Artificial Beacons with RGB-D Environment Mapping for Indoor Mobile Robot Localization

Chun-Chi Lai and Kuo-Lan Su[*]

Department of Electrical Engineering, National Yunlin University of Science and Technology,
123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan

The most important issue for intelligent mobile robot development is the ability to navigate autonomously in the environment to complete certain tasks. Thus, the indoor localization problem of a mobile robot has become a key component for real applications. In general, two categories of mobile robot localization technique are identified: one is robot pose tracking, and the other is robot global localization. In pose tracking problems such as the simultaneous localization and mapping (SLAM) process, the robot has to find its new pose using the observed landmarks or features in its knowledge base of its previous observations. In global localization methodology, a robot does not have knowledge of its previous pose. It has to find its new pose directly in the environment, such as by using a global positioning system (GPS). On the other hand, an artificial beacon-based localization technique, such as the received signal strength indicator (RSSI), causes higher pose uncertainty. However, the artificial beacon can provide a good initial reference for robot position mapping. The gyrocompass of a mobile robot is suited for short-term dead-reckoning. Also the RGB-D camera of a mobile robot can record meaningful features or landmarks in 3D space. The purpose of this work is to fuse the advantages of these sensors via strategy control by a particle filter to enhance the estimation accuracy of indoor mobile robot localization.

1. Introduction

The issue of intelligent mobile robot self-localization can be divided into two categories: robot pose tracking and robot global localization. In the pose tracking problem, such as the simultaneous localization and mapping (SLAM) process, the robot has to find its new pose using the knowledge of its previous pose i.e., to associate the observed landmarks in its knowledge base. In the global localization problem, a robot does not have knowledge of its previous pose. It has to find its new pose directly in the environment, such as by using a global positioning system (GPS). Furthermore, for real mobile robot applications, the kidnapping problem arises, which is a combination of those two localization problems. For example, while the mobile robot is performing an extended kalman filter (EKF)-SLAM, it may be suddenly kidnapped by moving to another pose, after which the robot cannot associate the landmarks in its knowledge base. Then the mobile robot will fail and be dangerous.

[*]Corresponding author: e-mail: sukl@yuntech.edu.tw

In this work, a fused concept including RGB-D visual landmark extraction, gyro-odometer, and artificial beacons for mobile robot indoor localization is proposed as shown in Fig. 1. First, a SLAM process with RGB-D camera is executed to construct an environment map with visual features. Ideally, the RGB-D camera captures all significant features in the indoor environment, but in practice, the visual feature extraction is constrained by the light, shapes, and colors of the environment. Hence, in real applications, we added some artificial beacons, such as an infrared receiver (IR) or radio beacons to the environment, and marked them on the map to help indoor mobile robot localization. Finally, the experimental results show that the proposed concept actually increased the accuracy for robot pose association and localization.

2. Related Work

2.1 Localization and mapping development of a mobile robot

An intelligent mobile robot is an artificial system that perceives the environment and its own status through sensors in order to navigate in an indefinite environment to complete certain tasks. Thus, the SLAM problem, also known as concurrent mapping and localization (CML), is one of the fundamental challenges of intelligent mobile robot development. The SLAM problem deals with the uncertainty in the pose of a robot when the environment is only partially known or completely unknown, and uses sensor measurements to estimate the robot pose and to construct an incremental environment map simultaneously. Many studies and technologies have been developed in this field.[1] Up to this point, two aspects of SLAM techniques have been classified: feature-based SLAM[2] and graph-based SLAM.[3] Feature-based SLAM applies estimation methods from Bayesian probability, and graph-based SLAM uses global optimal estimation techniques for alignment based on relative observations.
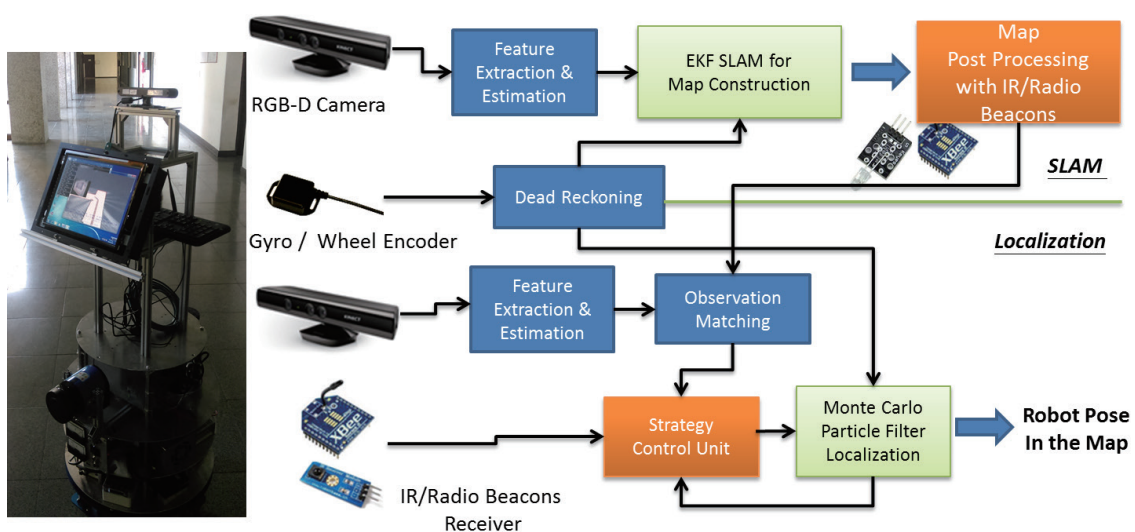

