Proceedings of the 2014 IEEE
International Conference on Robotics and Biomimetics
December 5-10, 2014, Bali, Indonesia

A Multi-sensor-based Mobile Robot Localization Framework

Peng Duan, Guohui Tian, and Hao Wu

***Abstract*— This paper presents a novel multi-sensor-based robot localization framework inspired by human coarse-to-fine recognition mechanism to realize fast and robust localization in the process of robot navigation. This localization framework consists of two parts: coarse place recognition and accurate location estimation. The coarse place recognition is realized using an onboard camera, whereas an image retrieval system is employed. The coarse localization system utilizes feature matching between the observed image and the re-ranked retrieval images to infer the possible locations of the robot. To obtain the accurate pose of robot, a modified particle filter which is mainly based on the laser radar data is implemented. To integrate the two localization stages, the current states of the particles are monitored. Once the information entropy changed greatly or the number of the effective particles is in a low level, a set of pointed particles is generated based on these estimated locations and then injected into the modified particle filter timely to ensure that enough particles are surviving around the correct pose of robot. Experiments are conducted extensively in an office environment and the results exhibits great improvement on speed and stability of mobile robot localization compared to conventional localization methods.**

I. INTRODUCTION

Mobile robot localization, which allows a mobile robot to achieve its position, is one of the fundamental problems in the field of mobile robot research [1]. Indeed, robust self-localization is the basis of autonomous navigation and task execution in complex environments. Although steady progress in mobile robot localization has been made in the past decades, it is still an open problem to achieve the efficient and robust localization [2].

The main difference among these localization methods is the pattern used to represent the position and the type of information employed for localization. Compared to the proximity sensors, cameras have lots of desirable properties. They are low-cost and passive, but can provide a wealth of information. In addition, the image information is more suitable for robot operation particularly in the populated environments rather than light or sound-based distance information. What's more, coarse place recognition can be achieved using a limited number of images, whereas image-retrieve makes place recognition more attractive to mobile robot localization because it will accelerate the speed of localization virtually [3]. However, due to lack of distance information, it is hardly to realize accurate robot location using camera alone.

Similarly, many robots carry time-of-flight ultrasonic [4] or laser-based range sensors [5] for dead reckoning or self-localization. Although these systems have accessed to applications, they are prone to restricted by environment conditions. Frequent corrupted observations, as a result of the existence of the dynamic targets, will lead to unrecoverable localization failure. Therefore, in view of its limitation using a single sensor, more and more researchers turn to multi-sensor information cooperation or data fusion based localization to realize reliable robot navigation. For example, Biswas [6] presented a sensor selection-based method to conduct mobile robot localization in diverse environments. By analyzing the strengths of localization, the robot selected the most effective sensors to collect environment information. In addition, Lee [7] contributed at proposing a robust pose tracking approach using hierarchical multiple sonar sensors in highly non-static environments. The proposed localization algorithm employs a non-corrupted window to sample the uncorrupted sensor data. Moreover, Karthick [8] presents a rule based algorithm by combining data from ultrasonic sensor, encoder and gyroscope to realize mobile robot localization. In brief, for most applications, the existed methods are focused on multi-sensor data fusion based on some filtering algorithms, which lack of an effective mechanism to monitor whether the collected information from sensors is corrupted or not. Hence, localization based these methods is not always robust, it is difficult to develop the advantages of each sensor fully.

In this paper, we propose an effective multi-sensor-based framework that integrates image information with laser data to realize fast and robust localization in the process of robot navigation. Just like a person walks into a new environment, to complete some tasks, he may estimate his location based on the scene in the front of him firstly. He usually then goes through a self-localization process from coarse location to fine position iteratively. The proposed framework is just to provide a closed-loop localization mechanism which consists of two stages: coarse place recognition and accurate location estimation. The coarse place recognition using image information will accelerate the speed of robot localization. In the accurate localization stage, a set of pointed particles is generated based on the locations estimate from place recognition and then injected into the motion update process of the particle filter timely to ensure that enough particles are surviving around the correct pose of the robot.

