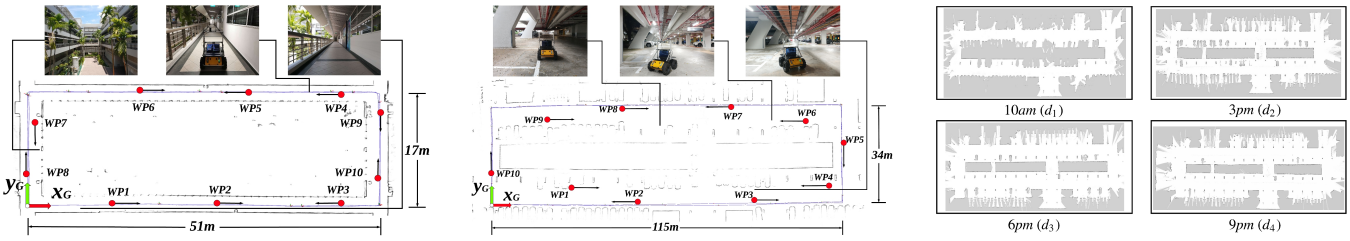
B. Vision-based Localization

Vision-based localization is often achieved through visual place recognition. A typical VBL system consists of three key components [18]: 1) an *image processing* module; 2) a *map*; 3) a *belief generation* module. Among the VBL methods, *ORB-SLAM2* is the most popular one and represents the state-of-the-art. *ORB-SLAM2* is a feature-based SLAM as well as relocalization system for monocular, stereo and RGB-D cameras [19]. The *ORB-SLAM2* system mainly has three parallel threads: 1) By matching features to the local map and minimizing the reprojection error, every frame of the camera is localized by the tracking; 2) Performing local bundle adjustment, the local map is managed and optimized by the local mapping; 3) Large loops are detected by the loop closing and the accumulated drift is corrected by pose-graph optimization.

Based on DBoW2 [17], the *ORB-SLAM2* system has incorporated a bags of words place recognition module to perform loop detection, relocalization for tracking failure, or reinitialization in an already mapped area. Visual vocabulary (i.e. visual words) is a discretization of the descriptor space. With the ORB descriptors extracted from large amount of general images, the vocabulary is created offline. Then the same vocabulary has a good capability to be utilized for various environments. In the localization mode, the local mapping and loop closing threads are disabled. By tracking as well as relocalization, the camera is consistently localized. Besides, the tracking matches to both visual odometry (VO) and feature point map (FPM) in this mode. VO matches ORB descriptors in the current frame with 3-D points created in the previous frame, then matches to FPM are implemented to eliminate the drift of localization.

C. Magnetic Field-based Localization

In general, the magnetic field $\mathbf{B} \in \mathbb{R}^3$ can be modeled as a 3-D vector field $\mathbf{B} = [B_x, B_y, B_z]^T$, where B_x, B_y, and B_z represent magnetic intensities in X, Y and Z direction, respectively. The measured B_x is always in the same direction with the heading of the robot. The earth magnetic field (EMF) sets a background for the MF, but anomalies caused by ferromagnetic objects deflect the EMF, which makes the MF unique at the local scale. For each MF data, the 3-D vector plus the 1-D magnitude are utilized to achieve higher matching accuracy. Since it is impossible for the robot to collect the MF data on all directions at every single location, the MF fingerprint database (MFFD) is built with the mobile robot teleoperated on different desired routes which are along the trajectories parallel to the straight fences or walls of the environments. So at each location, the MF data of only two directions (0° and 180°) are measured and stored, thus greatly reducing the workload for building the MFFD. By recording 4-D MF data, corresponding 2-D location, and robot body frame, the MFFD can be constructed. Then the particle filter [25] or *k*-NN classification [26] algorithms can be adopted to match the online MF data with the MFFD based on the Euclidean distance metric, thus accomplishing the MF-based localization task.



(a) Dataset 1: Rectangular corridor (b) Dataset 2: Semi-indoor carpark (c) OGMs of the carpark

Fig. 1. (a) and (b) demonstrate the occupancy grid maps (OGMs) of the rectangular corridor and semi-indoor carpark, respectively. Red dots denote waypoints (WP, starting locations) and black arrows represent the initial heading direction of the robot. Blue circle line denotes one typical robot trajectory. (c) depicts the OGMs of the carpark during different times of different days, where the first map is almost full of cars and the last map is almost empty. LBL is implemented on the four OGMs simultaneously and we adopt the result which has the minimum localization error.