Fig. 1. (Color online) The mobile robot platform and proposed system architecture.

2.2 RGB-D environment SLAM

Visual recognition is a critical component for a robot to behave autonomously. Since the RGB-D sensor of Microsoft Kinect,[4] as shown in Fig. 2(a), was announced for the Xbox-360 console in 2009, it has been possible to easily capture image and depth maps at a lower cost, and therefore to use Kinect for various visual recognition studies including SLAM, and recognition of objects, human gestures, and actions. Figure 2(b) shows results from the RGB-D SLAM.

3. SLAM with RGB-D Camera

3.1 Visual landmark and feature extraction

Numerous techniques for computer vision have been proposed to model landmarks for a robot navigating in indoor environments. They all rely on two assumptions: (1) landmarks have to be easily detected in the image signal; and (2) landmarks must be locally characterized to distinguish them from others features. Most recent work also makes use of points to define landmarks, taking advantage of new, powerful interest point detection and characterization algorithms such as the scale invariant feature transform (SIFT),[5] shown in Fig. 3, which has emerged as an effective methodology in general object recognition as well as for other applications relating to machine vision. An important aspect of this approach is that it generates large numbers of features in a local region such as location, scale, rotation, magnitude, and orientation in order to record information about key points.

(a)   (b)

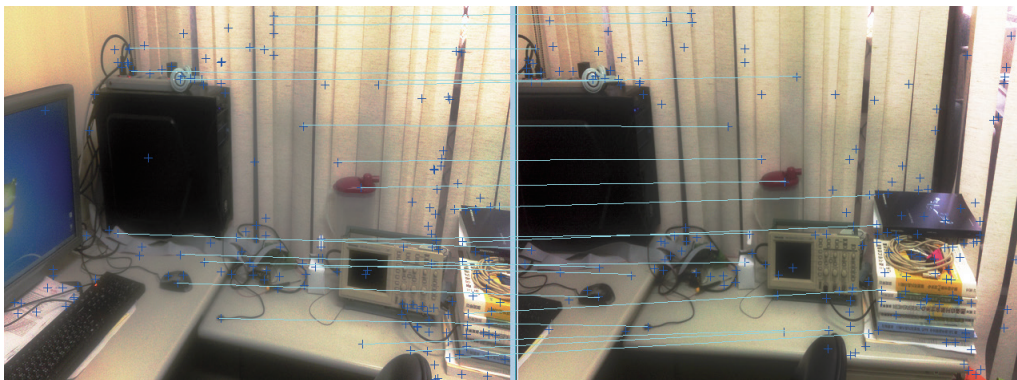Fig. 2.   (Color online) (a) Microsoft Kinect RGB-D camera and (b) the results of SLAM.

Fig. 3.   (Color online) The SIFT feature matching between two different views.

3.2 Bayesian filtering SLAM

SLAM using Bayesian filtering involves finding appropriate representations for modeling both an observation and motion, as shown in Fig. 4.

