The 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2015)
October 28 ∼ 30, 2015 / KINTEX, Goyang city, Korea

Image-based Localization using Image Database and Local 3D Maps

Hyungjin Kim[1], Kwangyik Jung[1], Taekjun Oh[1] and Hyun Myung[1]

[1]Urban Robotics Lab. (URL), Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro (373-1 Guseong-dong), Yuseong-gu, Daejeon 34141, Korea; (Tel : +82-042-350-5670;
E-mail: {hjkim86, ankh88324, buljaga, and hmyung†}@kaist.ac.kr)

Abstract - Image-based localization is one of the most important techniques for autonomous navigation. This paper proposes a 6-DoF localization method based on image database and local 3D maps. To perform localization, the image database and local 3D map of each image are collected in advance. By applying a fast place recognition algorithm such as FAB-MAP, an initial position is assumed to be available at a past position of a matched image from the database. Then, a 6-DoF camera pose is estimated using PnP algorithm by 2D-to-3D matching correspondences in a query image and a local 3D map. The proposed algorithm is demonstrated in real environment.

Keywords - Image-based localization, 2D-to-3D matching correspondences, FAB-MAP, camera pose estimation

1. Introduction

A localization is one of essential techniques for autonomous navigation in robotics [1]. GPS (global positioning system) is widely applied for localization due to provide a absolute position. However, there are still barriers to utilize in indoor environments or signal interfering areas. Therefore, Image-based localization is widely researched as a substitution of GPS in robotics field over recent years. Image-based localization method is necessary to collect database in advance for localization unlike GPS [2]. Then, it is important to search similar images rapidly in the database. M. Cummins et al. [3] proposed FAB-MAP as robust place recognition algorithm based on BoW (bag-of words). FAB-MAP is able to find the most similar image from enormous image database in real time. Although image-based localization is able to cover large areas from FAB-MAP, the algorithm is not available to estimate a exact location. Therefore, we propose a method to estimate 6-DoF (degree-of-freedom) camera pose by matching from a local 3D map of the database.

In this paper, an image database and local 3D map are collected in advance with a stereo vision system. At first, FAB-MAP algorithm is utilized to find the most similar image from the image database as initial search. Next, 2D-to-3D correspondences are generated from local 3D map of the most similar image. Finally, 6-DoF camera position is estimated by PnP (perspective-n-point) algorithm [4].

The remainder of this paper is organized as follows: section 2 introduces our propose method, section 3 presents experiments, and section 4 describes conclusion.

2. Proposed localization method

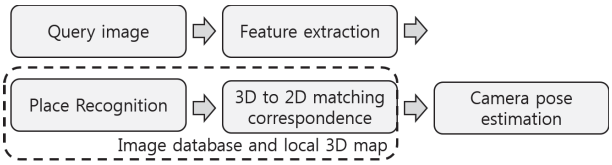


Fig. 1 Overview of the proposed image-based localization method.

A block diagram of the proposed method is shown in Fig. 1. In this paper, SURF (speeded up robust features) [5] is utilized as a image feature. A query image is collected from a current robot status. The camera position of query image is estimated by place recognition and 3D-to-2D matching.

2.1 Image database and local 3D maps generation

The image database is generated from a left camera of stereo vision system. Each image from database contains 3D local map using stereo vision technology. For generation of a 3D local map, image features are extracted from both sides of the stereo vision. As known the baseline and transformation of the stereo vision system, each depth value of the matched features on the image is estimated. Therefore, a 3D local map of each image from database is generated.

2.2 FAB-MAP place recognition

To search the most similar image in the image database, FAB-MAP algorithm is performed. Each BoW of the image database is learned as training set of FAB-MAP. For place recognition process, the best similar image is robustly found as shown in Fig. 2 by extracting BoW from the query image.

(a) (b)

Fig. 2 The FAB-MAP result of successful matched result.