IV. Experimental Evaluation

A. Experimental Setup

1) ***Platform****:* Fig. 2 demonstrates the experimental platform. The robot platform is a Husky A200 UGV. A Xsens MTi-20 IMU is mounted at the top and a ZED stereo camera is mounted at the middle of the central frame, respectively. A Hokuyo UTM-30LX 2-D Scanning Rangefinder is mounted at the front of the robot.

2) ***Experimental Environment****:* Two representative repetitive environments, one rectangular corridor and another semi-indoor carpark within Nanyang Technological University, are selected for experiments. Their characteristics are:

a) Dataset 1 - Rectangular Corridor: Rectangular-shaped corridor with repetitive and symmetric settings like the pillars and fences.

b) Dataset 2 - Semi-indoor Carpark: Mixed indoor and outdoor carpark with repetitive and symmetric settings like walls and pillars, including semi-static objects like the cars and pedestrians as well.

3) ***OGM & FPM & MFFD Building****:* 2-D LiDAR-based SLAM algorithms *Hector-SLAM* [8] and *Cartographer* [10] were both implemented to build the OGMs, where high portion of them will be failed due to the repetitiveness of the environments and lack of distinctive features. The OGM utilized in our experiments was the most accurate one chosen among the built OGMs, as shown in Fig. 1a and Fig. 1b. As demonstrated in Fig. 1c, different OGMs were built for the carpark environment due to its semi-static characteristic [12]. Besides, FPMs at day-time and night-time were both built by the *ORB-SLAM2*. Note that similar with the OGMs, it was quite hard to build the correct FPMs in such repetitive environments. As demonstrated in Fig. 3, preliminary test results showed that around 60% and 70% of the *LiDAR-based SLAM* and *ORB-SLAM2* algorithms failed in such repetitive environments, respectively. The resolution of the OGM and FPM was both set to $5cm$. Moreover, the robot was manually driven while carrying out 2-D LiDAR-based particle filter [5] to estimate the corresponding location of the MF fingerprint. MF fingerprint was collected every $5cm$ along the trajectories. There were 43,542 and 118,688 of fingerprints in the MFFD of two datasets, respectively.

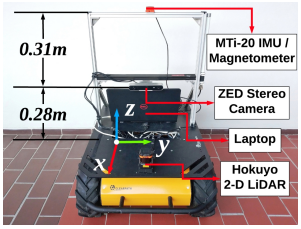


Fig. 2. *Husky A200* robot platform. To avoid possible electromagnetic interference, the IMU is installed 0.31m and 0.59m higher than the ZED stereo camera and the robot platform, respectively.

B. Evaluation Protocol

1) ***Comparison Baseline****:* Five different approaches are evaluated, which are: 1) *AMCL* [5]; 2) *Cartographer* [10]; 3) *ORB-SLAM2(Day)* [19]; 4) *ORB-SLAM2(Night)* [19]; 5) *PF-Mag*. Since no prior information will be given in the global localization, the initial pose was set with default parameters for *AMCL* and *Cartographer* by their own algorithm. *ORB-SLAM2(Day)* and *ORB-SLAM2(Night)* are evaluated based on day-time and night-time FPMs, respectively. Besides, *PF-Mag* is the particle filter (PF)-based MFL approach, which is

