Far East Journal of Electronics and Communications

Published Online: September 2016
http://dx.doi.org/10.17654/ECSV216131
Special Volume 2, 2016, Pages 131-138 **ISSN: 0973-7006**

FEATURE POINT ALGORITHM NECESSARY FOR COMPOSITION COGNITION OF GHOST IMAGE

Seongsoo Cho, Chung Hyeok Kim[*] and Woon Shik Choi

[*]Ingenium College of Liberal Arts
Kwangwoon University
20 Kwangwoon-ro, Nowon-gu
Seoul, 01897, South Korea

Department of Technology Education
Sehan University
1113 Noksek-ro, Samho-eub, Yeongam-gun
Cholla Namdo 58447, South Korea

Abstract

In this paper, we have attempted to classify necessary feature points for building cognition using adaptive splice critical more than once to improve the efficiency of building recognition technique on the level of cognition and matching. Due to the rapid change in urban environment, it is necessary to reflect the change in building by recognizing significant feature points of database. First of all, the feature points were extracted primarily through scale invariant feature transform (SIFT) and then eliminated falsely matched feature points. The random sample consensus (RANSAC) was applied to classify these points in the hidden area.

I. Introduction

Object recognition is used in a wide variety of applications in the field of computer vision. Especially, there are various programs which distinguish object inside image obtained by mobile phone cameras. The type of input

This paper initially presented at the ICACCI2016 held during May 27-29, 2016 at the University of San Carlos, Philippines.
Keywords and phrases: SIFT, RANSAC, occlusion region, feature point.
Lead Guest Editor: Gyanendra Prasad Joshi; Department of Information and Communication Engineering, Yeungnam University, South Korea.
[*]Corresponding author

image for building recognition is classified into aerial image and city building image. Most of the city building images coexist with pedestrians, tree, sky and other surrounding objects. Also, when excluding the buildings such as landmarks with particular shape, most of the buildings are in similar shape. Therefore, there is a need for a classification stage for specific building and surrounding environment, and the algorithm to process data with speed is required. Also, the studies are being conducted to find how to use GPS and other information to enhance accuracy [1]. The building recognition system is classified into feature point extraction and feature point matching. There are measures using global feature point such as scale invariant feature transform (SIFT) [2-3], and content based image searching measure using geometric specification, color, texture and shape of building in feature point extraction method [4-5]. When using global feature point to matching, 128 dimension or 62 dimension descriptor is in general use.

However, when the dimension is using high information, the process time prolongs. Methods such as principal component analysis (PCA) [6], linear discriminant analysis (LDA) [7], linear preserving projections (LPP) [8], supervised LPP (SLPP) [9], semi-supervised discriminant analysis (SDA) [10] propose measures to lower the dimension. Most of the building recognition technique extracts feature point from single image, and save the feature points after converging them as high-dimensional information like descriptor. Especially, in case of SIFT [2], the extracted feature point not only include the building information, but also the insignificant background information, the accuracy of feature point matching decreases.

This study utilized multi-frame images of the targeted building obtained in different times. The feature point based on SIFT [2] are classified and proposes a method to increase the accuracy in matching feature point. The proposed method utilizes the multi-view building input video rather than the existing method of using the singular period input video. While the feature point of buildings is reflected various input images, the feature point of background is only shown in limited image. Therefore, by using random sample consensus (RANSAC) to find relevance of feature point and obtain the repetitive feature point, and classify them to increase the accuracy of matching [11].

II. 9 Acquired Images Synchronization using NTP

Because the method utilizes multiple cameras, each camera must start recording at the same time. When recording the changing objects and the synchronization of camera do not accord, they might obtain different image.

Therefore, to achieve better accuracy of synchronization, the synchronization recording algorithm described in Figure 1 is used.

This study used 9 raspberry pies, and installed 1 camera module to each raspberry pie. The 9 camera modules are used to recording and each connected raspberry pie used Network Time Protocol (NTP) to estimate the recorded time. When the recorded time goes over certain threshold value, it is judged as not synchronized and obtains a new image.

