International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.11, (2016), pp.25-36
http://dx.doi.org/10.14257/ijsip.2016.9.11.03

Visual Location Recognition Based on Coarse-to-Fine Image Retrieval and Epipolar Geometry Constraint for Urban Environment

Guanyuan Feng[1,2], Lin Ma[1,2], Xuezhi Tan[1,2], Hao Xue[1,2] and Kai Guan[1,2]

[1]Communication Research Center, Harbin Institute of Technology,
Harbin 150080, China
[2]Key Laboratory of Police Wireless Digital Communication, Ministry of Public Security, People's Republic of China, Harbin 150080, China
fengguanyuan@126.com

Abstract

Visual based location recognition of a mobile device is an important problem in many applications, such as visual navigation, auto-piloted driving and augmented reality. In this paper, a visual location recognition system based on the Coarse-to-Fine image retrieval and the epipolar geometry constraint is proposed. The basic idea of this system is to match a user captured image against some geo-tagged images in the database, and then estimate the user's location by the epipolar geometry constraint. The process of the Coarse-to-Fine image retrieval is necessary to select some database images in the same scene with the user captured image. The epipolar geometry constraint is utilized to determine the refined location using the geographical location information of the database images. The specific experiments for the visual location recognition are performed and the results show that this system can achieve the excellent performance of the location recognition.

Keywords: *visual location recognition, image retrieval, features matching, epipolar geometry constraint*

1. Introduction

Location recognition is an important problem with many civilian and military applications, such as vehicle navigation and path planning. The most commonly used device is the GPS in the outdoor environment in recent years. The point is that the information provided by the GPS device is not stable sometimes, especially in the business district where tall buildings block the satellite view and the multi-path effects are severe. Therefore, an effective location recognition method is required as an alternate approach of the GPS device. Visual localization is a great choice.

Visual localization methods derive from the field of robotics control. The computer vision groups have researched visual location recognition for a long time, mainly in the field of mobile robots [1-3]. In recent ten years, visual localization researchers mainly focus on the image retrieval for location recognition using mobile terminals, such as smartphones and tablet PCs. Recent advances in the field of the content-based image retrieval (CBIR) have made it facilitative to quickly search large image databases using pictures or video sequences captured by the user's smartphone as a query. With appropriately tagged images of precise location information, the CBIR technique is able to be applied to the visual location system. Some of these works focus on the outdoor environments. The authors in [4-5] propose the localization approaches based on street view images. When the user takes a picture of an unfamiliar place, the proposed algorithm can recognize the

ISSN: 2005-4254 IJSIP
Copyright ⓒ 2016 SERSC

user's current location by the means of the images retrieval from an image database. Location recognition methods based on the CBIR technique continue to develop and evolve. A rapid image retrieval method used in location recognition is proposed in [6], the process of the image retrieval based on Bag-of-Feature achieves low query time. The Multiple Hypothesis Vocabulary Tree is introduced to reduce the complexity of the feature quantization. The authors in [7] present a completed image retrieval based pipeline for the visual location recognition.

In the location recognition methods based on the CBIR technique, the location of query image is usually assigned to the most neighboring location of the database image. However, the performance of the location recognition based on the CBIR technique is restricted to the density of the database images. That is to say, the more densely database images are captured, the more exactly user's location can be determined. But in most cases, it is inconvenient to collect a great number of database images. What is more, sometimes the query image and the database image contain common objects while query image is far away from the database image. In this case, the location errors are magnified.

Inspired by these problems, we propose a visual location system based on the Coarse-to-Fine image retrieval and the epipolar geometry constraint which achieves a higher location precision in a scene of relatively sparse database images, and the location errors are effectively limited. The basic idea of the epipolar geometry-based method for location recognition is proposed in [8] by Sadeghi for the indoor environment. But in Sadeghi's paper, only the indoor environment is considered and the timeliness of the image retrieval is not mentioned. For the urban environment, the image databases usually are very large, and the operation time of the database searching is long. So in this paper, an efficient image retrieval called the Coarse-to-Fine image retrieval is proposed. Combined with the Coarse-to-Fine image retrieval and the epipolar geometry-based approach, an integrated visual location recognition system in the urban environment is presented. In our system, the Coarse-to-Fine image retrieval is used to select the database images in the same scene with the query image. In this phase, the database images which contain common objects with the query image are selected. Subsequently, the epipolar geometry constraint is utilized to refine the location of the query image. In this phase, a precise location of the query image can be determined in the situation with relatively sparse database images.