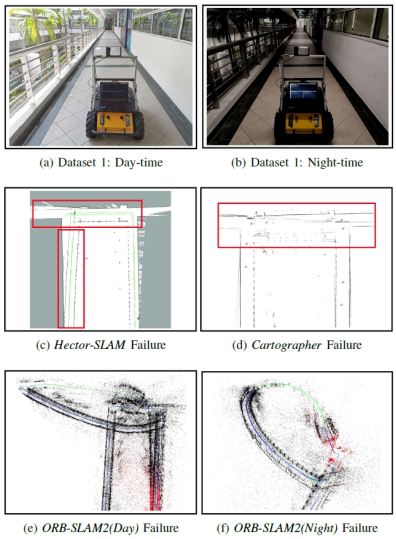
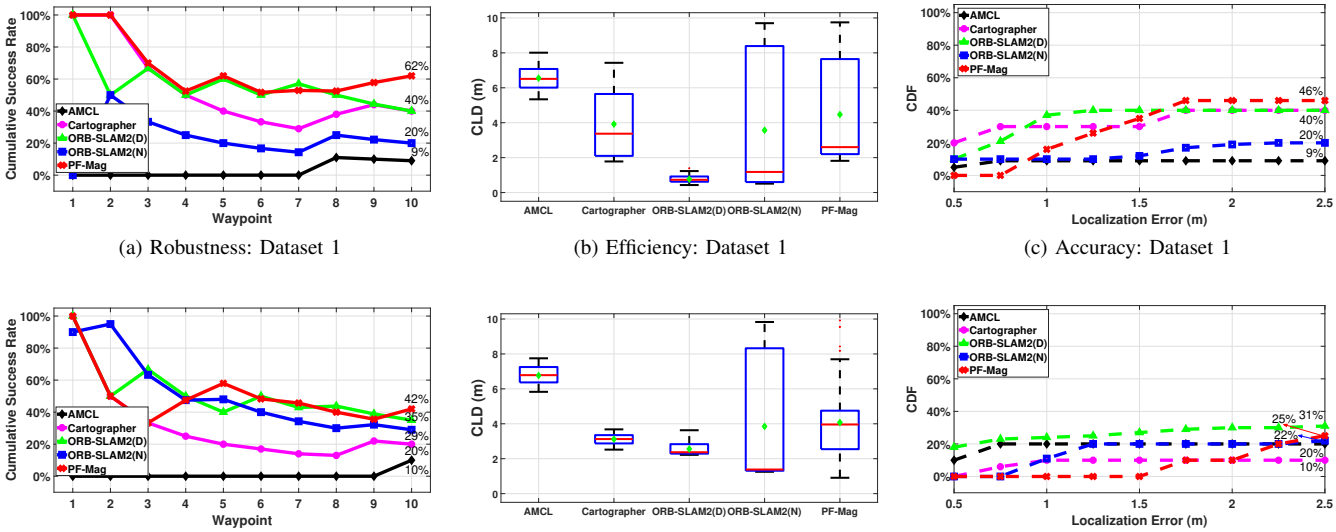


Fig. 3. (a) and (b) demonstrate the illumination change in day-time and night-time at the same location. (c)/(d) and (e)/(f) show examples of the failure when building OGMs and FPMs of the corridor environment, respectively, due to highly similar structures with few distinctive features.



(a) Robustness: Dataset 1 (b) Efficiency: Dataset 1 (c) Accuracy: Dataset 1

(d) Robustness: Dataset 2 (e) Efficiency: Dataset 2 (f) Accuracy: Dataset 2

Fig. 4. **Left**: (a) and (d) are the cumulative localization success rate shown in line graph. **Middle**: (b) and (e) are the average correct localization distance (*CLD*) shown in boxplots. The mean value, median value, and the outlier are denoted in green diamond, horizontal red line and red cross mark, respectively. **Right**: (c) and (f) are the Cumulative Distribution Function (*CDF*) plot of the localization error.

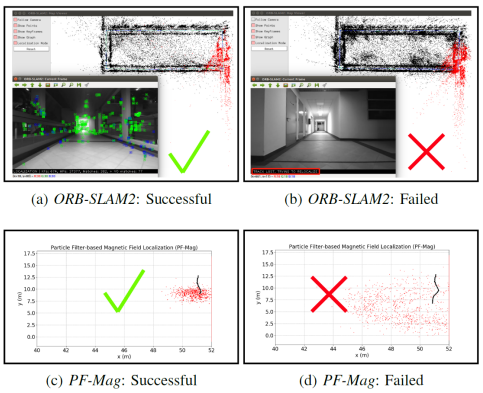
implemented based on the general idea of non-legged robot localization part of [25]. As shown in Fig. 1a and Fig. 1b, ten waypoints (WPs) were randomly selected for each dataset to evaluate the global localization performance. The ground truth locations of the WPs were manually measured by a tape ruler. For each evaluated approach of each WP, ten individual random tests were conducted. Besides, tests for odd-numbered WPs were done at day-time, while tests for even-numbered WPs were done at night-time so as to evaluate the robustness to illumination changes.

*2) **Evaluation Protocol**:* Mobile robot global localization performance is evaluated by the following three criterions: 1) Robustness; 2) Efficiency; and 3) Accuracy. Robustness is evaluated by the cumulative localization success rate, which is the average success rate over the ten WPs. Efficiency is assessed by measuring the actual distance the robot has travelled to successfully localize itself (correct localization distance - (*CLD*)). The localization is considered failed if the *CLD* is longer than $10.0m$. Furthermore, accuracy is visualized by the cumulative distribution function (*CDF*) plot of the absolute localization error. The localization is considered failed if the error is larger than $10.0m$.