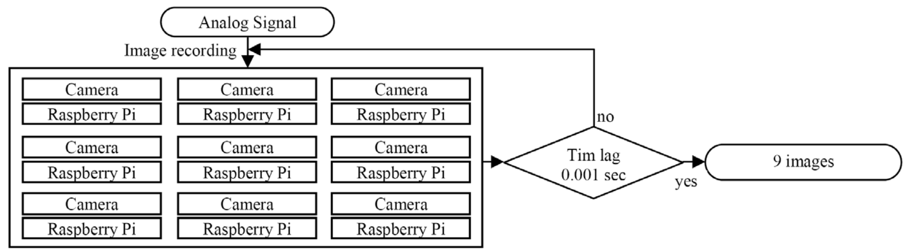

Figure 1. Synchronization recording algorithm.

III. Brightness and Smoothing

As shown in Figure 2, the brightness differs in accordance with the angle of camera, so the images with totally different brightness are obtained. Therefore, to increase the definition and image fusion accuracy, the brightness standardization is conducted. When using different method for different images, the better definition can be obtained [12]. We applied various methods including brightness standardization setting the brightness value of central camera among 9 cameras, average/central value of images, brightness value by white balancing as reference, and verified that histogram smoothing showed comparatively better performance.

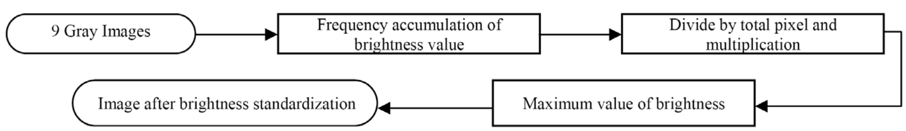

Figure 2. Histogram smoothing algorithm.

The brightness frequencies of 9 images are accumulated like in Figure 3 using equation (1). It shows the brightness of image is concentrated between 75~150.

$$sum[i] = \sum_{j=0}^{i} hist[j]. \qquad (1)$$

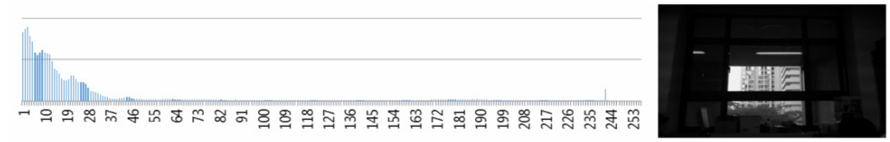

Figure 3. Brightness frequency accumulation.

The accumulated frequency obtained in equation (1), is divided and equation (2) which multiples maximum brightness is applied to show comparatively well distributed brightness as in Figure 4. It was clearer than the original image. They are applied equally to all 9 images to have image with equivalent brightness in the stage of fusion.

$$n[i] = sum[i] \times \frac{1}{N} \times 255. \tag{2}$$

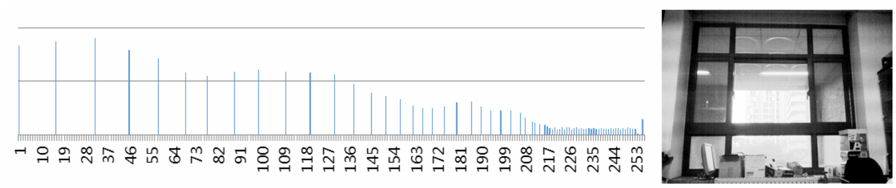

Figure 4. Histogram smoothing result.

IV. Image Fusion using Panorama Technique

As depicted in Figure 5, the panorama technique is divided into 5 stages. Fixing 9 cameras to set ROI for detecting feature point, and detect points that are the feature of images. Calculate the feature amount of the images, transform the images using values, to fuse the transformed image to obtain fused image ultimately. In the algorithm extracting feature point, there are various feature point extract algorithms such as SIFT, Harris corner, Features from accelerated segment test (FAST) [13], Speeded up robust features (SURF) [14], Binary robust independent elementary features (BRIEF) [15], Oriented FAST and rotated BRIEF (ORB) [16]. Fast retina keypoint (FREAK) [17] has the best speed and uses the strong feature in distortion of intensity difference images. When extracting SURF feature point, the Hessian matrix is used and the function of Hessian shows second differential.

