The 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2014)
Nov. 12-15, 2014 at Double Tree Hotel by Hilton, Kuala Lumpur, Malaysia

Image-based Localization Using Prior Map Database and Monte Carlo Localization

Hyongjin Kim, Taekjun Oh, Donghwa Lee, and Hyun Myung
Department of Civil and Environmental Engineering, KAIST, Daejeon, 305-701, Korea
(Tel : +82-42-350-5670; E-mail: {hjkim86, buljaga, leedonghwa, hmyung}@kaist.ac.kr)

Abstract – The aim of this paper is to propose an image and map data-based localization method applicable to a variety of environments. For the localization, we use prior map database, image-based localization method, and MCL (Monte Carlo Localization). The results were confirmed by open data set in a variety of environments. The experimental results show the feasibility of the proposed method for the robot localization.

Keywords – Image-based Localization, Image Feature, Prior Map Database, MCL

1. Introduction

In order for the robot to do something on its own, the robot has to know where the robot is. This problem is referred to as the localization and important in robotics area [1, 2]. The robot localization has received much attention and a number of studies are underway.

For the robot localization, a method using a GPS (Global Positioning System) and an IMU (Inertial Measurement Unit) is usually used [3]. However, in places like the tunnel or around high building, GPS error is inevitable because GPS signal is not received correctly. Also the IMU accumulates the error in a long run time. In this case, another method is necessary and the localization method using prior map database is proposed as a solution to this problem.

There are only a few works related with localization method using prior map database. In Oxford University, Newman’s research team presented a work on robust image matching using BoW (Bag of Words) by generating vocabulary of feature [4, 5]. However this approach is not able to find the exact location by just finding the image that is most similar in the collected data in the past. Therefor to overcome this problem Napier et al. [6] presented a research to save the 3D information using stereo vision, and localize the robot pose using features in the road. However, in areas other than roads and in roads with no characteristic points, it is impossible to apply the algorithm.

In this paper, we propose a novel robot localization method using the image feature-based algorithm applicable to a variety of environments in order to overcome this problem.

The remainder of this paper is structured as follows. In Section 2, the details on the method for image-based localization are described. To validate the performance of our method, the overall experimental setup is provided and localization experiments are conducted and described in Section 3. In Section 4, conclusions and future works are discussed.

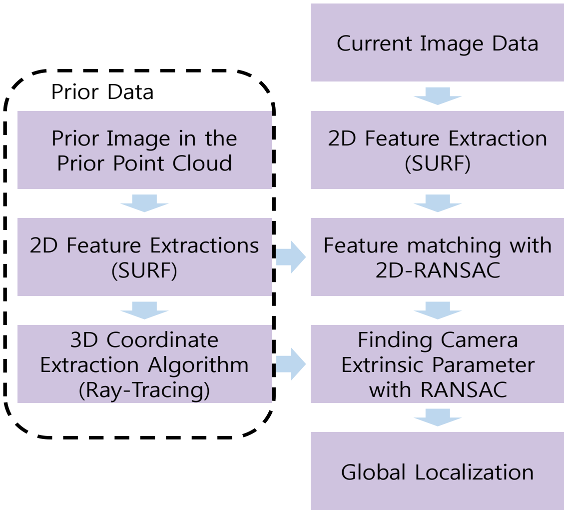


Fig. 1. The overview of the image-based global localization algorithm.

2. Image-based Localization

This section provides key concepts used in our method. We introduce the method for image-based localization. First, we present how to generate the map database as a prior data for robot localization. Then we present the global localization algorithm. Finally, we present the localization algorithm incorporating the MCL (Monte Carlo Localization) for more robust and efficient operation.

2.1 Map Database

The map database is composed of the 3D Point cloud and image data. The image data are feature points extracted by the SURF (Speeded Up Robust Features) algorithm [7]. The extracted feature points can be connected to the camera origin by lines in 3D point cloud using camera parameters. The 3D coordinates of feature points are defined as 3D point cloud data coordinate met by ray-tracing algorithm from camera origin. The method of applying the ray-tracing algorithm for all point cloud data has a computation burden. Therefore, we use an octree structure to reduce the computation time. The running time of the algorithm is decreased by constructing a map database combining 2D and 3D coordinates of extracted feature points in preprocessing step.