Coarse-to-fine mechanism, as an excellent strategy, has been applied many systems such as vision based pedestrian localization [9], object detection [10], and scene classifi-

This work is supported by the National Natural Science Foundation of China (61075092, 61203330), by the Natural Science Foundation of Shandong Province (ZR2011FM011), and by the Independent Innovation Foundation of Shandong University-IIFSDU (2011JC017).

Guohui Tian is with the School of Control Science and Engineering, Shandong University, Jinan 250061, China, corresponding author: 86-0531-88392032; e-mail: g.h.tian@sdu.edu.cn

Peng Duan and Hao Wu are with the School of Control Science and Engineering, Shandong University, Jinan 250061, China.

978-1-4799-7397-2/14/$31.00 © 2014 IEEE

cation [11], et al. What's more, Park [12] contributed at presenting a coarse-to-fine global localization method for mobile robots navigation based on hybrid maps of objects and spatial layouts. Different from the methods mentioned above, the localization approach presented in this paper is not just a coarse-to-fine process alone. Its advantages lies on the robust and feedback characteristics. On one hand, for example, for most of the existing coarse-to-fine systems, if the coarse recognition is failed, the fine localization will not be achieved generally. However, our localization is robust due to a particle based localization representation is employed in the accurate location estimation stage. The coarse place recognition stage is just to provide a guidance of particles injection. If the place recognition is failed, it does not mean that all the particles are missing the correct pose. In contrast, our place recognition mechanism will be activated rapidly and continuously when the major particles mistakes the pose of the robot. On the other hand, the localization is not open-loop localization process. The current states of the particles are monitored to determine the time of making place recognition in the robot navigation process. In other words, the coarse place recognition stage is only conducted when necessary. The motivation of this paper is to present a closed-loop multi-sensor-based framework for mobile robot localization. Unlike the open-loop coarse-to-fine localization systems [7], [9], [10], [12], we provide a localization state monitoring mechanism to fuse multi-sensor measurement. Experiments are conducted in an office environment, and the results exhibits great improvement on speed and stability of robot localization.

The remainder of this paper is organized as follows. After discussing the overview of the proposed method in Section 2, we describe the details about the coarse place recognition scheme in Section 3. Next, the accurate location estimation based on a modified particle filter is introduced in Section 4. In Section 5, experiments are presented to verify the robustness and reliability of the proposed method. Finally, we conclude this paper in Section 6.

II. OVERVIEW OF THE MULTI-SENSOR-BASED ROBOT LOCALIZATION FRAMEWORK

Fig. 1 gives an overview of the proposed closed-loop multi-sensor-based localization framework. Firstly, a camera on board is employed to capture pictures in the working environment. The grabbed picture is then fed to an image retrieval system to infer the possible locations of the robot. In the accurate location estimation stage, a particle-based robot position representation which is mainly based on the laser radar data is implemented. To combine the two localization stages, the current states of the particles are monitored. Once the effective number of particles below a specified level, the localization system orders the onboard camera to capture images, and activates the coarse place recognition mechanism immediately. A set of pointed particles are generated based on the locations estimated from place recognition and then injected into the motion update process of particle filter timely. Such an injection of pointed particles will not only

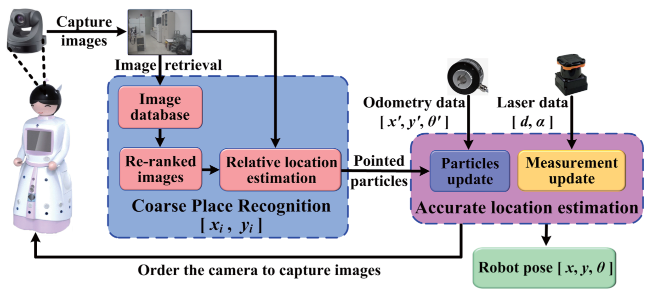

Fig. 1. Overview of the proposed closed-loop localization framework.

increase the speed of the robot localization globally, but also improve the robustness in position tracking.