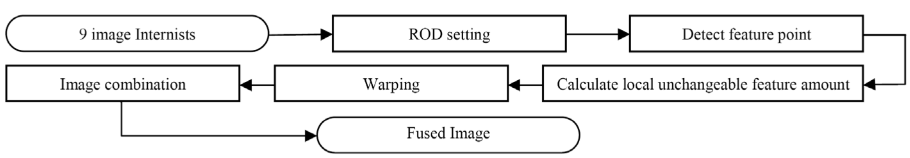

Figure 5. Image fusion algorithm.

Hessian is a matrix showing curve rate feature of the function. Hessian matrix detects feature points in maximum matrix formula. When matrix is positive and has same eigenvalue, it is judged by point of interest. Centering around the found points of interest, the Haar wavelet response surrounding are calculated. As depicted in Figure 6, the corresponding points of images are used to conduct geometric process. For the fusion, the uneven lengthening process of image is conducted which is called warping. It is to transform the image and fuse to connect images naturally [2].

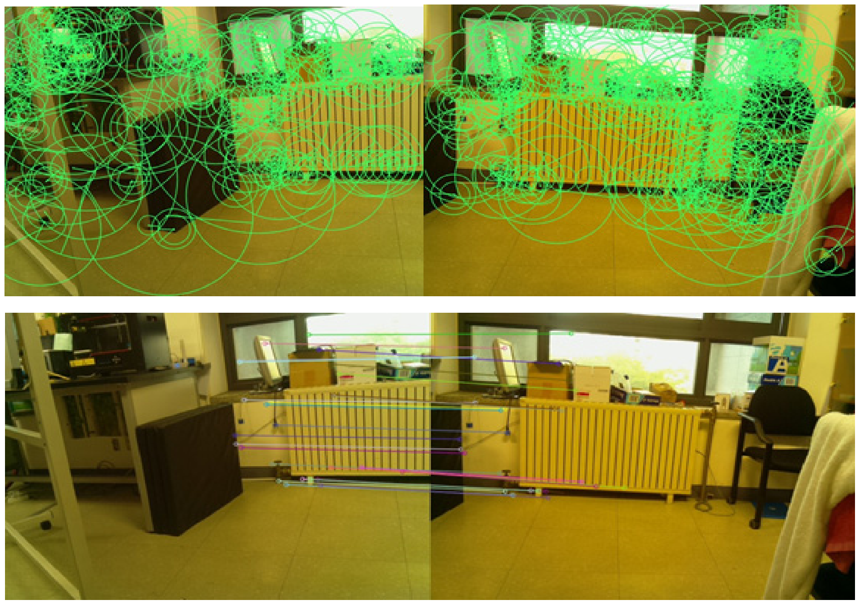

Figure 6. Connect feature points equivalent to local unchangeable feature amount.

(a)

(b)

Figure 7. Image set and fused panorama image recorded by raspberry camera (a), (b).

9 images are recorded at the same time like (a), (b) of Figure 7 and is able to obtain high-definition image with one image. At the same point, it has wider view of angle than raspberry camera and smart phone camera. It goes through preconditioning called brightness standardization, so it can conduct much accurate image fusion different from other panorama technique. Other than that, by adjusting the brightness, the image gets better definition. When fusing 9 images, the image is distorted. As shown in image (b) of Figure 7, the left and right side is a straight line originally, but is curved. This shows that there is a need for improvements in image obtaining process. When using the low price camera, the high-definition image with lower price was obtained. Also, in case of high-definition cameras, the super high definition images were able to be obtained by numerous studies to increase accuracy.

Table 1. Conudcted time for image set 1~10

Image set	Conduct time for open CV	Conduct time for proposed method
1	86.12 sec	45.87 sec
2	92.71 sec	49.54 sec
3	71.35 sec	46.78 sec
4	83.09 sec	50.41 sec
5	81.45 sec	30.98 sec
6	79.52 sec	41.22 sec
7	82.32 sec	71.08 sec
8	81.17 sec	64.92 sec
9	105.36 sec	101.11 sec
10	83.35 sec	44.30 sec
Average	84.64 sec	54.62 sec