The observation model describes the probability of making an observation in time $k(z_k)$, when a vehicle's location x_k and a map (including the landmark locations) are known. It is generally described as

$$p(z_k|x_k, m). \tag{1}$$

The motion model describes the probability distribution of transitions in the robot's state as

$$p(x_k|x_{k-1}, u_k). \tag{2}$$

The state transition is assumed to be a Markov process in which the next pose x_k depends only on the previous state x_{k-1} and the applied control input u_k. The SLAM algorithm is implemented in a standard two-step recursive (sequential) prediction and correction as follows.

$$\text{Prediction: } p(x_k, m|z_{0:k-1}, u_{0:k}, x_0) = \int p(x_k|x_{k-1}, u_k) \times p(x_{k-1}, m|z_{0:k-1}, u_{0:k-1}, x_0)dx_{k-1} \tag{3}$$

$$\text{Correction: } p(x_k, m|z_{0:k}, u_{0:k}, x_0) = \frac{p(z_k, x_k, m)p(x_k, m|z_{0:k-1}, u_{0:k}, x_0)}{p(z_k|z_{0:k-1}, u_{0:k})} \tag{4}$$

The solution involves an appropriate representation for both the motion model and the observation model that allows efficient and consistent computation of the prior and posterior distributions. The most popular of the state-of-the-art SLAM methods is the application of the EKF. The EKF linearizes the nonlinear motion model at an estimated linearization point, uses a first-order approximation to represent the state, and involves a Jacobian matrix calculation. The SIFT RGB-D feature map can be constructed directly with the EKF-SLAM, as shown in Fig. 5.
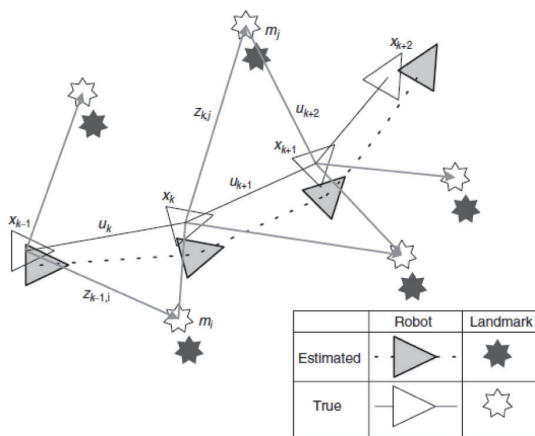

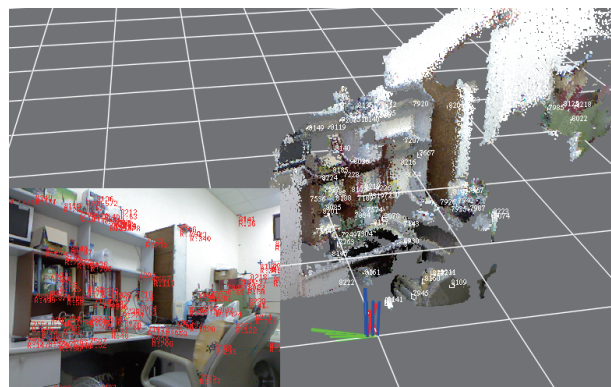Fig. 4.   SLAM process with observation and motion.

Fig. 5.   (Color online) RGB-D map with SIFT features.

4.   System Implementation

4.1   EKF construction of an indoor environment and beacons post processing

There are a number of SLAM application programming interfaces (APIs), such as the mobile robot programming toolkit (MRPT), for real mobile robot SLAM implementation. In this work, the EKF-SLAM toolbox using the classical EKF implementation was applied in advance to simulate a Kinect RGB-D camera to construct an environmental feature map.[6] Figure 6(a) shows the EKF-SLAM process with point landmarks in a global view, and Fig. 6(b) shows the point landmarks captured in the image of an RGB sensor on a mobile robot.

Ideally, the RGB-D camera is expected to capture all significant features of the indoor environment, but practically, the visual feature extraction is constrained by light, shapes, and colors. Hence, in real applications, it is expected that some artificial beacons, such as IR or radio beacons, will be added in the environment and marked on the map to help indoor localization. The idea and concept are shown in Fig. 7, where the diamond shapes on the map are mapped to the visual features in the environment, and the blue circles are man-made beacons which are placed in corresponding positions in the environment.

4.2   Monte Carlo particle filter (PF)

A PF, also called CONDENSATION (conditional density propagation),[7] is based on Monte Carlo and Bayesian methods. The PF uses random sampling. Each particle presents an assumption of the location (x,y) and orientation (θ) of the robot. For example, 1000 green particles are initialized with uniform distribution as shown in Fig. 8. The z-axis represents the robot head angle in radians.

The advantage of PF is that it can eliminate background noise. For mobile robot applications, two types of data are distinguished: perceptual data, such as landmark measurements, and odometer data or control, which carries information about robot motion. The PF algorithm for mobile robot self-localization is shown in Table 1, in which x_t is robot pose at time t, u_t is the robot motion command, and z_t is measurement of the robot.