III. SCHEME OF COARSE PLACE RECOGNITION

The coarse place recognition discussed in this paper is mainly concerned as a coarse location estimation problem, where the map of environment is known in advance. Most of the images in the image database are captured purposefully ahead of time and each of them has been bounded with a coarse location and a non-precise direction where it was grabbed.

The architecture of the presented coarse place recognition scheme is shown in Fig. 2. The scheme consists of two parts: offline process and online process. In the offline procedure, representative images are collected in advance. Invariant features on images are then extracted based on Speeded Up Robust Features (SURF) detectors and descriptors [13]. Using these invariant features, a visual vocabulary is then created based on the *k*-mean clustering and *k-d* tree structure. Each image in the database is afterwards expressed using the terms of the visual vocabulary based on an Image Model Vector (IMV). In the online process, invariant features on the image captured by the mobile robot are extracted. The closest visual words are then assigned to build the image model vector. Similar images are retrieved by comparing the model vector to those in the image database. To facilitate accurate retrieval, geometric constraints on the detected features are taken into account. Finally, the coarse location of the robot

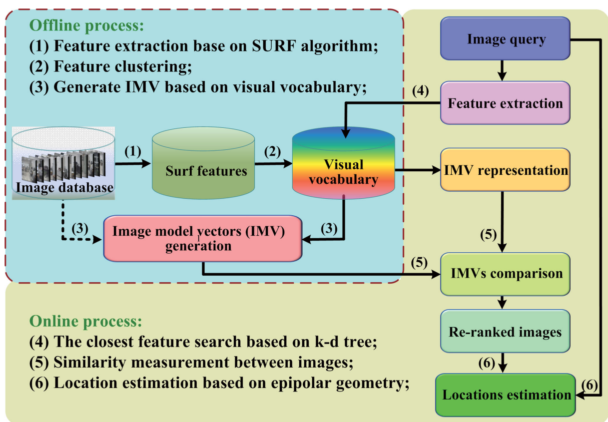

Fig. 2. The architecture of the proposed place recognition scheme.

is estimated based on the epipolar geometry.

A. SURF Detectors and Descriptors

Due to it is hardly to have an image from the camera that fully matches an image in the database, some invariant features are desirable to find the one with most similar contents. It is expected that the features should be invariant to translation, rotation, and scaling. To extract these invariant features on images, SURF detectors and descriptors are employed in this paper. Hence, each image in the database can be expressed by (1).

$$I_j = (x_j^c, y_j^c, \alpha_j^c, inv\_surfvector_j^i | i = 1, 2, 3 \cdots n) \quad (1)$$

where the three parameters ahead describe the location where the image j was grabbed. What's more, for each invariant SURF vector, it includes: the interest point coordinates (u, v), the dominant orientation ω of the interest point, and a 64 dimensional descriptor.

$$inv\_surfvector_j^i = (u, v, \omega, surf_{64})_j^i \quad (2)$$

B. Visual Vocabulary Building

Just like a text is made of a set of words, an image can also be expressed by a set of visual words. More specially, these visual words are invariant features. In general, the number of invariant features is significantly large. Using all features to create vocabulary is redundancy. In order to create an appropriate visual vocabulary, the well known *k*-means clustering algorithm is employed to group similar features into clusters. The centroid of each cluster is regarded as a visual word of the vocabulary. We should clear that in order to make the visual vocabulary able to perform well, the image database has to be diverse and covers all scene of the working environment.