Table 2. Average fusion rate for various image set

	Average fusion rate
Adobe Photoshop CS6	Approximately 70%
openCV Stitching	Approximately 82.5%
Proposed method	Approximately 90%

This study verified that the conduct time has remarkably reduced as in Table 1 when comparing to provided function in openCV [18]. This is considered to be due to the influence of setting ROI. Also, in the level of Registration, by using standardized image, the much accurate feature point detection is available. However, in a certain image set, there was no time difference, in which case it is difficult to detect the feature point. When the brightness of images is high, there are some cases when the lack of a feature point makes it hard to fuse. When using panorama technique of photoshop, it is not fused with very small difference. On the contrary, the proposed method as in Table 2 used histogram standardization to clarify the brightness difference to increase accuracy and lead to better rate of fusion.

V. Conclusion

This study classified the feature points using multi-frame in recognition stage and proposed a method to improve the matching performance. The existing openCV based building feature point extraction method frequently dealt with lower accuracy of matching due to the extraction of feature points of surrounding environments other than the recognition target. The proposed method used images recorded at the regular gap of time to obtain homography matrix, applied RANSAC method to classify the feature point. One homography matrix classifies the one flat feature points of image. By repeating the process of obtaining homography matrix with rest of the feature points after classification, the accurately matched feature points are all classified. The value of proposed method has increased compared to the value using openCV.

Acknowledgment

This work is supported by the research grant of Kwangwoon University in 2016.

References

[1] J. Li, W. Huang, L. Shao and N. Allinson, Building recognition in urban environments: a survey of state-of-the-art and future challenges, Information Sciences 277 (2014), 406-420.

[2] D. G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60(2) (2004), 91-110.

[3] H. P. Moravec, Obstacle avoidance and navigation in the real world by a seeing robot rover, Ph. D. Thesis, Department of Computer Science, Stanford University, 1980.

[4] Y. Li and L. G. Shapiro, Consistent line clusters for building recognition in CBIR, Proceedings of 16th International Conference on Pattern Recognition 3 (2002), 952-956.

[5] S. Cho, B. Shrestha and J. Lee, Method of recognising facial image using fuzzy measure, International Journal of Security and Networks 9(4) (2014), 185-189.

[6] I. Jolliffe, Principal Component Analysis, John Wiley & Sons, Ltd., 2002.

[7] G. McLachlan, Discriminant Analysis and Statistical Pattern Recognition, vol. 544, John Wiley & Sons, 2004.

[8] X. Niyogi, Locality preserving projections, Neural Information Processing Systems, vol. 16, MIT, 2004.

[9] D. Cai, X. He and J. Han, Using Graph Model for Face Analysis, Department of Computer Science, University of Illinois at Urbana Champaign, 2005.

[10] D. Cai, X. He and J. Han, Semi-supervised discriminant analysis, IEEE 11th International Conference on Computer Vision, 2007, pp. 1-7.

[11] M. A. Fischler and R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Communications of the ACM 24(6) (1981), 381-395.

[12] K. Mikolajczyk and C. Schmid, Indexing based on scale invariant interest points, Proceedings of the Eighth IEEE International Conference on Computer Vision, 2001, pp. 525-531.

[13] E. Rosten and T. Drummond, Machine learning for high-speed corner detection, European Conference on Computer Vision, Springer, Berlin, Heidelberg, 2006, pp. 430-443.

[14] H. Bay, T. Tuytelaars and L. Van Gool, Surf: Speeded up robust features, European Conference on Computer Vision, Springer, Berlin, Heidelberg, 2006, pp. 404-417.

[15] M. Calonder, V. Lepetit, C. Strecha and P. Fua, Brief: Binary robust independent elementary features, European Conference on Computer Vision, Springer, Berlin, Heidelberg, 2010, pp. 778-792.

[16] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, ORB: An efficient alternative to SIFT or SURF, IEEE 2011 International Conference on Computer Vision, 2011, pp. 2564-2571.

[17] A. Alahi, R. Ortiz and P. Vandergheynst, Freak: Fast retina keypoint, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012, pp. 510-517.

[18] K. Pulli, A. Baksheev, K. Kornyakov and V. Eruhimov, Real-time computer vision with OpenCV, Communications of the ACM 55(6) (2012), 61-69.