A NOVEL FILTERING APPROACH FOR ROBUST AND FAST KEYPOINT MATCHING IN MOBILE ENVIRONMENT

Jonghoon Seo

LG Electronics
Software Platform R&D Lab.
19 Yangjae-daero 11 gil, Seoul, Korea

*Seungho Chae, Yoonsik Yang, Tack-Don Han**

Yonsei University
Department of Computer Science
50 Yonsei-ro, Seoul, Korea

ABSTRACT

Local keypoint matching method is widely used because it can provide robust to environment change and occlusion. In this paper, we propose a novel keypoint filtering approach, which accomplish not only fast, but also robust and reliable matching. The proposed approach, in offline learning phase, evaluates detected keypoints with proposed criteria and stores only high-distinguishable keypoints. This approach reduces the number of stored keypoints, also reduces false matching rate and attains fast and robust matching quality. We evaluated this approach in mobile augmented reality application, and proved the proposed approach is effective in terms of both matching speed and quality.

Index Terms— Keypoint filtering, Keypoint matching, Feature selection, Keypoint saliency, Image matching

1. INTRODUCTION

Image matching is a fundamental problem in a variety of computer vision applications, including object recognition[1], panorama stitching[2], and augmented reality[3]. To accomplish image matching, keypoint-based local matching is widely used, since it can provide relatively high matching quality against severe occlusion and do not require segmentation for regions of interest. Also, recent work has concentrated on making invariant to image transformation with low computing power[4, 5].

The overall flowchart of keypoint matching and recognition is shown in Fig. 1. This procedure can be divided into two main phases: offline learning (or training) and online testing procedure. Offline learning is a prerequisite to online matching process. In offline learning phase, a set of reference images to be recognized is analyzed and stored a set of descriptors in a database. In online testing phase, a newly captured image is analyzed and compared with the reference images in the database to find a nearest reference image.

Conventional keypoint matching methods stores almost every keypoints which are detected by keypoint detection process. Keypoint detection processes are designed to extract repeatable keypoints and robust against arbitrary image transformation. Then, detected keypoints are independent to the follow matching procedures, and do not reflect quality of descriptors. Therefore, as seen Fig. 2, some keypoints are not distinguishable, and they tend to cause inter-keypoint confusion and miss matching. Also, those detected keypoints are stored in database and are compared with keypoints in query images in every frame while matching. Then, it decreases the speed of matching. To overcome these problem, in offline learning procedure, detected keypoints are evaluated with respect to proposed matching quality criteria and filtered by the goodness score. With this filtering method, only a small subset of keypoints is stored in the database. Accordingly, it provides more improved matching performance with faster matching speed.

Especially, mobile devices still have insufficient computing power and limited memory compared to desktop, so there is an urgent need of effective processing methods for image matching. However, conventional keypoint matching approaches stored redundant keypoints into database, and these