C. Visual Vocabulary Building

For most features detected on the image, the most similar visual words are assigned. Each image can be represented using the visual words. To describe the image using the visual words, the appearance frequencies of each visual words of the vocabulary on the image are counted using a histogram. In addition, inspired by retrieval techniques applied in the task of text search [14], the inverse document frequency is also taken into account. The image can be represented using an image model vector [15]:

$$I_j^{mv} = (n_1 \cdot f_1^{id}, n_2 \cdot f_2^{id}, \cdots n_i \cdot f_i^{id} \cdots n_{N_C} \cdot f_{N_C}^{id})_j \quad (3)$$

where n_i denote the appearance frequency of the visual word i on the image I_j, and N_C is the size of the visual vocabulary. The inverse document frequency is calculated by (4).

$$f_i^{id} = \log \frac{N_I}{N_i} \quad (4)$$

where N_I is the total number of the images in the database and N_i is the number of the images that contains the visual word i.

D. Similarity Measurement

To compute the similarity between two images, two similarity measures are employed. One is the inner product of the model vector. The similarity score between two images can be computed as follows [15].

$$S_{score}(I_q, I_{db}) = \sum_{i=0}^{N_C} (I_q^{mv}(i) \cdot I_{db}^{mv}(i)) \quad (5)$$

where I_q denotes the query image, and I_{db} represents an image in the image database. N_C is the size of the visual vocabulary. Another similarity measure is based on the appearance frequencies of each visual words of the vocabulary [15].

$$S_{af}(I_q, I_{db}) = \sum_{i=0}^{N_C} \min(I_q^{mv}(i), I_{db}^{mv}(i)) \quad (6)$$

The image retrieval system ranks the retrieved image based on the similarity value, and the top three is taken as the input of the next stage.

E. Geometric Consistency Check

In the image retrieval process aforementioned, images are compared based on visual words. Due to lack of the geometry spatial layout information, the retrieval system may fail to produce the accurate results. Two images may be assigned to the same visual words, but in a totally different spatial layout. Hence, the geometric consistency check is a re-ranking stage to improve the accuracy of the image retrieval system.

A typical implementation is to use the RANSAC estimator [16]. This estimator works by selecting a small set of samples and estimating the model parameters from the samples. The computation of this process is time-consuming and expensive. In this paper, we pursue a simpler method which implements based on statistics. Some geometrical transformation parameters of each match are the computed and the distribution of them are then collected. As we know, the dominant orientation of the interest point, which is extracted in the process of feature detection, is a readily available as the difference between the principal direction of the query and target SURF features. Hence, the statistical distribution of rotation for each match is computed. The distribution of the observed angles is counted using a histogram. The corresponding matches with the highest peak on the histogram are considered as inliers. The matches that deviate too much from the mode of that distribution are eliminated.

F. Location Estimation

As discussed above, the image retrieval stage products the top three candidates. To determine the relationship of the viewpoint between the matched images, the epipolar geometry is employ. The epipolar geometry is the intrinsic projective geometry between two views [17]. It is independent of scene structure, and only depends on the internal parameters of the cameras and relative pose. The relative pose with respect to the reference view can be computed by decomposing the essential matrix based on SVD [17].

Based on the epipolar geometry, the location estimation stage will output three locations which correspond to the three input images, respectively. In the stage of the geometric consistency check, retrieval images are re-ranked based on feature geometric consistency, which means the higher rank the more similar to the query image. So the importance of the three locations is different. We define the location importance proportion is 6:3:1 according to the rank of images. No matter whether they indicate the same location, a set of pointed particles is generated around these locations and then injected into the next accurate localization stage to increase the diversity of the samples.

IV. ACCURATE LOCATION ESTIMATION BASED ON MODIFIED PARTICLE FILTER

In this section, we will turn our attention to the proposed accurate localization algorithm, which is based on a modified particle filter. Particle filter represents a robots belief by a set of weighted particles. Due to its easy to implement and tend to work well across a broad range of problems, it has already become one of the most popular localization algorithms in robotics. Particle-based localization is not bound to a limited parametric subset of distribution [18]. In its present form, it deals with the global localization problem but hardly affords robot kidnapping or global localization failures. All particles survive at the most likely pose and others gradually disappear as the position is acquired. Once observations are corrupted for a long period of time or the pose of robot happens to be incorrect, it will inevitably lead to particle impoverishment or localization failure. Fortunately, this problem can be solved by adding a set of additional particles to the resample process [19], [20]. However, the approach of injecting particles raises a series of questions as follows:

(a) From which distribution should we generate the particles, random or not?