(a)   (b)
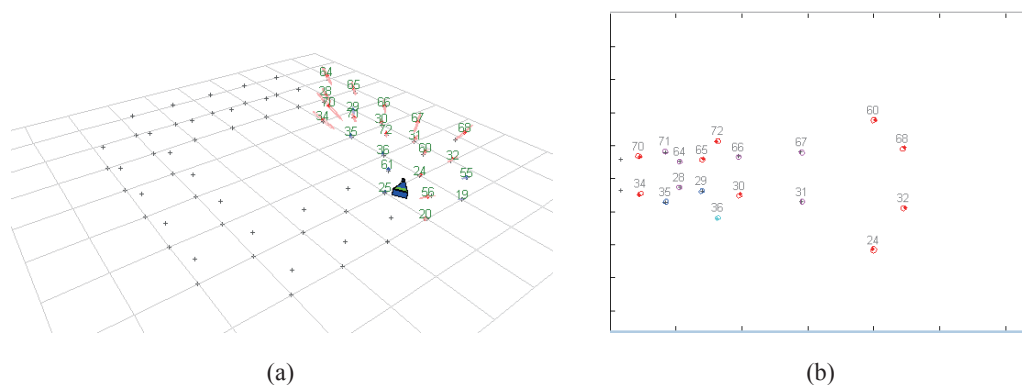
Fig. 6.   (Color online) EKF-SLAM toolbox for map construction: (a) global mapping in EKF-SLAM and (b) feature landmark in camera view.
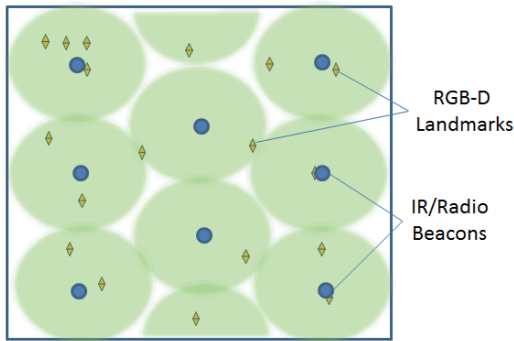
Fig. 7. (Color online) Visual features/landmarks and IR or radio beacons on the map.

Table 1
PF algorithm.

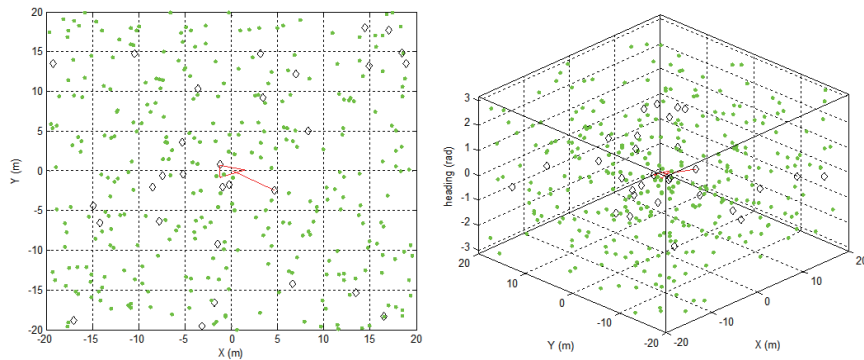| PF algorithm for mobile robot localitzation |
| --- |
| Input: X_{t-1}, u_t, z_t, m |
| Output: X_t |
| $\overline{X}_t = X_t = \phi$ |
| for $m = 1$ to M do |
| 1. $x_t^{[m]} = $ sample_motion_model(u_t, $x_{t-1}^{[m]}$) |
| 2. $w_t^{[m]} = $ measurement_model(z_t, $x_t^{[m]}$, m) |
| 3. $X_t = X_t + \langle x_t^{[m]}, w_t^{[m]} \rangle$ |
| endfor |
| for $m = 1$ to M do |
| 1. draw i with probability $\propto w_t^{[i]}$ |
| 2. add $x_t^{[i]}$ to X_t |
| endfor |
| return X_t |

Fig. 8.   (Color online) PF initialization with uniform distribution.

4.3   Strategy control with beacons

The strategy is to control the field of sampling when PF is initialized for robot pose estimation. The main idea is that the wireless beacon's radio signal strength or the infrared beacon's transmit scope can be pre-measured or pre-determined as a circular field. Figure 9(a) shows the green particles with uniform sampling of the entire map, because it has no cues to guess the initial pose of the robot in the map for PF initialization. However, if there is a beacon mounted on the ceiling and upon the robot as shown in Fig. 9(b), then we can constrain all the green particles with uniform distribution inside the circular boundary, and speed up PF initialization.