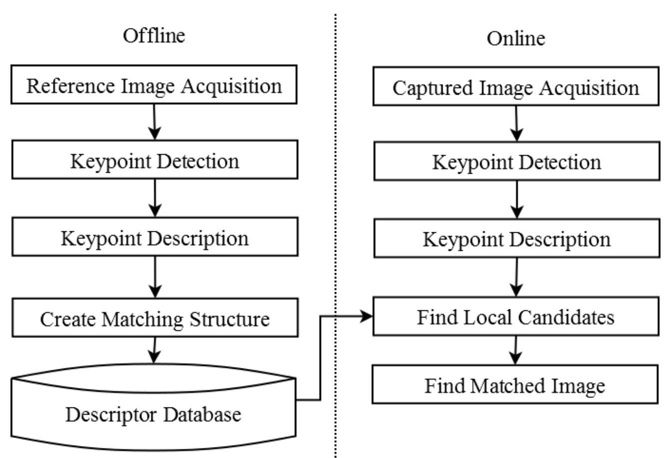


Fig. 1. Overall process of conventional keypoint-based matching

*This

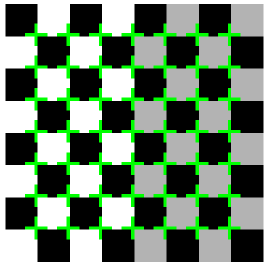


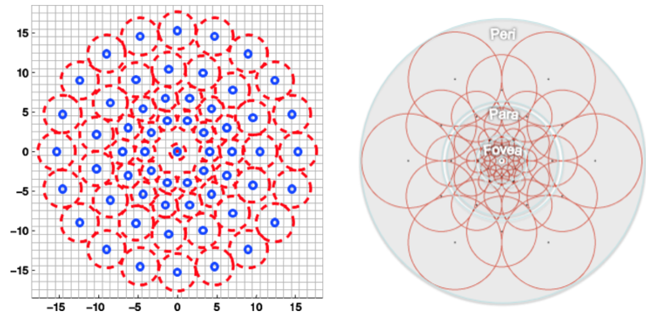
Fig. 2. Example of high repeatable but poor distinguishable keypoints. Conventional keypoint matching systems do not consider the discriminability of keypoints, so these keypoints usually stored and negatively affected matching.

redundant keypoints may are compared in every frame. So, matching speed will be decreased and this causes problem in the mobile computing devices. On the other hand, the proposed method removes redundant keypoints in the database, it reduces the number of comparisons while matching and increase matching speed even in the mobile computing environment.

This paper is structured as follows: In Chapter 2, we discuss literature on light-weight keypoint-matching algorithms. Chapter 3 describes the proposed keypoint score function. In Chapter 4, we executed experiments to prove the proposed keypoint filtering method in various algorithms and compared over several evaluation metrics. Finally, Chapter 5 presents the conclusion.

2. RELATED WORKS

In spite of robust matching quality, keypoint-based local matching requires a large amount of computation, especially in mobile computing environment. To overcome this limitation, various lightweight algorithms are proposed. At first, in the study of keypoint descriptors, the vector-value based description methods such as SIFT[6] or SURF[7] provide robust recognition quality, but complex computation is required to calculate descriptors. In recent years, variant binary-value based descriptors, such as BRIEF[8], ORB[9], BRISK[10], FREAK[11], are proposed. These binary descriptors compare the brightness values with focus on the keypoints as shown in Figure 3 by using a wide range of form patterns and express the results in binary codes. Since descriptors are computed only by a simple comparison computation, its computational speed is significantly faster than vector-based descriptors. Moreover, since the orientation and scale are normalized on the basis of generation patterns, they show significantly robust performance against a wide range of distortion. In particular, studies using the binary descriptors that can be processed by simple comparison computation rather than complex vector-value descriptor based computation are increasing in the environment of limited performance such as a mobile computing.



(a) BRISK[10] Sampling Pattern (b) FREAK[11] Sampling Pattern

Fig. 3. Sampling Patterns of Binary Descriptors

The other researches accelerates nearest neighbor matching with designing an efficient matching data structure. The conventional brute force matching compares all the keypoints of query images with all the keypoints of reference images, so it shows the slowest, but has the advantage of being able to detect the most accurate nearest neighbors. kD tree based approximation method is proposed in [12]. This method shows good performance in the vector-value descriptor method such as SIFT or SURF with relatively low dimension of features but the improvement of its performance is unlikely to be achieved in the latest binary descriptor with high dimension. A method enabling it to accelerate matching using Hashing based structure in LSH[13] was proposed[9]. As for this method, it is critical to compose appropriate hash function set to distribute the keypoints points evenly in the offline training phase. In addition, 'Random Forest[14]' or 'Random Fern[15]' composed a matching structure by applying a binary description method to the tree structure or list structure. To increase the recognition speed and obtain more accurate approximation values, these matching methods generate efficient matching structure by way of the application of supplementary computation using the descriptors computed in the offline training phase. However, since the supplementary structure that uses this method is complex and large, its the matching structure is too heavy to use in a mobile environment. Moreover, since the properties of the detected keypoints are not taken into consideration, the set for which it is difficult to classify the detected keypoints may lead to performance degradation. To solve this problem, this dissertation proposes a keypoints filtering based matching methods.

3. PROPOSED METHOD

The proposed filtering method stores only good keypoints by analyzing the characteristics of detected keypoints and measuring the degree of effectiveness for image matching.