(b) How many particles should be added at each iteration of the algorithm?

(c) When should we inject the particles?

In this paper, we introduce a particles injection scheme that addresses all of three questions. To enhance the speed of convergence in mobile robot localization, we designed a particles production mechanism which is based on the estimation of the robots location. A set of pointed particles is generated according to the results of the place recognition. We should point out that these particles are not generated randomly around these three locations. They are limited to circle areas with the locations as the centers within a 1.5 m radius. These particles centered the locations and obeyed to Gaussian distribution with an average weight. These pointed particles will be injected into the localization system when necessary instead of random. To determine the time to inject the additional particles, the current probability distribution of the particles is monitored in real time. In this paper, Information Entropy (IE) is employed to describe the uncertainty of probability distribution [21]. The information entropy is defined as follows.

$$H = -\sum_{i=1}^{m} w_i \cdot \log w_i \tag{7}$$

where w_i is the weight of the particle, and m is the number of the particles. For example, if the pose of the robot is achieved, most of the particles are rounded together. The weights of the particles are in similar value. However, if most of the particles are shifted to a wrong location as the time goes on due to corrupted observations in the populated environments, the weights of the particles are assigned with a low value. Hence, the value of the information entropy when the localization has failure is smaller than that in the stage of localization success. Therefore, when the information entropy changed greatly in the process of localization, the robot is more inclined to localization failure. Moreover, the Effective Sample Size (ESS) is used to count the number of the effective samples in the pose estimation. ESS is calculated by (7) [22].

$$ESS = \frac{m}{1 + cv^2} \tag{8}$$

$$cv^2 = \frac{\mathrm{var}(w_i)}{E^2(w_i)} = \frac{1}{m}\sum_{i=1}^{m}(m \cdot w_i - 1)^2 \tag{9}$$

Once the ESS below a certain threshold which means the number of the effective particles is in a low level, the sample particles are more likely degeneracy. Hence, we can examine these two parameters periodically to determine whether to inject pointed particles or not. In addition, the number of the added particles can be calculated according to the difference between the total number of particles and ESS. We should clear that once the number of the effective particles is below a certain threshold, the coarse place recognition is activated immediately. It does not mean that there are no effective particles. At the same time, even if the coarse place recognition is failed in this stage and the additional particles may lead to the robot kidnapped problem, the coarse place recognition will be conducted in the next stage continuously. The robot will recover its position from the perspective of iteration finally.

V. EXPERIMENTAL RESULTS

The approach described above has been implemented on a real mobile robot (see Fig. 3) and tested intensively in an office environment. The robot is nurse-robot that developed by service robots laboratory of Shandong University. The size of the robot is about 0.5m x 0.5m x 1.4m, and it has two differential driving wheels and a universal wheel. The camera used to grab images is equipped at the head of the robot. The laser radar is embedded in the front of the robot. It gathers 180 degrees laser data ahead of the robot. The image database contained 120 images. They were captured purposefully through the environment. Each of images was bounded with a coarse location in advance, which was estimated using laser radar. The localization method is processed on a 2.0 GHz laptop with 1G memory, which is the algorithm process unit of the robot.

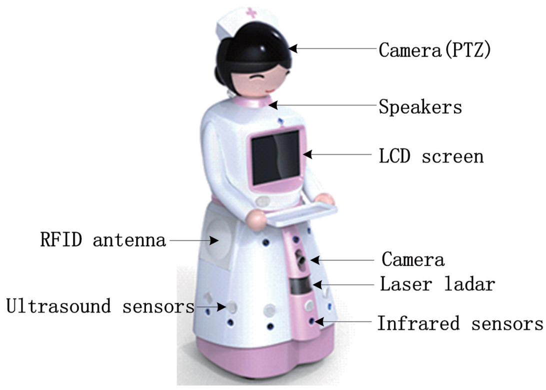

Fig. 3. Nurse-robot used to validate the proposed localization approach.