2.3 Camera pose estimation

After searching the best similar image, a camera pose is estimated using PnP algorithm [4] from 2D-to-3D

978-1-4673-7971-7/15/$31.00 ©2015 IEEE

matching correspondences. To obtain the correspondences, the extracted features from the query image match to a local 3D map of the best similar image. To neglect outlier matching, RANSAC (random sample consensus) based PnP algorithm is performed for localization.

3. Experiments

The experiment was conducted at the National Science Museum located at Daejeon, South Korea. A survey vehicle equipped with two camera systems as stereo vision and a GPS receiver for ground truth. The survey vehicle run two laps as one for training set and one for test set. The red numbers in Fig. 3 denote estimated depth value using stereo vision algorithm. Fig. 4 shows confusion matrices of FAB-MAP result with 98% and 99% of similarity. Table 1 represents matching success or fail result of FAB-MAP. Fig. 5 shows matched results of FAB-MAP on Google map. Final camera pose estimation result is 1.47m as a mean value of 367 positions.

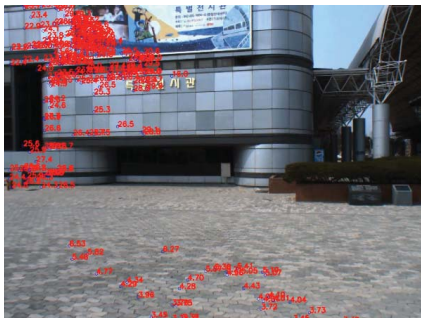
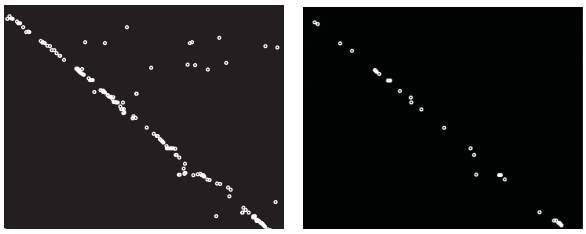


Fig. 3 Estimated depth values from the stereo vision system. The red numbers denote a depth value of each feature.



(a) (b)

Fig. 4 Confusion matrices results of FAB-MAP with 99% and 98% of similarity.

Table 1 True and false matching results of FAB-MAP with 99% and 98% of similarity.

Similarity	Total number	True positive	False positive
98%	703	159	18
99%	703	74	1

4. Conclusion

This paper proposed a image-based localization using image database and local 3D maps. The image database and local 3D maps were generated by a stereo vision system in advance. For a initial estimation of pose, FAB-MAP was conducted as searching the best similar image. 2D-to-3D matching correspondences were found from the query image and local 3D map of the best similar image on the image database. Using PnP algorithm, the final camera pose was estimated. The proposed algorithm was demonstrated from a real experiment.

Fig. 5 Matched position results of FAB-MAP. The red and blue dots denote positions of the query image and matched image respectively.

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). This research was also supported by the project, "Development of basic SLAM technologies for autonomous underwater robot and software environment for MOSS-IvP" sponsored by Korea Research Institute of Ships & Ocean Engineering (KRISO), Korea. The students are supported by Ministry of Land, Infrastructure and Transport (MoLIT) as U-City Master and Doctor Course Grant Program.

References

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

[2] H. Kim, T. Oh, D. Lee, Y. Choe, M. J. Chung, and H. Myung, "Mobile robot localization by matching 2d image features to 3d point cloud", *In Proc. of Int'l Conf. on Ubiquitous Robots and Ambient Intelligence*, pp. 266-267, 2013.

[3] M. Cummins and P. Newman, "FAB-MAP: Probabilistic localization and mapping in the space of appearance," *Int'l J. of Robotics Research*, Vol. 27 No. 6, pp. 647-665, 2008.

[4] R. Hartley and A. Zisserman, Multiple view geometry in computer vision, Cambridge University Press, 2003.

[5] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346-359, 2008.