The reminding parts of this paper are organized as follows. We explain the database generation in Section 2, and describe the Coarse-to-Fine image retrieval in Section 3. Section 4, presents the visual location recognition method in detail. We evaluate our experiment results in Section 5, and conclude in Section 6.

2. Database Generation Analysis

This section provides a detailed description of the database generation. We mount three CCD cameras, a laser scanner, a GPS device (sub-meter precision) and an inertial measurement unit (IMU) onto our mobile data acquisition platform which is carried by a human operator. Figure 1 shows the schematic diagram of the mobile data acquisition platform. Using the mobile data acquisition platform, the database images are captured by the CCD cameras. At the same time, the geographical location information and the 6 DOF information of the database camera are measured separately by the GPS device and the IMU.

In this paper, we consider the focus lengths of the each database camera are same, and the focus lengths are always fixed. All the cameras have been calibrated already. The goal of calibration is to measure the intrinsic and extrinsic parameters of the camera. We also assume that the query image plane and the database image plane are perpendicular to the ground plane. The calibration process between the CCD camera coordinate and the IMU device coordinate has already been done. The relation of translation and rotation between the world coordinate system (IMU device coordinate system) and the database camera coordinate system can be described as:

$$\mathbf{X}_w = \mathbf{R}_w \mathbf{X}_c + \mathbf{t}_w \tag{1}$$

where $\mathbf{R}_w$ and $\mathbf{t}_w$ are the rotation matrix and the translation vector between the world coordinate system and the database camera coordinate system. The 3D points coordinates in the world coordinate system and the database camera coordinate system are denoted by $\mathbf{X}_w$ and $\mathbf{X}_c$.

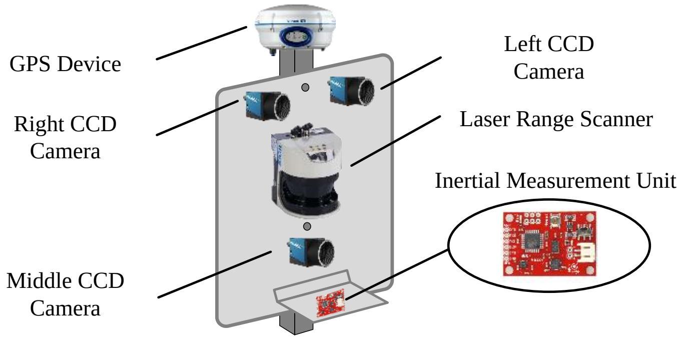


Figure 1. Schematic Diagram of Mobile Data Acquisition Platform

There are 5 scenes have been selected in the Harbin Institute of Technology (HIT) campus. For the each scene, 20 database camera locations are chosen uniformly. In the each camera location, six database images are captured in different orientations. At the same time, the database camera locations (the longitude and latitude values in the direction of X_W and Y_W) in the world coordinate system are recorded. In addition, 18 query images are also captured in the each scene. The query camera locations are recorded for the experiments. Figure 2, shows an example of the database generation in the Harbin Institute of Technology Main Building Scene. In this scene, the green circles denote the database camera locations and the yellow squares denote the query camera locations. There are some locations of the query camera are selected, and only one query image is captured to test the location recognition accuracy in the each location.