A. Coarse Place Recognition Results

The size of the image is 720 x 576. The camera is calibrated in advanced and the internal parameters do not change during the stage of mobile robot navigation. Fig. 4 illustrates several image retrieval results in the stage of the coarse place recognition. In the each row, the first image is the image captured in the process of mobile robot navigation, and the others are re-ranked retrieval results produced as the most similar top three matches. The correct ration of the image retrieval system is about 90.1% for the top three re-ranked images. As shown in Fig. 4 (a), there are some occlusions, and the correct images are re-ranked by our retrieval system. The image taken under bad illumination condition (in Fig. 4 (b)) is correctly retrieved from the database. The image grabbed at a different viewpoint (in Fig. 4 (c)) is also correctly found in the database.

The localization results show that the proposed coarse localization is robust against different viewpoint and illumination changes. The computation time of this stage is less than 1.3 seconds for a vocabulary that contains 2048 visual words. Indeed, the time for visual words assignment is the most consuming process which depends on the number of features detected on the input image and the number of visual words in the vocabulary.

B. Accurate Location Estimation

Based on the epipolar geometry, the coarse place recognition stage will output three locations which correspond to all three input images, respectively. At these locations, a set of pointed particles is generated around these locations and then injected into the accurate localization stage to increase the diversity of samples.

To show the effectiveness of the proposed accurate location estimation, comparative experiments among particle filter based localization [18], mixture Monte Carlo Localization (MCL) [19], and the proposed localization method are conducted for mobile robot global localization. The size of the office room is 15.3 x 9.5 m^2. In these experiments, the mobile robot is ordered to walk about 250 mm/s in speed in an office room. At first, the initial pose of the robot is unknown in global localization, and we initialized 1000 samples with a uniform distribution. All particles are distributed randomly over the map originally. After several times of coarse place recognition, the particles have almost centered on the correct position of the robot. Finally, the position of the robot is uniquely determined after about thirty-two iterations of the proposed algorithm. In these experiments discussed above, the real trajectory of the robot is recorded using a landmark detection sensor: StarGazer [23]. This sensor analyzes infrared ray image which is reflected from the passive landmark to calculate the relative position between the sensor and the landmark. It directly outputs the position and heading angle of the sensor with a precise resolution and high speed. The localization accuracy based on this kind of sensor is below 5 cm. At the same time, the estimated position of the robot is determined by the proposed localization method. The localization error is defined as the difference of the two trajectories. Fig. 5 shows a comparison on the localization errors among three localization methods. As shown in Fig. 5, the proposed

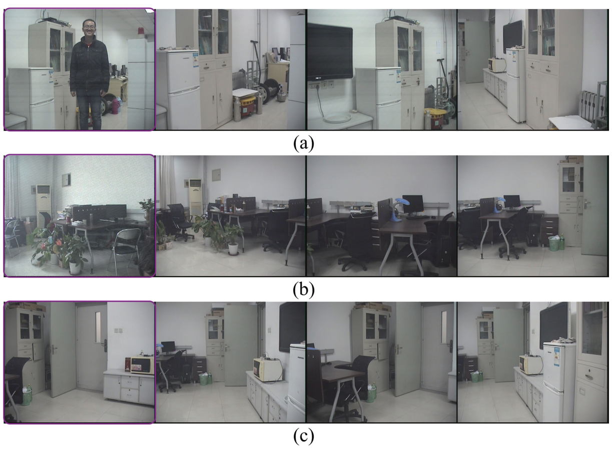

Fig. 4. Localization results in the stage of the coarse place recognition. (a) image with occlusions. (b) image with bad illumination. (c) image with different viewpoint.