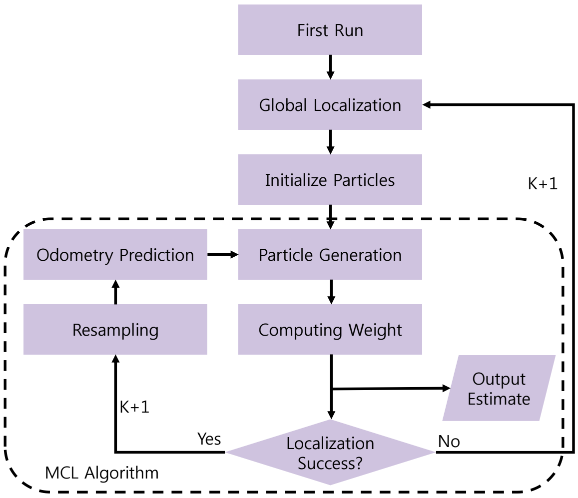


Fig. 2. The flowchart of the localization using the MCL algorithm.

2.2 Global Localization

The global localization algorithm is to estimate extrinsic camera parameters using 2D and 3D matching pairs and its overview is shown in Fig. 1. The image from the camera matches features which are extracted using the SURF algorithm in advance. The outliers which are not the valid matching pairs are removed by 2D-RANSAC (RANdom SAmple Consensus) algorithm to improve the accuracy in this process. As mentioned in Section 2.1, the feature points of the database can match 2D with 3D data because feature points of the database have a 3D coordinate data. Then we estimate 6D pose of camera origin using the 6-Point-DLT [8] algorithm because we know the 2D-to-3D matching pairs and camera's intrinsic parameters. In order to remove the error existing in the matched pairs, RANSAC algorithm based 6-Point-DLT is used and the accuracy is improved. Finally, the robot pose is estimated using the transformation between the camera and robot's origin.

2.3 Localization using MCL

The global localization algorithm has a computational burden and fails in some cases because the algorithm processes for all databases. Thus the failure of the pose estimation and the computation time are reduced with the minimum amount of data by using the MCL algorithm to estimate the position. The flowchart of the localization using the MCL algorithm is shown in Fig. 2. On the basis of the MCL framework, the algorithm consists of the particle generation, the weight calculation, and the resampling. The initial value of the particle is result of the global localization in Section 2.2 and the particles are generated with the addition of Gaussian noise. The most important part of the MCL algorithm is to select the weight of each particle. In this paper, we select weight by simplifying the global position estimation algorithm of Section 2.2. Similar to Section 2.2, the 2D-to-3D matching pairs are calculated. x_i , X_i are defined coordinates on both 2D and 3D, respectively. The weight w_i of each particle is defined by following equation:

$$\frac{1}{w_j} = \frac{1}{N}\sum_{i=1}^{N} Dist(x_i, PTX_i) \qquad (1)$$

where, j is the index of particle and i is index of the 2D-to-3D matching pair and N is the number of matching pairs. *Dist(x, y)* indicates Euclidean distance between x and y. P is the matrix composed of the camera's intrinsic parameter and T is the transformation matrix of each particle. The weight is generated by comparing the distance between the 2D coordinate and the projection of the 3D coordinate of each feature point on the image plane. The pose estimation is processed by finding the best particle using highest weight. After pose estimation, resampling is performed using weights and stochastic universal sampling method. The success of the pose estimation is determined by the weight of best particle. In case of a success of pose estimation, the particle is reproduced using prediction part, otherwise the particle is reproduced using the global localization result.

3. Experiments

In this section, we describe our experimental setup and present the performance of our approach. For the experiments, we used an open data set called KITTI data set [9] and conducted experiments.

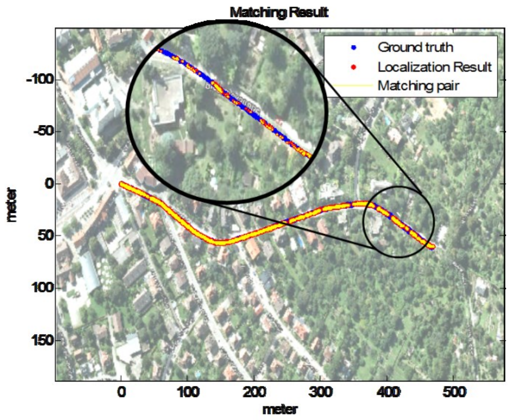
3.1 Experimental Setup

In order to assess the performance of our proposed method, KITTI data set [9] was used. KITTI data set [9] is a data set collected using the vehicles equipped with the camera, Velodyne, and RTK-GPS in the city of Karlsruhe. The data catalog is composed with a total of 152 data sets which include city, residential, road campus and person. In this paper, city and residential catalog data set from a total of eight trials were extracted whose data acquisition time is 18 minutes and total travel distance is 9,141 m.