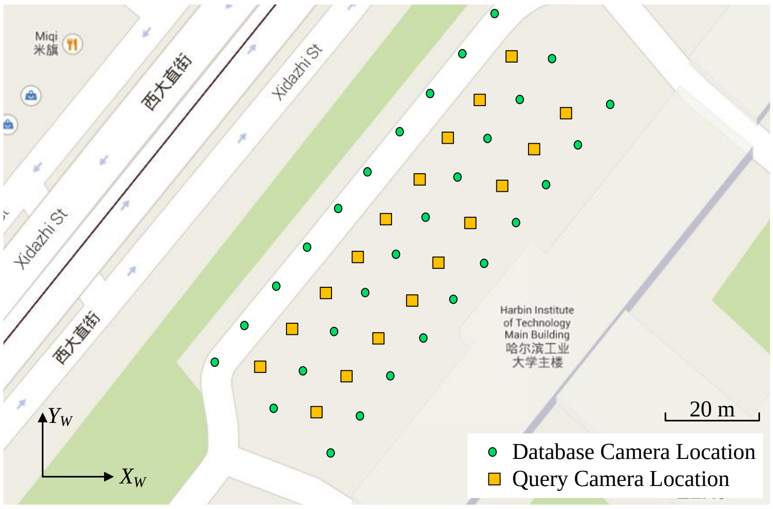


Figure 2. Description of Database Generation in HIT Main Building Scene

3. Coarse-to-Fine Image Retrieval

The aim of image retrieval is to select the matching images which contain the same object with the query image from the database. The image retrieval based on the global features (such as gist features) is fast, but not accurate enough. Whereas the image retrieval based on the local features (such as sift, surf or ORB features) is relatively slow, but more accurate. The integration of the global features and the local features makes fast and reliable retrieval possible. So a Coarse-to-Fine image retrieval is proposed in this paper.

In the Coarse image retrieval stage, the gist features [9] are utilized to select the database images which are similar with the query image. The gist features are extracted from every database images, and then we put these gist features into a kd-tree [10]. After that, we extract the gist features from the query image. For the gist features extracted from the query image, we find its top N neighbors in the kd-tree. In this paper, N is defined as 10. That means ten database images which are most similar to the query image are selected as the result for the Coarse image retrieval. But indeed, there may be some result images are not in the same scene with the query image. So the Fine image retrieval is needed to remove the database images which do not contain the same object with the query image.

The Fine image retrieval stage is carried out based on the results of the Coarse image retrieval. The top ten database images computed by the Coarse image retrieval are filtered in this stage by ORB features matching [11]. Specifically, we exact the ORB features from the query image and the database images. Next, the query image and the top ten database images are matched by the Euclidean standard metric in the ORB feature space, and the sum-of-squared pixel differences (SSD) is utilized. Using the procedure of the PROSAC [12] for random sampling of correspondence, we can obtain the inliers of the feature matching. Let MR denote the matching rate between the query image and the database image.

$$MR = \frac{N_i}{N_k} \tag{2}$$

where N_i denotes the number of the inliers of the feature matching, and N_k denotes the number of the ORB features of the query image. MR reflects the similarity between the query image and the database image. In this paper, the matching threshold ε is set to 0.52. If $MR \geq \varepsilon$, we consider that the database image in the Coarse image retrieval result is in the same scene with the query image (the query image contains the same object with the database image). Otherwise, we consider the database image is in the different scene with the query image.

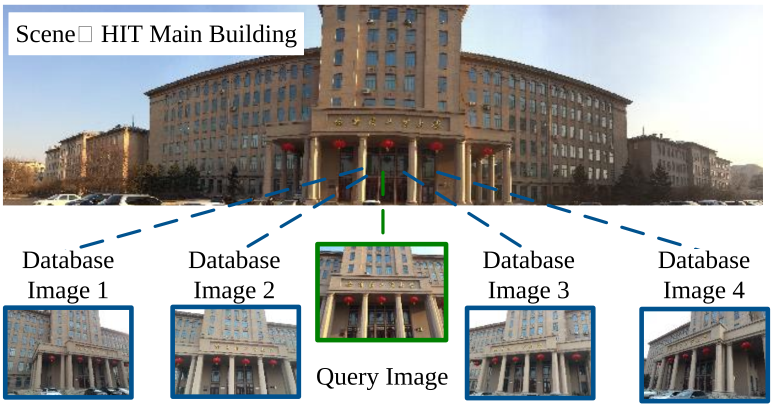


Figure 3. The Result of Coarse-to-Fine Image Retrieval

Figure 3, shows the result of the Coarse-to-Fine image retrieval result in the scene of the HIT Main Building. In this scene, the image of front door is captured as the query image by the user. By means of the Coarse image retrieval, ten most similar database images are selected. But some of these database images do not contain the same object with the query image. So after the Fine image retrieval, only four database images are remained as the image retrieval result. The database images in the image retrieval result contain the common objects with the query image.

4. Visual Location Recognition

By the Coarse-to-Fine image retrieval, the database images which are in the same scene with the query image are selected. In this stage, we try to determine the location of the query image by the database images in the retrieval result. The geographical locations of the database cameras are also used to calculate the location of the query image. Since the database images in the same scene contain the same object, the corresponding ORB features obey the epipolar geometry constraint.