Furthermore, for a consistent estimate of PF, we modified the concept from Fig. 9(b). i.e., when the statistical standard deviation of all particles for estimation of the robot pose exceeds a threshold, *e.g.*, half the radius of the beacon's scope, as shown in Fig. 10(a), and if the nearest beacon is identified as shown in Fig. 10(b), then the particle is re-sampling in the scope and its new mean position is the same as the last mean position. The complete flow diagram for re-sampling control is shown in Fig. 10(c).
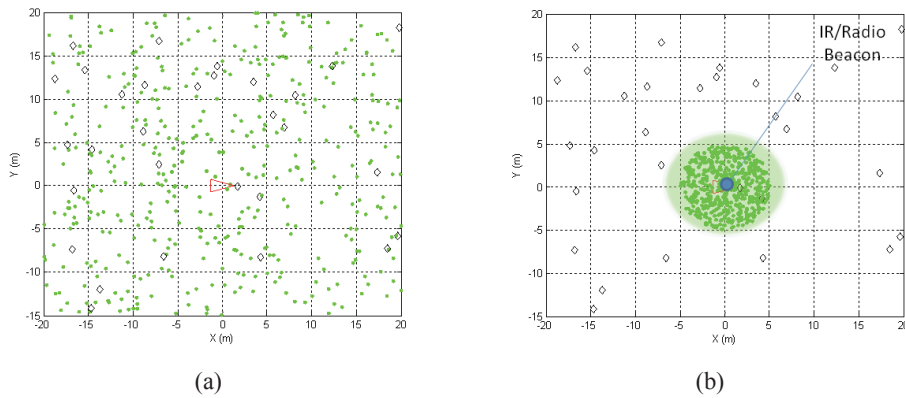
(a)   (b)

Fig. 9.   (Color online) PF initialization with/without beacon: (a) uniform sampling of the entire map and (b) uniform sampling of the beacon's field.
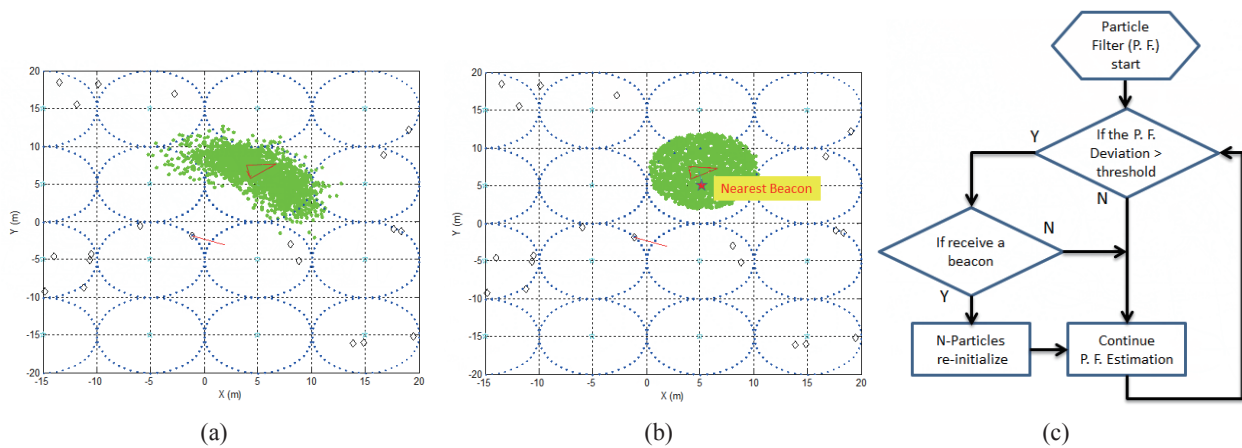

(a)   (b)   (c)

Fig. 10. (Color online) Consistent particles re-sampling control: (a) the estimates of deviation increase, (b) re-sampling with last mean position, and (c) flow diagram of particle re-sampling.