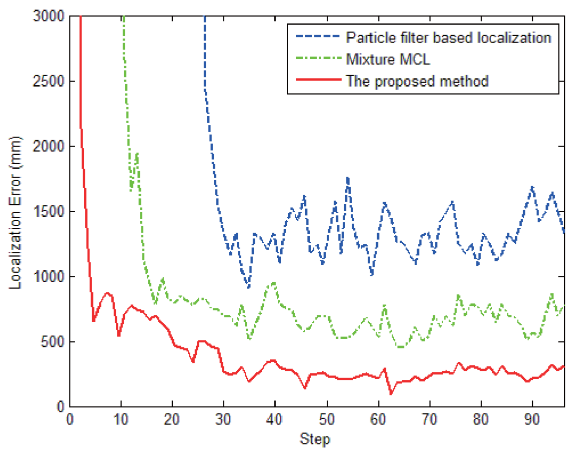

Fig. 5. Localization errors among the three methods in global localization.

localization method is able to determine the pose of the robot rapidly compared to the two other localization methods, and to keep track of the position reliably afterwards. The average localization error of the proposed method is less than 40 cm. Moreover, the error variation is far less than the other two methods. There are two reasons that may contribute to the fast and robust mobile robot localization. One is the closed-loop coarse-to-fine mechanism, which is the basis of the successful localization. Another is the particle-based position representation and the current particles state monitoring implementation. In this way, the status of the localization is known timely. Moreover, this pattern representation is more suitable for populated environment.

The estimated trajectory of the robot in the experiment of global localization is illustrated in Fig. 6. In this part, we employed the mean of the particles to estimate the position of the robot. Hence, the estimate position is close to the center of the map at the beginning. After several steps of place recognition, a certain number of pointed particles are injected into particles set. A growing number of particles tend to converge to the real position of the robot. Due to high uncertainty at the beginning of global localization, the estimated trajectory contains straight lines initially. After steps of iteration, the estimated pose of the robot will be restrained near to the correct location.

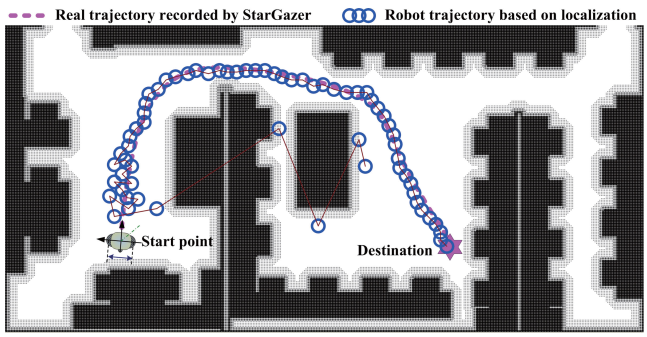

Fig. 6. The estimated robot trajectory in global localization.

VI. CONCLUSIONS

In this paper, we present a novel multi-sensor-based robot localization framework inspired by a closed-loop coarse-to-fine recognition mechanism which consists of two parts: coarse place recognition and accurate location estimation. On one hand, the coarse place recognition implemented using an onboard camera is just to provide a guidance of particles injection. On the other hand, a modified particle filter which is mainly based on the laser radar data is employed to access to the accurate location estimation. The current states of the particles are monitored to determine the time of making place recognition. The integration of coarse and accurate localization just like people recognize environment makes fast and reliable mobile robot localization fast. Compared to conventional localization methods, the presented approach has exhibited great improvement on speed and stability of robot localization.

REFERENCES

[1] P. Jensfelt, "Approaches to Mobile Robot Localization in Indoor Environments", *Tekniska hogsk*, 2001.

[2] S. Thrun, D. Fox, W. Burgard and F. Dellaert, "Robust Monte Carlo Localization for mobile robots", *Artificial Intelligence*, vol. 128, pp. 99-141, 2001.