In this paper, the database image camera and the query image camera have already been calibrated. The matrices $\mathbf{K}_1$ and $\mathbf{K}_2$ denote the database camera calibration matrix and the query camera calibration matrix, respectively. The ORB feature points $\mathbf{x}_{database}$ and $\mathbf{x}_{query}$ exacted from the database images and the query image can be normalized by the following form [13]:

$$\hat{\mathbf{x}}_{database} = \mathbf{K}_1^{-1}\mathbf{x}_{database} \tag{3}$$

$$\hat{\mathbf{x}}_{query} = \mathbf{K}_2^{-1}\mathbf{x}_{query} \tag{4}$$

The relations between the ORB feature points of the query image and the database image can be described as Equation (5):

$$\hat{\mathbf{x}}_{database}\mathbf{E}\hat{\mathbf{x}}_{query} = 0 \tag{5}$$

The essential matrix $\mathbf{E}$ is first proposed by Longuet-Higgns [14] for the structure-from-motion. The essential matrix is determined completely by the rotation matrix and the translation vector. In this paper, the essential matrix is used to present the location relationship between the database camera and the query camera. $\mathbf{E}$ contains the camera translation parameters between the database camera coordinate and the query camera coordinate in the following form:

$$\mathbf{E} \square \hat{\mathbf{t}}_E \mathbf{R}_E \tag{6}$$

where $\mathbf{t}_E$ denotes the translation vector and $\mathbf{R}_E$ denotes the rotation matrix. In Equation (6), $\mathbf{t}_E = [t_x, t_y, t_z]^{-1}$ and

$$\hat{\mathbf{t}}_E = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \tag{7}$$

The essential matrix $\mathbf{E}$ can be computed from the ORB feature correspondences of the query image and the database image. One efficient method to compute the essential matrix is Nister's five point algorithm proposed in [15]. According to the essential matrix $\mathbf{E}$, the translation vector $\mathbf{t}_E$ and rotation matrix $\mathbf{R}_E$ can be extracted using the singular value decomposition (SVD). The decomposition of $\mathbf{E}$ into $\mathbf{R}_E$ and $\mathbf{t}_E$ is described in Nister's paper.

In this paper, the essential matrix $\mathbf{E}$ is computed by the feature correspondences of the query image and the database image. The translation vector $\mathbf{t}_E$ and the rotation matrix $\mathbf{R}_E$ decomposed from the essential matrix $\mathbf{E}$ reflect the relationship of translation and rotation between the database camera coordinate and the query camera coordinate, which is shown in Figure 4. As shown in Figure 4, the epipolar geometry constraint is the intrinsic projective geometry between the database camera and the query camera. The database image and the query image are in the same scene, and they contain the same object P. In the epipolar geometry constraint, the object P, the feature point x_1 (in the database image) and the feature point x_2 (in the query image) are in the same epipolar plane. The form of the epipolar plane is a triangle. The object P, the database camera center o_1 and the query camera center o_2 are vertexes of the epipolar plane.

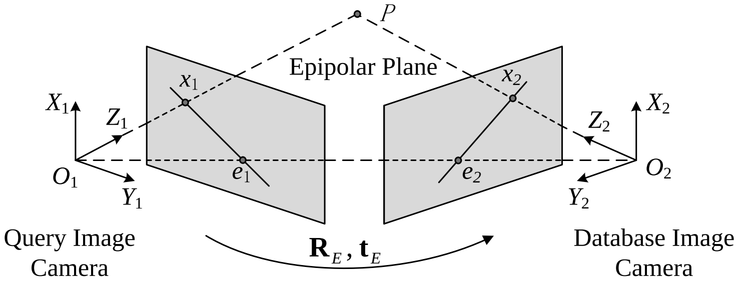


Figure 4. The Result of Coarse-to-Fine Image Retrieval

Under the condition of the epipolar geometry constraint, the 3D points in the database camera coordinate and the query camera coordinate systems are denoted by $\mathbf{X}_q$ and $\mathbf{X}_d$ respectively. $\mathbf{X}_d$ and $\mathbf{X}_q$ give a constraint of the following form:

$$\mathbf{X}_q = \mathbf{R}_E \mathbf{X}_d + \mathbf{t}_E \tag{8}$$

Then, we can obtain the translation vector between the database camera center and the query camera center by Equation (9):

$$\mathbf{X}_q = \mathbf{R}_E (\mathbf{X}_d + \mathbf{R}_E^{-1} \mathbf{t}_E) \tag{9}$$

where $\mathbf{R}_E^{-1}\mathbf{t}_E$ is the translation vector in the database camera coordinate system. If we project database camera center and query camera center to the ground plane, $\mathbf{R}_E^{-1}\mathbf{t}_E$ is the slope factor of the connection line between the projection point of the database camera center and the projection point of the query camera center on the ground plane [8].