5.   Experimental Results

5.1   Comparison of number of landmarks for PF localization

The first experiment indicated that the number of landmarks/features and number of particles affected the accuracy of the estimated pose. There are 100 landmarks and 2000 particles on the map for mobile robot self-localization, as shown in Fig. 11(a). The red line represents the PF's estimate and the blue line represents actual trajectory of the robot for a 2500 time-step (250 s) simulation. Figure 11(b) shows the standard deviation of robot pose estimation (blue, x; green, y; red, θ), and the PF achieved excellent pose estimation over the time step. Figure 11(c) is simulated with 100 landmarks and 1000 particles, where the initial estimation could not locate the actual robot pose, but after 80 time steps, the particles converge to robot's actual trajectory. Figure 11(d) is simulated with 30 landmarks and 1000 particles, where the estimate of the mobile robot's trajectory becomes difficult.
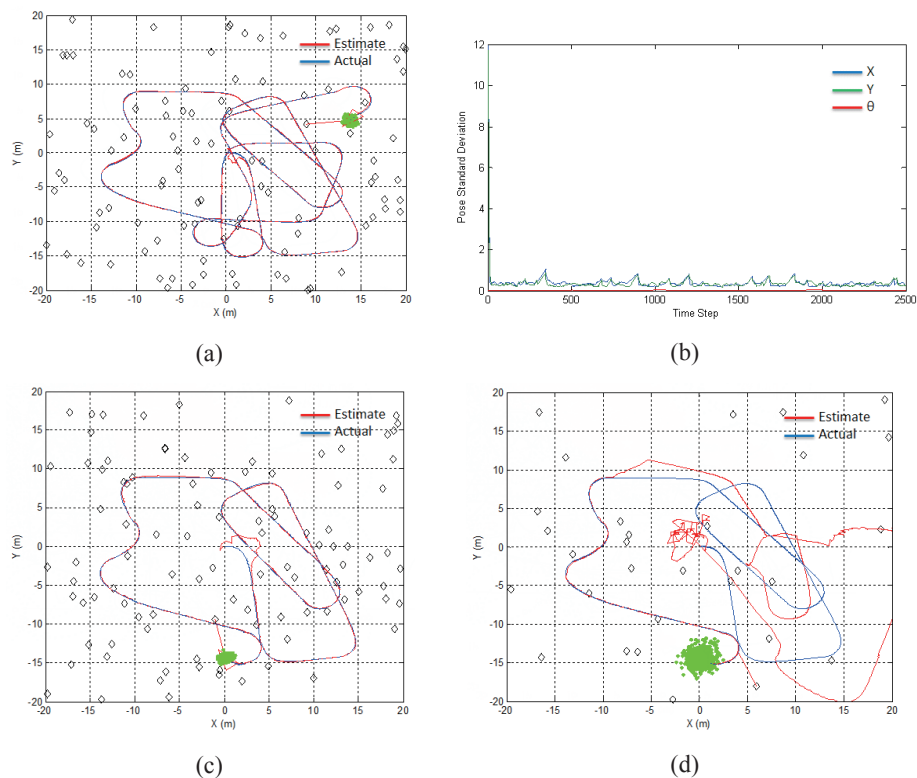
Fig. 11. (Color online) The comparison of estimates for different landmarks and particles; the RGB-D sensor observation ability is limited to a 6 m range and −45 to 45 deg with a deviation of 0.008 m and 1 deg. The odometer deviation is 0.005 m and 2 deg. (a) landmarks = 100, particles = 2000; (b) the pose standard deviation in each time step; (c) landmarks = 100, particles = 1000; and (d) landmarks = 30, particles = 1000.

5.2   Estimation with initial beacon

From the experiment in § 5.1, the number of landmarks is the key to increasing the accuracy of the PF estimation. In real applications however, the visual landmark is constrained by the light, shapes, and colors in the environment. Thus, the artificial beacons were added to assistant robot self-localization. The concept and comparison are shown in Figs. 9 and 12. Without artificial beacons as shown in Fig. 9(a), the initial estimation of robot position is a uniform distribution over the map, and it is difficult to converge to an actual pose in the beginning phase with fewer visual landmarks observed. When more landmarks were observed in the middle time step, the PF converged to an actual robot trajectory as shown in Fig. 12(a). As shown in Fig. 9(b), an artificial beacon was added, and it restricted the initial particle sampling field. Figure 12(b) shows an acceptable estimation result for matching an actual trajectory.

5.3   Multiple beacons with strategy control

From the experiment in § 5.2, an initial beacon restricted the initial particle sampling field to reach an acceptable estimation for matching an actual trajectory. In this experiment, multiple beacons were simulated with the strategy control to assist mobile robot localization in an environment with fewer less visual landmarks (landmarks = 20). Figure 13 shows the comparisons
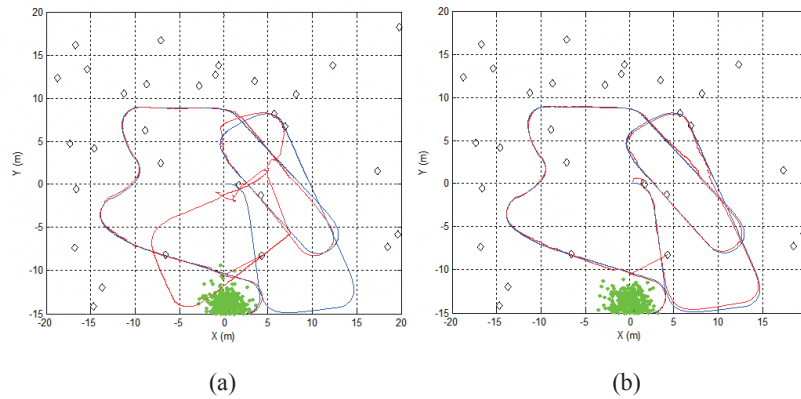
(a)