C. Experimental Evaluations

The rectangular corridor environment is depicted in Fig. 1a. Firstly, the localization success rate is shown in Fig. 4a, where the *PF-Mag* yields the highest rate of **62%** and the other approaches yield rates ranging from 9% to 40%. As for the semi-indoor carpark depicted in Fig. 1b, the localization success rate is shown in Fig. 4d, where the *PF-Mag* yields the highest rate of **42%** and the other approaches yield rates ranging from 10% to 35%. Regarding the localization efficiency, it is obvious that *ORB-SLAM2(Day)* has the best *CLD* as shown in Fig. 4b and Fig. 4e . Note that the efficiency of *ORB-SLAM2(Night)* is worse than *Cartographer* due to the poor illumination at night-time. As for the accuracy evaluations shown in Fig. 4c and Fig. 4f, the performance of *PF-Mag* is comparable with *ORB-SLAM2(Day)*, both of which outperform *ORB-SLAM2(Night)* and LBL approaches. Besides, examples of the successful and failed localization cases of *ORB-SLAM2* and *PF-Mag* are demonstrated in Fig. 5.

In summary, it is observed from the experimental results that VBL method *ORB-SLAM2(Day)* performs slightly better than LBL methods, including the *AMCL* and *Cartographer*. However, the results also reveal that the localization performance of *ORB-SLAM2(Night)* is much worse than the performance of *ORB-SLAM2(Day)*, demonstrating that VBL method is brittle to the illumination changes. Moreover, *ORB-SLAM2* suffers from viewpoint changes where the robot viewpoint is not exactly same with the previous one when



(a) *ORB-SLAM2*: Successful (b) *ORB-SLAM2*: Failed

(c) *PF-Mag*: Successful (d) *PF-Mag*: Failed

Fig. 5. Demonstrations of successful and failed localization cases in WP#9 of dataset 1. (a) and (b) demonstrate the results of *ORB-SLAM2*. (c) and (d) show the results of *PF-Mag*, where red points and black curves denote the particles and robot trajectories, respectively.

building the map, as the sudden drop of the cumulative success rate of *ORB-SLAM2* at WP#4 of both datasets shown in Fig. 4a and Fig. 4d. Furthermore, the offline preparation cost of *ORB-SLAM2* is higher than the other approaches, since two sets of FPMs, both day-time and night-time, have to be built. Besides, illumination changes seem to have no influence on LBL methods. In general, the performance of the evaluated LBL and VBL methods drops dramatically in such repetitive environments. As for the MFL approach, *PF-Mag* is the most robust technique among all the evaluated methods. However, *PF-Mag* is not accurate enough and its performance deteriorates with the increased size of the environments. Finally, it is obvious that *AMCL* performs the worst among all the evaluated approaches.

V. Conclusion

This paper investigated and briefly reviewed the problem of infrastructure-free mobile robot global localization with low-cost and efficient sensors in repetitive environments. Different state-of-the-art infrastructure-free localization approaches, including LiDAR-based localization (LBL), vision-based localization (VBL), and magnetic field-based localization (MFL) methods, were analyzed and evaluated. Extensive global localization experiments were conducted in real-world repetitive scenarios and the overall evaluations indicated that VBL methods perform slightly better than LBL and MFL approaches. However, VBL methods are brittle and suffer from illumination and viewpoint changes. LBL methods are robust to lighting variations and have good accuracy, but they take more iterations and longer distances to be localized. MFL approach has the most robust performance but its accuracy is not good enough. Hence, infrastructure-free global localization in repetitive environment is still a challenging problem which requires more research efforts. Future work will develop new algorithms to tightly couple the MFL with LBL method to enhance the localization performance in repetitive environments and to serve as a feasible alternative to the existing infrastructure-based solutions.

References

[1] K. Zinchenko *et al.*, "A study on speech recognition control for a surgical robot," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 607–615, 2016.

[2] X. Zang *et al.*, "Behavioral indoor navigation with natural language directions," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 283–284.

[3] P. Nazemzadeh *et al.*, "Indoor localization of mobile robots through QR code detection and dead reckoning data fusion," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 6, pp. 2588–2599, 2017.