According to the Equation (1), we can transfer $\mathbf{R}_E^{-1}\mathbf{t}_E$ to the world coordinate system. The relation between the database camera coordinate system and the world coordinate system can be described as:

$$\mathbf{T} = \mathbf{R}_w \left(\mathbf{R}_E^{-1} \mathbf{t}_E \right) + \mathbf{t}_w \tag{10}$$

We assume $X_1O_1Y_1$, $X_2O_2Y_2$, and $X_3O_3Y_3$ planes (in camera coordinates) are parallel to the $X_WO_WY_W$ plane in the world coordinate, as shown in Figure 5. In the 2D situation (ground plane), we consider $\mathbf{R}_w \left(\mathbf{R}_E^{-1} \mathbf{t}_E \right)$ as the slope factor of the connection line (such as line l_{13} or l_{23} in Figure 5) which connects the projection points of the database camera center and the query camera center [8]. Since we know the locations of database cameras (such as location point O_1' and O_2' in Figure 5, we can determine the connection lines with the slope factor and the database camera locations.

Each pair of the database image and the query image can determine a connection line (we call it the projection connection line). Consequently, the projection point of the query camera center should be on every projection connection line. Therefore, if we obtain two projection connection lines, the projection point of the query camera center should be determined in the intersection of two projection connection lines. By two database images and their locations, we can obtain two projection connection lines, such as l_{13} and l_{23} in Figure 5. The l_{13} connects the projection points of database camera 1 center and the query camera center, and l_{23} connects the projection points of database camera 2 center and the query camera center. The intersection of l_{13} and l_{23} is the projection point of the query camera center. We consider this projection point of the query camera center is the location of the query camera on the ground plane.

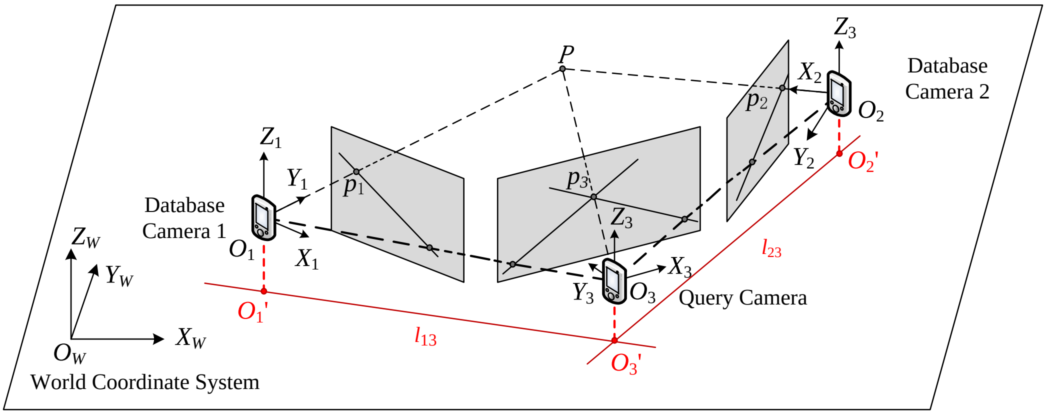


Figure 5. An Explanation of Location Recognition Principle

By the Coarse-to-Fine image retrieval, usually more than two database images are in the retrieval result, so there are more than two projection connection lines we can obtain. But these projection connection lines sometimes cannot intersect on one point due to the locating errors. So there may be more than one possible location of the query camera. In this paper, the mean value L_e of the possible locations is considered as the location of the query camera.

$$L_e = \frac{1}{N_s} \sum_{i=1}^{N_s} L_i \tag{11}$$

where L_e donates the location coordinates of the query camera, and L_i donates the possible location (intersection of any two projection connection lines) coordinates. N_s is the number of the projection connection line intersections.

5. Implementation and Performance Analysis

The potential of the proposed location recognition system is evaluated through comprehensive experimental tests conducted on a wide variety of dataset collected by the mobile data acquisition platform in the urban scene. In each scene, 30

Copyright ⓒ 2016 SERSC

database images and corresponding locations are captured. In order to test our location recognition system, query images and corresponding locations are also captured. In the process of capturing query images, the user walks in the region including database camera location marks, and the true locations of the query camera are recorded by the GPS device to compare with the estimation locations. The estimation locations are determined by our visual location recognition system.