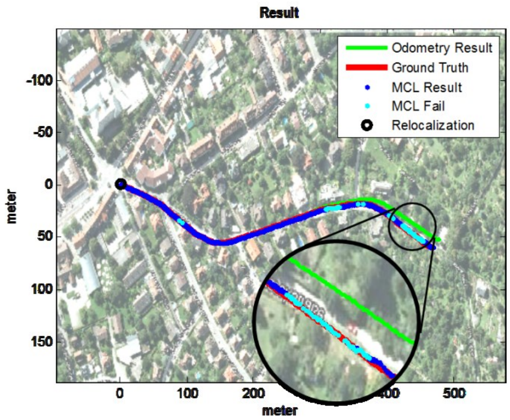
3.2 Results

Fig. 3 (a) shows the results of global localization in Google Map using data set #5. The blue dots represent the ground truth and the red dots global localization results. The yellow line indicates the matching pairs between the blue dots and the red dots. The results of MCL are shown in Fig. 3 (b). The green line indicates the odometry results and the red line ground truth. The blue dots are MCL results and the blue-green dots the results of MCL in areas where the global position estimation fails.

Table 1 shows the computation time of each algorithm. The computation time of MCL algorithm is shorter than the global localization. Table 2 shows the results for each data set. The error of the MCL is higher than the global localization error, because we used a few feature points for running MCL with reduced computation time. It is impossible for global localization to use a few features. As shown in Table 2, the global localization fails despite the



(a)



(b)

Fig. 3. Image-based localization results using KITTI data set (a) global localization algorithm, (b) MCL algorithm.

use of many feature points. Thus it is expected to obtain better performance by combining global localization and MCL algorithm.

Table 1 Computation time.

	Global localization	MCL
Computing time (s)	2.012	0.195

Table 2 Results with each data set

Data Set	Global localization		MCL
	Failure rate (%)	Position error (cm)	Position error (cm)
Data Set #1	4.55	2.4	26
Data Set #2	2.57	0.79	45
Data Set #3	1.07	4.4	36
Data Set #4	10.67	1.	44
Data Set #5	41.30	1.2	35
Data Set #6	39.38	5	17
Data Set #7	22.44	0.4	43
Data Set #8	6.31	2.3	14

4. Conclusion and Future Work

In this paper we presented the method for the image-based localization applicable to a variety of environments. It was confirmed that results of the global localization is accurate for the robot localization and the computation time of the MCL is small. The global localization algorithm and MCL algorithm are complementary to each other.

As a future work, we will fuse global localization and MCL algorithm.

Acknowledgement

This work was supported by the R&D program of the Korea Ministry of Trade, Industry and Energy (MOTIE) and the Korea Evaluation Institute of Industrial Technology (KEIT). (The Development of Low-cost Autonomous Navigation Systems for a Robot Vehicle in Urban Environment, 10035354).

References

[1] R. Smith, and P. Cheeseman, "On the Representation and Estimation of Spatial Uncertainty," International Journal of Robotics Research (IJRR), Vol. 5, No. 4, pp. 56-68, Winter 1987.

[2] M. Obst, S. Bauer, P. Reisdorf and G. Wanielik, "Multipath Detection with 3D Digital Maps for Robust Multi-constellation GNSS/INS Vehicle Localization in Urban Areas," in Proc. Intelligent Vehicles Symposium (IV), pp. 184-190, 2012.

[3] E. S. Jones and S. Soatto, "Visual-inertial navigation, Mapping and Localization: A Scalable Real-time Causal Approach," International Journal of Robotics Research (IJRR), Vol. 30, No. 4, pp. 407-430, 2011.

[4] M. Cummins, and P. Newman, "FAB-MAP: Probabilistic localization and mapping in the space of appearance," International Journal of Robotics Research (IJRR), Vol. 27, No. 6, pp. 647-665, 2008.

[5] R. Paul, and P. Newman, "FAB-MAP 3D: Topological mapping with spatial and visual appearance," " in Proc. of the IEEE International Conference on Robotics and Automation (ICRA), pp. 2679-2656, 2010.

[6] A. Napier and P. Newman, "Generation and exploitation of local orthographic imagery for road vehicle localisation," in Proc. Intelligent Vehicles Symposium (IV), pp. 590-596, 2012.

[7] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, "SURF: Speeded Up Robust Features," Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346–359, 2008.

[8] R. I. Hartley, and A. Zisserman, "Multiple View Geometry in Computer Vision," Cambridge Univ. Press, 2nd edition, 2004.

[9] A. Geiger, P. Lenz and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in Proc. of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3354-3361, 2012.