There are several factors which good keypoints should follows:

Repeatability: Good keypoints need to be stably detected even in various environment. In actual matching environments, a wide range of transformed image degradation may occurred, such as the rotation, size, noise and lighting of the targeted images. The good keypoints have to be stably extracted against those transformation.

The detection of stable keypoints can be measured by *Repeatability* condition. *Repeatability* is calculated by the ratio between the total number of synthesized images and the number of cases where the transformed keypoints are existent in the synthesized images.

$$p_{rep}(p_i) = \frac{n_i^{overlap}}{N} \tag{1}$$

where $p_{rep}(p_i)$ represents repeatability score of given point p_i, $n_i^{overlap}$ is calculated by the frequency of the existence of transformed keypoint(p_i) in the set of keypoints($T(p_i) \in K_t'$) of synthesized images $T_t(I)$; N is the total number of synthesized images ; and all keypoints have single value.

Similarity: Good keypoints need to be well-matched with identical keypoints even though targeted images change in various ways. With regard to a certain keypoint(p_i) of reference images, genuine distribution' and imposter distribution' for the corresponding keypoint can be measured by calculating the matching between the descriptors of all the sets of keypoints(p_i) in images($T_t(I)$) transformed in various ways during the training process. At this time, to reduce the failure in matching the corresponding keypoints and the descriptors in the transformed images, the genuine distribution needs to have small value, being far enough away from match distance threshold. To this effect, it was measured using the mean of genuine distribution. As shown in Equation (2), the keypoints with the decreasing the genuine distribution are better, so the evaluation function was calculated by normalizing the mean of the genuine distribution and subtracting its value from 1.

$$p_{sim}(p_i) = 1 - \frac{\mu_{gen,i} - \min_i \mu_{gen,i}}{\max_i \mu_{gen,i} - \min_i \mu_{gen,i}} \tag{2}$$

where $p_{sim}(p_i)$ represents similarity score of given point p_i, $\mu_{gen,i}$ represents mean of genuine distribution of give point p_i.

Separability: The trained keypoints and other keypoints shall not be matched, which is associated with the imposter distribution of each keypoint. Of the keypoints extracted from the images converted in various images, the distribution of the matching with other keypoints rather than the converted keypoints themselves are referred to as imposter distribution. Thus for a specific keypoint to show the low success rate of matching with other keypoints rather than themselves, in is necessary that the genuine distribution and imposter distribution are well classified. To this effect, in this paper, *Fisher's Discriminant Ratio[16]* was used. It measures the distance between two classes by the mean and distribution of sample in 1-dimensional, two class problems. Since the second Similarity condition ensures the genuine distribution is small enough, the nonexistence of the matching with the keypoints in the imposter distribution is ensured if the importer distribution is far enough away compared with the genuine distribution. Separability value also requires the normalization process as shown in equation (4).

$$FDR(p_i) = \frac{(\mu_{gen,i} - \mu_{imp,i})^2}{\sigma_{gen,i}^2 + \sigma_{imp,i}^2} \tag{3}$$

$$p_{sep}(p_i) = \frac{FDR(p_i) - \min_i FDR(p_i)}{\max_i s_i} \tag{4}$$

where $p_{sep}(p_i)$ represents similarity score of given point p_i, $\mu_{gen,i}, \mu_{imp,i}$ represent mean of genuine and impostor distribution of give point p_i, respectively. $\sigma_{gen,i}, \sigma_{imp,i}$ represent standard deviation of genuine and impostor distribution, respectively.

The score functions of each keypoint can be defined using 3 criteria calculated as above. The 3 conditions are dependent, so can be defined as shown in Equation (5).

$$gf(p_i) = p_{rep}(p_i)p_{sim}(p_i)p_{sep}(p_i) \tag{5}$$

4. EXPERIMENTS

In this chapter, we present several experiments that demonstrate the effectiveness of the proposed method. As for the experimental images, 16 images of Seoul Guide Map Pamphlet was selected. These images were deformed by way of rotating($0.5 \sim 2.0$-folds, at the interval of 0.1-fold) scaling (($0^\circ \sim 360^\circ$, at the interval of 10° intervals) and blurring (Gaussian blur, $r = 0, 3, 5, 7$ pixels). As a result, 32,256 images were obtained. Of them, 16,114 training images and 16,142 test images were selected at random. With training images, we detected keypoints and calculated the score function($gf(p_i)$). Then, we stored filtered keypoints and calculated nearest neighbor with test images.