In the world coordinate system, we consider that $X_W - Y_W$ plane coincides with the ground plane, and the directions of X_W and Y_W are same as the directions of the longitude and the latitude, respectively. The location coordinate values on $X_W - Y_W$ plane are the longitude and latitude values. In our experiment, the location of the query image is represented by the longitude and latitude values. To better reflect the location errors of our system, we also transform the longitude and latitude values to the distance values to present the location errors on the ground plane.

Figure 6, shows the location result of the HIT Main Building scene. In this scene, 30 locations (blue dots) of the database camera are captured uniformly and 18 locations (green squares) of the query camera are selected. Using the database images and the corresponding locations, we can determine the estimation locations (red rhombuses) of the query camera by our visual location recognition system. In this experiment, the estimation locations are presented by the longitude and latitude values. Comparing with the true locations and the estimation locations, we can clearly find out the performance of our visual location recognition system.

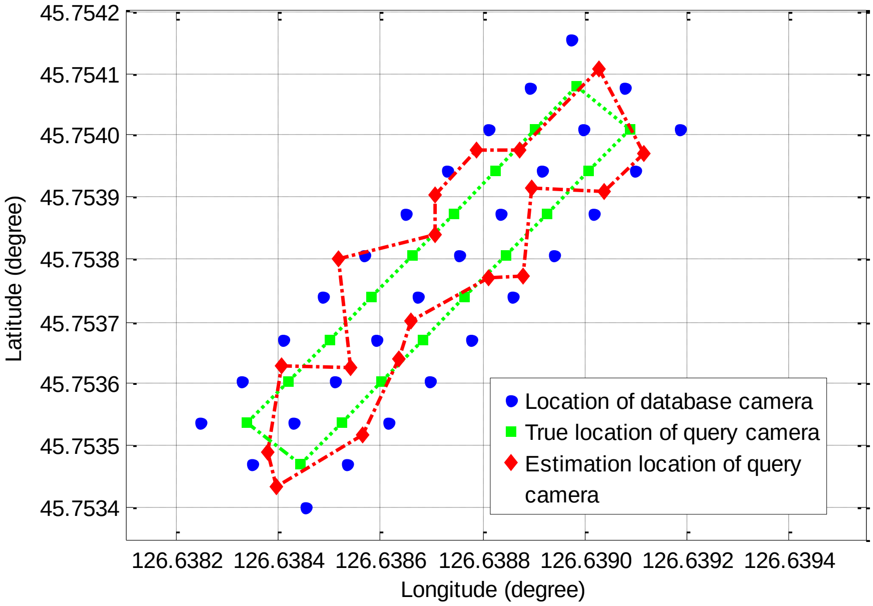


Figure 6. The Result of Coarse-to-Fine Image Retrieval

In order to show the location errors clearly, we transform the longitude and latitude values to the distance values. Figure 7, shows the location errors for the query images in different scenes. For 90 query images, the mean location errors in direction of the longitude and latitude are 3.2607 m and 3.2793 m, respectively. The Euclidean distances between the true locations and the estimate locations of the query images are also calculated. The mean location error for the Euclidean distance is 4.8194 m for 90 query images.