[3] J. Wolf, W. Burgard, H. Burkhardt, "Robust Vision-Based Localization by Combining an Image Retrieval System with Monte Carlo Localization", *IEEE Transactions on Robotics*, vol. 21, pp. 208-216, 2005.

[4] Minyong Choi, Jinwoo Choi and Wan Kyun Chung, "Correlation-Based Scan Matching Using Ultrasonic Sensors for EKF Localization", *Advanced Robotics*, vol. 26, pp. 1495-1519, 2012.

[5] Tingting Liu, et al., "A Laser Radar Based Mobile Robot Localization Method", *Proceeding of the IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Shenzhen, China, pp. 2511-2514, December 2013.

[6] Joydeep Biswas and Manuela Veloso, "Multi-sensor Mobile Robot Localization for Diverse Environments", *Proceedings of the RoboCup Symposium*, August, pp. 1-11, 2013.

[7] Jung-Suk Lee and Wan Kyun Chung, "Robust Mobile Robot Localization in Highly Non-static Environments", *Autonomous Robots*, vol. 29, pp. 1-16, 2010.

[8] Karthick Srinivasan, Jason Gu, "Multiple Sensor Fusion in Mobile Robot Localization", *Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1207-1210, April, 2007.

[9] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Coarse-to-fine Pedestrian Localization and Silhouette Extraction for the Gait Challenge Data Sets", *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, Madison, pp. 1009-1012, 2006.

[10] M. Pedersoli, A. Vedaldi, J. Gonzlez, "A Coarse-to-fine Approach for Fast Deformable Object Detection", *Proceedings of the 24th IEEE Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, CO, June, pp. 1353-1360, 2011.

[11] X. Feng, A. Hadid, and M. Pietikainen, "A Coarse-to-fine Classification Scheme for Facial Expression Recognition", *Proceedings of the International Conference on Image Analysis and Recognition (ICIAR)*, Porto, Portugal, pp. 668-675, 2004.

[12] Soonyong Park, Howon Cheong, and Sung-Kee Park, "Coarse-to-Fine Global Localization for Mobile Robots with Hybrid Maps of Objects and Spatial Layouts", *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, St. Louis, USA, pp. 3993-4000, 2009.

[13] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features", *Computer Vision-ECCV 2006*, Springer Berlin Heidelberg, pp. 404-417, 2006.

[14] Y. Rui, T. Huang, and S. Mehrotra, "Content-based Image Retrieval with Relevance Feedback in MARS", *Proceedings of the International Conference on Image Processing*, vol. 2, 1997.

[15] Yannis Kalantidis, Giorgos Tolias, Yannis Avrithis, et al., "Visual Image Retrieval and Localization", *Multimedia Tools and Applications*, Volume 51, Issue 2, pp 555-592, January 2011.

[16] M. Fischler and R. Bolles, "Random Sample Consensus - a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", *Communications of the ACM*, vol. 24, no. 6, pp. 381C395, 1981.

[17] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision", *Cambridge University Press*, 2001.

[18] Sebastian Thrun, Wolfram Burgard, Dieter Fox, "Probabilistic Robotics", *MIT Press*, 2005.

[19] S. Thrun, D. Fox, and W. Burgard, "Monte Carlo localization with mixture proposal distribution", *Proceedings of the AAAI National Conference on Artificial Intelligence*, pp. 859-865, 2000.

[20] D. Fox, "KLD-Sampling: Adaptive Particle Filters", *Advances in Neural Information Processing System*, MIT Press, 2001.

[21] Zhiwei Liang, Xudong Ma, Xianzhong Dai, "An Extended Monte Carlo Localization Approach Based on Collaborative Distributed Perception", *ROBOT*, vol. 30, no. 3, pp. 210-216, 2008.

[22] J. S. Liu, R. Chen, T. Logvinenko, "A Theoretical Framework for Sequential Importance Sampling and Resampling", *Sequential Monte Carlo in Practice*, New York, USA: Springer-Verlag, 2001.

[23] *http://www.hagisonic.com/*.