(b)

Fig. 12. (Color online) PF comparisons with/without an initial beacon: (a) initial estimations are wrong and (b) an acceptable estimation with an initial beacon.
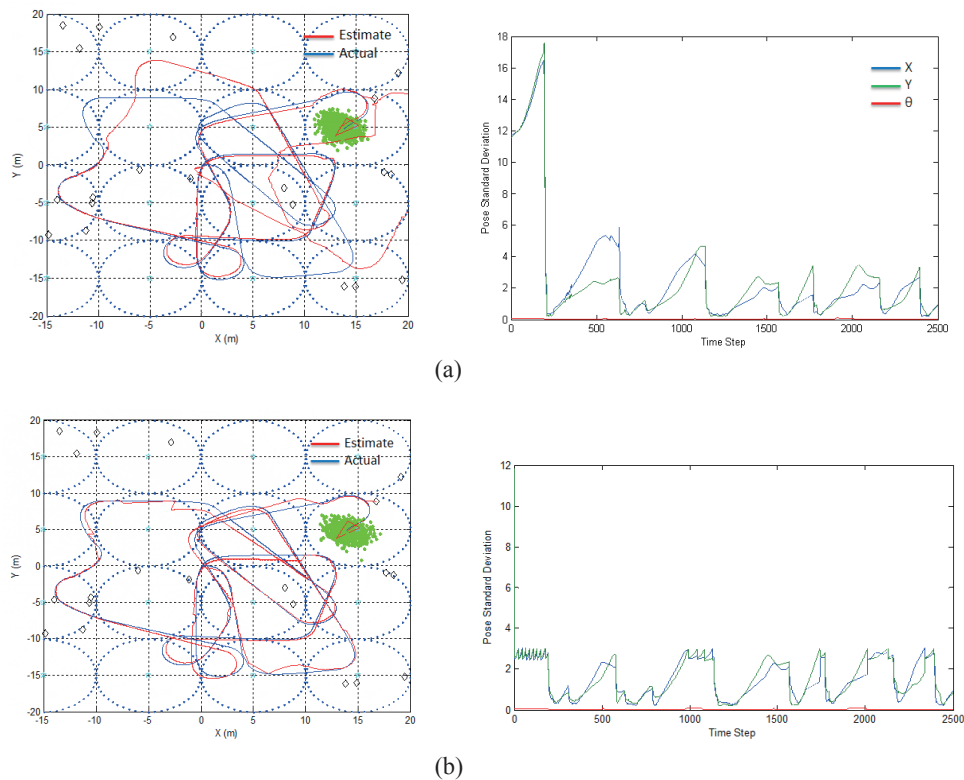

(a)

(b)

Fig. 13. (Color online) PF comparisons with/without multiple beacons: (a) landmarks = 20, particles = 800, without beacon control, the position estimation mean error = 3.5510 m; the standard deviation of the estimation increased when landmarks were not observed, and (b) landmarks = 20, particles = 800, with multiple beacon control, the position estimation mean error = 0.4248 m; the standard deviation of the estimation was maintained within the threshold when the beacon signal was received.

when landmarks = 20, particles = 800. Without the artificial beacons to assistant mobile robot localization in Fig. 13(a), the mean error in position was 3.5510 m. With the artificial beacons to assistant mobile robot localization in Fig. 13(b), the mean error in position was 0.4248 m, which is