Copyright ⓒ 2016 SERSC

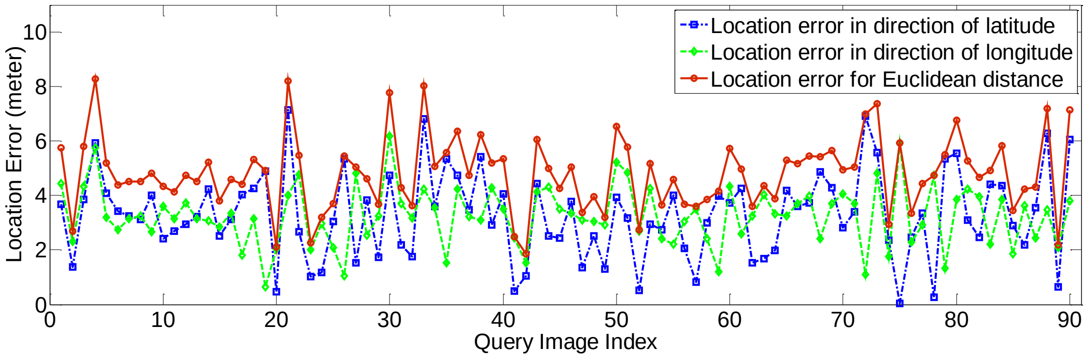


Figure 7. Location Errors for Query Images

The cumulative probability curves of the location errors are shown in Figure 8. The cumulative probabilities of our visual location recognition system within 3 m errors in the direction of the longitude and latitude reach 35.56% and 43.33%, respectively. The cumulative probabilities within 6 m errors in the direction of the longitude and latitude reach 98.89% and 94.44%, respectively. For Euclidean distance location errors, the cumulative probabilities within 4 m and 8 m are 26.67% and 96.67, respectively. These results demonstrate that our visual location recognition system can achieve high location accuracy in urban environment.

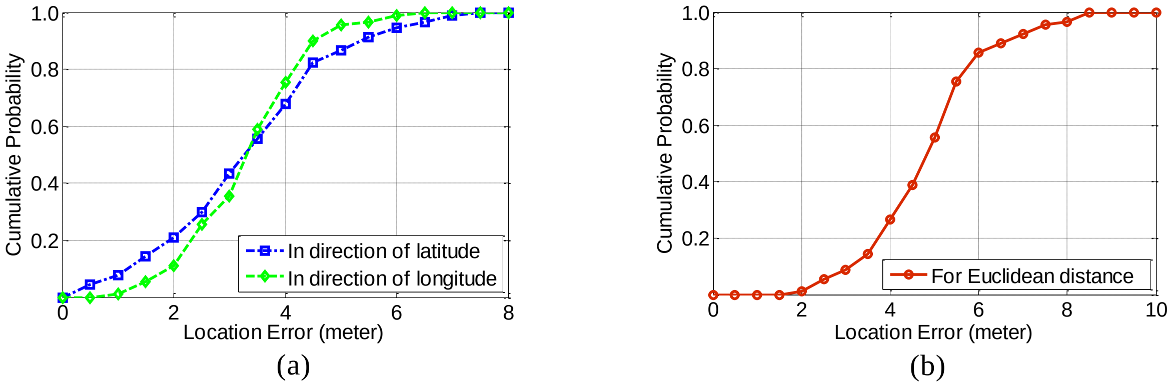


Figure 8. Cumulative Probabilities for Location Errors. (a) Cumulative Probabilities for Location Errors in Directions of Latitude and Longitude; (b) Cumulative Probabilities for Location Errors for Euclidean Distance

6. Conclusion

In this paper, a visual location recognition system based on the Coarse-to-Fine image retrieval and the epipolar geometry constraint is proposed. First, we present a database generation process which contains two aspects, the database images capturing and the corresponding location collection. Secondly, a Coarse-to-Fine image retrieval method is introduced. By image retrieval, the database images which are familiar to the query image are selected. Third, a visual location recognition algorithm is utilized. Based on the geometry constraint, the query camera location can be estimated by the database images and the corresponding positions. Different from existing researches focusing on the image retrieval for visual location recognition, the epipolar geometry constraint is used to refine the location of the query camera in the urban environment. The performance analysis and experiments in different scenes show the promising results for the location accuracy. The average error for the Euclidean distance between the true locations and the estimation locations of the query images is 4.8194 m, which satisfies the most requirements of the location based service. Compared with other localization methods, our method needs less localization infrastructures, but achieves satisfactory location accuracy. Future work is needed to extend the pose estimation algorithms which can precisely determine the user's orientation. Combing location recognition and pose estimation algorithms, an integrated state of user's location and orientation can be obtained.

Acknowledgments

This paper is supported by (2014ZX03004003-005), Municipal Exceptional Academic Leaders Foundation (2014RFXXJ002).

References

[1] H. Temeltas and D. Kavak, “SLAM for Robot Navigation”, IEEE Aerospace and Electronic Systems Magazine, vol. 23, no. 12, **(2008)**, pp. 16-19.

[2] S. Hwang and J. Song, “Monocular Vision-Based SLAM in Indoor Environment Using Corner, Lamp, and Door Features from Upward-Looking Camera”, IEEE Transactions on Industrial Electronics, vol. 58, no. 10, **(2011)**, pp. 4804-4812.

[3] R. Munguia and A. Grau, “Closing Loops with a Virtual Sensor Based on Monocular SLAM”, IEEE Transactions on Instrumentation and Measurement, vol. 58, no. 8, **(2009)**, pp. 2377-2384.