At first, to prove the proposed algorithm can enhance recognition performance, we conducted a comparison experiment(see, Fig. 4). Keypoints are detected with AGAST[17] detector and described with FREAK[11] descriptor. The figure shows comparison among top 50 keypoints based on the proposed score function(*Proposed*), 50 keypoints filtered randomly(*Random*), top 50 keypoints based on the corner response function(*C.R.F.*), and non-filtered $3,000$ keypoints set(*Full*). The figure shows that the proposed approach generates the most robust matching performance.

At second, to prove the proposed approach works among several algorithm suits, we performed the comparison tests with changing detection and description algorithms(see, Fig.

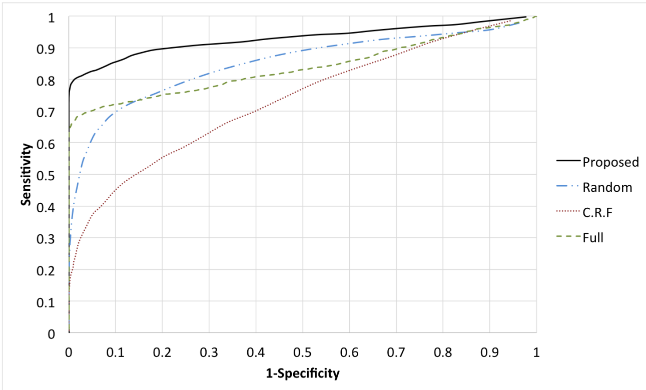
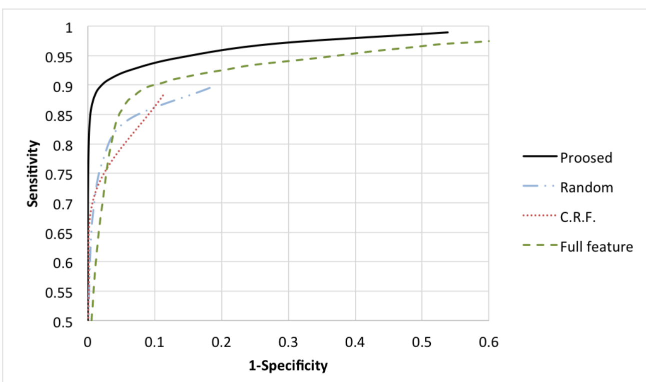
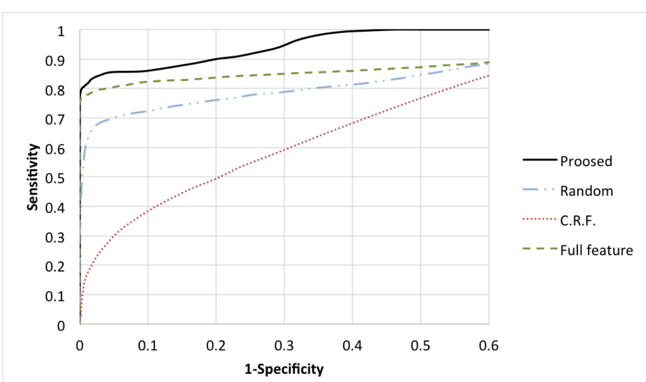


Fig. 4. ROC curve comparison between proposed filtering function and other methods

??). These figures show the proposed method generates robust matching performance even though the algorithm is changed.



(a) AGAST detector and BRISK descriptor



(b) SURF detector and FREAK descriptor

Fig. 5. ROC curve comparison between different algorithm sutirs

The last experiment is performed to calculate the optimal number of keypoints. The experiment is performed among non-filtered 3,000 keypoints set(K_{all}), top 500, 300, 100, 50 keypoints subsets($K_{500,300,100,50}$). As shown Fig. 6, the result showed slight degradation of recognition rate whereas K_{100} and K_{50} showed the improvement of performance.