clearly a better result than that without beacons. Figure 14 shows a repeat of the experiment with particles = 2000. This experiment indicated that more particles enhance the estimation accuracy. In Fig. 14(a) without beacon control, the mean position error was decreased to 1.0791 m; in Fig. 14(b) with beacon control, the mean position was decreased to 0.2589 m. All the comparisons are shown in Table 2.
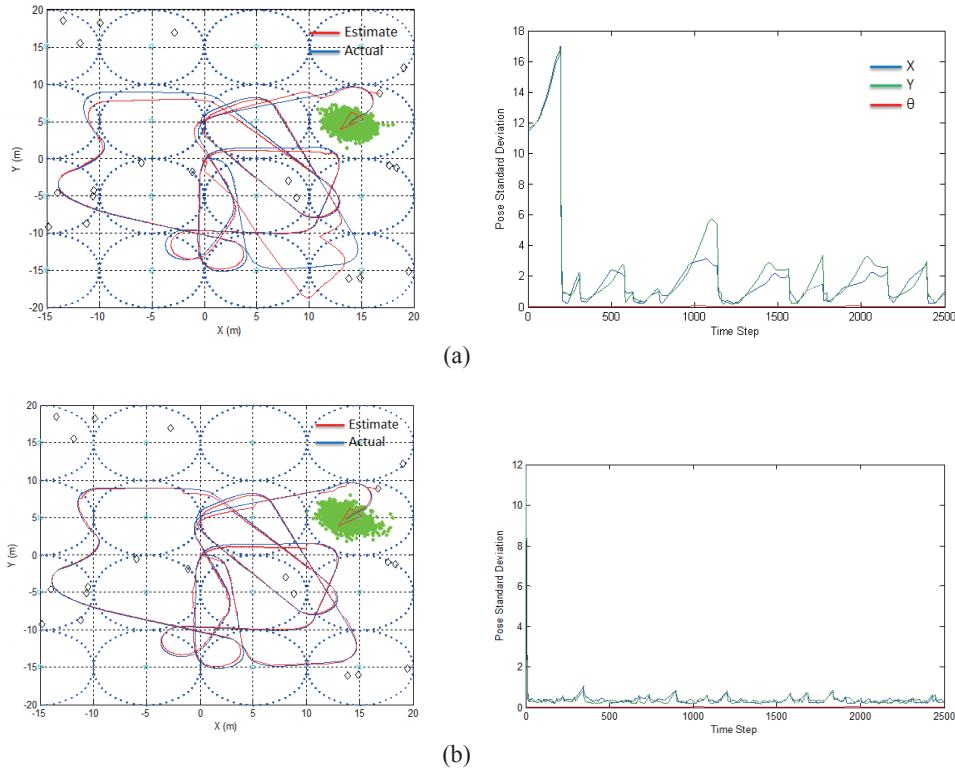

(a)

(b)

Fig. 14. (Color online) PF comparisons with/without multiple beacons: (a) landmarks = 20, particles = 2000, without beacon control, the position estimation mean error = 1.0791 m; the standard deviation of the estimation increased when landmarks were not observed, and (b) landmarks = 20, particles = 2000, with multiple beacon control, the position estimation mean error = 0.2589 m; the standard deviation of the estimation was maintained within the threshold when beacon signal was received.

Table 2
Position mean error comparison with/without beacons in 20 landmarks.

With beacons	No. of particles	
	800	2000
No	Position mean error 3.5510 m	Position mean error 1.0791 m
Yes	Position mean error 0.4248 m	Position mean error 0.2589 m

6.   Conclusions

In this work, a fused concept of RGB-D visual landmarks, gyro-odometer, and artificial beacons was proposed to increase the accuracy of mobile robot indoor localization. First, an EKF-SLAM process using RGB-D camera was carried out to construct an environment map with visual features. Ideally, the RGB-D camera captured all significant features in the indoor environment, but practically, the visual feature extraction was constrained by the light, shapes, and colors of the environment. Hence, in real applications, we added some artificial beacons, such as IR or radio beacons, in the environment and marked them on a map to help indoor mobile robot self-localization. In addition, for consistent estimation of robot trajectory, a re-sampling control strategy using a PF was proposed. Finally, all the experimental results show the proposed method actually increases the accuracy of mobile robot association and localization when fewer visual features are in the environment.

References

1   T. Bailey and H. Durrant-Whyte: IEEE Rob. Autom. Mag. **13** (2006) 108.
2   M. R. Walter, R. M. Eustice, and J. J. Leonard: Int. J. Rob. Res. **26** (2007) 335.
3   G. Grisetti, R. Kummerle, C. Stachniss, and W. Burgard: Intell. Transp. Syst. Mag. **2** (2010) 31.
4   Z. Zhang: Microsoft Kinect Sens. Eff. **19** (2012) 4.
5   P. Scovanner, S. Ali, and M. Shah: Proc. 15th Int. Conf. Multimedia (2007) p. 357.
6   J. Sola: IEEE Int. Conf. Rob. Autom. (2010) p. 3513.
7   C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau: IEEE Int. Conf. Med. Biol. (2006) p. 6384.