[4] G. Schindler, M. Brown and R. Szeliski, “City-Scale Location Recognition”, IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, America, **(2007)** June 17-22.

[5] I. H. Jhuo, T. Chen and D. T. Lee, “Scene Location Guide by Image-Based Retrieval”, Advances in Multimedia Modeling, vol. 5916, **(2010)**, pp. 196-206.

[6] G. Schroth, A. A. Nuaimi, R. Huitl, F. Schweiger and E. Steinbach, “Rapid Image Retrieval for Mobile Location Recognition”, IEEE International Conference on Acoustics, Speech and Signal Processing, Prague, Czech, **(2011)** May 22-27.

[7] J. Z. Liang, N. Corso, E. Turner and A. Zakhor, “Image Based Localization in Indoor Environments”, Fourth International Conference on Fourth International Conference on Computing for Geospatial Research and Application, San Jose, CA, America, **(2013)** July 22-24.

[8] H. Sadeghi, S. Valaee and S. Shirani, “A Weighted KNN Epipolar Geometry-Based Approach for Vision-Based Indoor Localization”, 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop, A Coruna, Spain, **(2014)** June 22-25.

[9] A. Oliva and A. Torralba, “Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope”, International Journal of Computer Vision, vol. 42, no. 3, **(2001)**, pp. 145-175.

[10] A. Oliva and A. Torralba, “An Improved Algorithm Finding Nearest Neighbor Using Kd-trees”, LATIN 2008: Theoretical Informatics Lecture Notes in Computer Science, vol. 4957, **(2008)**, pp. 387-398.

[11] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, “ORB: An Efficient Alternative to SIFT or SURF”, IEEE International Conference on Computer Vision, Barcelona, Spain, **(2011)** November 6-13.

[12] O. Chum and J. Matas, “Matching with PROSAC - Progressive Sample Consensus”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, America, **(2005)** June 20-25.

[13] R. Hartley and A. Zisserman, “Multiple View Geometry in Computer Vision (Second Edition)”, Cambridge University Press, Cambridge, **(2004)**.

[14] H. C. Longuet-Higgins, “A Computer Algorithm for Reconstructing a Scene from Two Projections”, Readings in Computer Vision: Issues, Problems, Principles, and Paradigms, **(1987)**, pp. 61-62.

[15] D. Nister, “An Efficient Solution to the Five-point Relative Pose Problem”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 6, **(2004)**, pp. 756-770.

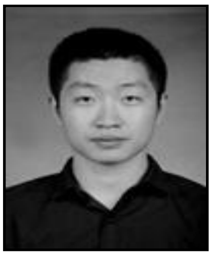
Authors

Guanyuan Feng, he received the B.S. degrees from Southwest Jiao Tong University, Chengdu, China, in 2011, received the M.S. degrees from Harbin Institute of Technology, Harbin, China, in 2013. He is currently working toward the Ph.D. degree in the School of Electronics and Information Engineering, Harbin Institute of Technology. His research interests include visual location recognition, visual navigation and cognitive radio networks.

Lin Ma, he received the B.S., M.S., and Ph.D. degrees from Harbin Institute of Technology, Harbin, China, in 2003, 2005 and 2009. He worked for Department of Electronics and Communication Engineering HIT as a doctoral tutor and associate professor. His research areas include the location and navigation technologies and cognitive radio technology.

Xuezhi Tan, he received the B.S., M.S., and Ph.D. degrees from Harbin Institute of Technology, Harbin, China, in 1982, 1986, and 2005, respectively. He is currently a Professor with the School of Electronics and Information Engineering, Harbin Institute of Technology. His research interests include wireless communications, digital trunking communication, and cognitive radio.

Hao Xue, he achieved the B.S. degree from Harbin Institute Technology, China, in 2014. He is currently pursuing the M.S. degree of information and communication engineering in the School of Electronics and Information Engineering, Harbin Institute of Technology. His researches are mainly on the aspect of vision-based indoor localization and navigation.



Kai Guan, he achieved the B.S. degree from Harbin Institute Technology, China, in 2014. He is currently devoting to the M.S. degree on information and communication engineering in the School of Electronics and Information Engineering, Harbin Institute of Technology. His researches are mainly on the aspect of image-based indoor localization and navigation.

Copyright ⓒ 2016 SERSC

Copyright ⓒ 2016 SERSC