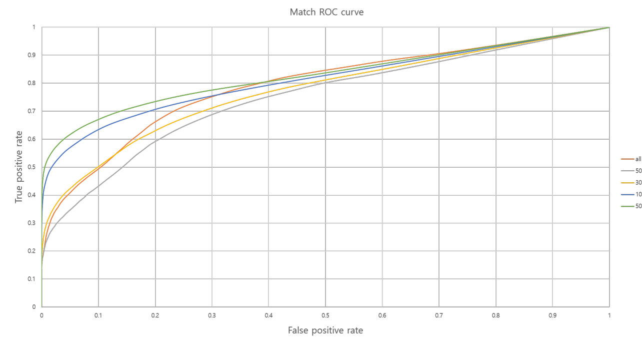


Fig. 6. ROC curve for match rate

5. CONCLUSION

In this paper, we propose a keypoint filtering approach to accomplish not only fast, but also robust and more reliable matching for mobile computing environment. Conventional keypoint matching methods rarely consider about quality of stored keypoint database. So, to accomplish robust matching quality, they require more keypoints. However, those redundant keypoints degrade matching quality. To overcome this problem, we evaluate detected keypoints and store only filtered keypoints. These filtered keypoints are repeatedly detected despite the change of images; show high match similarity with identical keypoints; and show low match similarity with other keypoints. Experimental results show that our filtering approach is effective in terms of both matching speed and quality thus applying this approach feasible for mobile image matching applications such as mobile object recognition or aumented reality.

6. REFERENCES

[1] D. Nister and H. Stewenius, “Scalable recognition with a vocabulary tree,” in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, vol. 2, pp. 2161–2168.

[2] M. Brown and D.G. Lowe, “Recognising panoramas,” in *Ninth IEEE International Conference on Computer Vision, 2003. Proceedings*, Oct. 2003, pp. 1218–1225 vol.2.

[3] Daniel Wagner, D. Schmalstieg, and H. Bischof, “Multiple target detection and tracking with guaranteed framerates on mobile phones,” in *8th IEEE International Symposium on Mixed and Augmented Reality, 2009. ISMAR 2009*, Oct. 2009, pp. 57–64.

[4] Gerardo Carrera, Jesus Savage, and Walterio Mayol-Cuevas, "Robust feature descriptors for efficient vision-based tracking," in *Progress in Pattern Recognition, Image Analysis and Applications*, Luis Rueda, Domingo Mery, and Josef Kittler, Eds., number 4756 in Lecture Notes in Computer Science, pp. 251–260. Springer Berlin Heidelberg, Jan. 2007.

[5] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[6] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[7] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, June 2008.

[8] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua, "BRIEF: Binary robust independent elementary features," in *Computer Vision – ECCV 2010*, Kostas Daniilidis, Petros Maragos, and Nikos Paragios, Eds., number 6314 in Lecture Notes in Computer Science, pp. 778–792. Springer Berlin Heidelberg, Jan. 2010.

[9] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *2011 IEEE International Conference on Computer Vision (ICCV)*, Nov. 2011, pp. 2564–2571.

[10] S. Leutenegger, M. Chli, and R.Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *2011 IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 2548–2555.

[11] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast retina keypoint," in *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2012, pp. 510–517.

[12] J.S. Beis and D.G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," in *, 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997. Proceedings*, June 1997, pp. 1000–1006.

[13] Aristides Gionis, Piotr Indyk, and Rajeev Motwani, "Similarity search in high dimensions via hashing," in *Proceedings of the 25th International Conference on Very Large Data Bases*, San Francisco, CA, USA, 1999, VLDB '99, pp. 518–529, Morgan Kaufmann Publishers Inc.

[14] V. Lepetit and P. Fua, "Keypoint recognition using randomized trees," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1465–1479, Sept. 2006.

[15] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast keypoint recognition using random ferns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 448–461, Mar. 2010.

[16] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, Sept. 1936.

[17] Elmar Mair, Gregory D. Hager, Darius Burschka, Michael Suppa, and Gerhard Hirzinger, "Adaptive and generic corner detection based on the accelerated segment test," in *Computer Vision – ECCV 2010*, Kostas Daniilidis, Petros Maragos, and Nikos Paragios, Eds., number 6312 in Lecture Notes in Computer Science, pp. 183–196. Springer Berlin Heidelberg, Jan. 2010.