[4] J. Tiemann *et al.*, "ATLAS FaST: Fast and simple scheduled TDOA for reliable ultra-wideband localization," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 2554–2560.

[5] D. Fox *et al.*, "Monte carlo localization: Efficient position estimation for mobile robots," *AAAI/IAAI*, no. 343-349, pp. 2–2, 1999.

[6] P. Jensfelt and S. Kristensen, "Active global localization for a mobile robot using multiple hypothesis tracking," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 5, pp. 748–760, 2001.

[7] D. Fox, "Adapting the sample size in particle filters through KLD-sampling," *The International Journal of Robotics Research*, vol. 22, no. 12, pp. 985–1003, 2003.

[8] S. Kohlbrecher *et al.*, "A flexible and scalable SLAM system with full 3D motion estimation," in *2011 IEEE International Symposium on Safety, Security, and Rescue Robotics*. IEEE, 2011, pp. 155–160.

[9] J. Röwekämper *et al.*, "On the position accuracy of mobile robot localization based on particle filters combined with scan matching," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2012, pp. 3158–3164.

[10] W. Hess *et al.*, "Real-time loop closure in 2D LIDAR SLAM," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 1271–1278.

[11] J. Li *et al.*, "Deep learning for 2D scan matching and loop closure," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 763–768.

[12] G. D. Tipaldi *et al.*, "Lifelong localization in changing environments," *The International Journal of Robotics Research*, vol. 32, no. 14, pp. 1662–1678, 2013.

[13] Y. Yue *et al.*, "Collaborative semantic perception and relative localization based on map matching," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, to be published, 2020.

[14] Y. Yue *et al.*, "Hierarchical probabilistic fusion framework for matching and merging of 3-D occupancy maps," *IEEE Sensors Journal*, vol. 18, no. 21, pp. 8933–8949, Nov 2018.

[15] J. Wolf *et al.*, "Robust vision-based localization by combining an image-retrieval system with monte carlo localization," *IEEE Transactions on Robotics*, vol. 21, no. 2, pp. 208–216, 2005.

[16] B. Williams *et al.*, "A comparison of loop closing techniques in monocular SLAM," *Robotics and Autonomous Systems*, vol. 57, no. 12, pp. 1188–1197, 2009.

[17] D. Gálvez-López and J. D. Tardos, "Bags of binary words for fast place recognition in image sequences," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1188–1197, 2012.

[18] S. Lowry *et al.*, "Visual place recognition: A survey," *IEEE Transactions on Robotics*, vol. 32, no. 1, pp. 1–19, 2015.

[19] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An open-source SLAM system for monocular, stereo, and rgb-d cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

[20] X. Tang *et al.*, "Place recognition using line-junction-lines in urban environments," in *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*. IEEE, 2019, pp. 530–535.

[21] Y. Yue *et al.*, "Collaborative semantic understanding and mapping framework for autonomous systems," *IEEE/ASME Transactions on Mechatronics*, to be published, 2020.

[22] J. Haverinen and A. Kemppainen, "Global indoor self-localization based on the ambient magnetic field," *Robotics and Autonomous Systems*, vol. 57, no. 10, pp. 1028–1035, 2009.

[23] J. Moore and R. Tedrake, "Magnetic localization for perching UAVs on powerlines," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2011, pp. 2700–2707.

[24] B. Li *et al.*, "How feasible is the use of magnetic field alone for indoor positioning?" in *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*. IEEE, 2012, pp. 1–9.

[25] M. Frassl *et al.*, "Magnetic maps of indoor environments for precise localization of legged and non-legged locomotion," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2013, pp. 913–920.

[26] Z. Wu *et al.*, "Magnetic-assisted initialization for infrastructure-free mobile robot localization," in *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*. IEEE, 2019, pp. 518–523.

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

[28] ——, "Day and night collaborative dynamic mapping in unstructured environment based on multimodal sensors," in *2020 International Conference on Robotics and Automation (ICRA)*. IEEE, to be published, 2020.

[29] S. Thrun *et al.*, *Probabilistic robotics*. The MIT press, 2005.

[30] E. Rublee *et al.*, "ORB: An efficient alternative to SIFT or SURF," in *2011 International Conference on Computer Vision (ICCV)*. IEEE, 2011, pp. 2564–2571.

[31] H. Badino *et al.*, "Real-time topometric localization," in *2012 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2012, pp. 1635–1642.