

MEGH: A New Affine Invariant Descriptor

Xiaojie Dong¹, Erqi Liu², Jie Yang¹ and Qiang Wu³

¹ Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University
Shanghai, 200240 - China

[e-mail: xjdong_2010, jieyang@sjtu.edu.cn]

² China Aerospace Science & Industry Corp
Beijing, 100048 - China

[e-mail: everyth_ok@yahoo.com.cn]

³ University of Technology Sydney
Sydney, 2007 - Australia

[e-mail: qiang.wu@uts.edu.au]

*Corresponding author: Xiaojie Dong

*Received March 31, 2013; revised May 20, 2013; revised June 14, 2013; accepted June 19, 2013;
published July 30, 2013*

Abstract

An affine invariant descriptor is proposed, which is able to well represent the affine covariant regions. Estimating main orientation is still problematic in many existing method, such as SIFT (scale invariant feature transform) and SURF (speeded up robust features). Instead of aligning the estimated main orientation, in this paper ellipse orientation is directly used. According to ellipse orientation, affine covariant regions are firstly divided into 4 sub-regions with equal angles. Since affine covariant regions are divided from the ellipse orientation, the divided sub-regions are rotation invariant regardless the rotation, if any, of ellipse. Meanwhile, the affine covariant regions are normalized into a circular region. In the end, the gradients of pixels in the circular region are calculated and the partition-based descriptor is created by using the gradients. Compared with the existing descriptors including MROGH, SIFT, GLOH, PCA-SIFT and spin images, the proposed descriptor demonstrates superior performance according to extensive experiments.

Keywords: Local Descriptor, SIFT, Image Registration, Performance Evaluation

This research is partly supported by NSFC, China (No: 61273258, 61105001), Ph.D. Programs Foundation of Ministry of Education of China (No: 20120073110018). We would like to thank the anonymous reviewers for the valuable comments.

<http://dx.doi.org/10.3837/tiis.2013.07.010>

1. Introduction

Local features that are robust to photometric transformations and geometric transformations are crucial to many image understanding and computer vision applications [1]. Such local features basically consist of detecting interest features in an affine covariant manner and describing the characteristic structure of a local image region around the detected features.

In recent years, a number of feature detectors extracting interest regions have been investigated, such as MSER (maximally stable extremal region) detector [2], EBR (edge-based region) detector and IBR (intensity extrema-based region) detector [3], Harris-Affine and Hessian-Affine detectors [4] and salient region detector [5]. Comprehensive study on these existing detectors indicates that MSER and Hessian-Affine are the two best detectors [6].

After the regions of interest are detected, a descriptor is needed to describe the characteristic structure of the detected regions. Many techniques for describing local image regions have been developed. One of the popular descriptors is SIFT [7]. Inspired by the discriminability and robustness of the SIFT, lots of variants have been proposed, such as GLOH (gradient location and orientation histogram) [8], SURF [9], PCA-SIFT (principal component analysis-SIFT) [10], DAISY [11] and PPD (phase-space partition based descriptor) [12]. According to the study results of [8], SIFT and GLOH have been demonstrated to be superior to others used in literature on a number of measures.

A good local image descriptor is expected to have high discriminative ability and to be robust to various image transformations, such as scaling, rotation, viewpoint changes, image blur, JPEG compression and illumination changes.

To achieve rotational invariance, the widely adopted approach is to determine a reference orientation for each local region around its interest point, such as SIFT, SURF and DAISY. Although the gradient histogram representation is robust in somehow to deformations of the image pattern, estimation of the reference orientation, according to the properties of local area, is still problematic. It affects the performance of the local descriptor [13, 14]. More detailed study can be seen in [13].

To address the problem mentioned above, a new feature descriptor is proposed in this paper, MEGH (**M**ulti-support region **E**llipse-partition based **G**radient **H**istogram). In the proposed method, the affine covariant region is divided into 4 sub-regions according to ellipse orientation (the major axis orientation). Meanwhile, the affine covariant regions are normalized to the circular regions. Then, the gradients of pixels in circular region are computed in a locally rotation invariant coordinate system. Finally, the descriptor is created by gradient histogram based on the sub-regions. Compared with the existing descriptors including MROGH, SIFT, GLOH, PCA-SIFT and spin images, the proposed descriptor shows superior discriminability according to extensive experiments.

The main contribution of this paper is that a special image segmentation based on affine covariant region is proposed. An affine covariant region is divided into 4 sub-regions from ellipse orientation with equal angles. To achieve invariance to scaling, each affine covariant region is normalized to a canonic region with a common size. To further improve the discriminative ability of the proposed descriptor, the scheme of utilizing multi-support region as [13, 17] is also followed in this paper.

The rest of the paper is organized as follows: Section 2 discusses the related work. The proposed algorithm is presented in Section 3. Experiments and conclusions are given in

Section 4 and 5 respectively.

2. Related Work

There are two major approaches to construct rotation-invariant descriptors. One approach is to rotate the normalized region to align a reference orientation [7, 9, 10, 12, 15] and then the feature descriptor is built up. The other is to directly design rotation-invariant descriptor [13, 14, 16]. A brief review is given below to explore their advantages and disadvantages which inspire the newly proposed descriptor in this paper. Comprehensive reviews can be found in [8].

SIFT [7] is regarded as one of the most popular rotation invariant feature descriptor. SIFT description of an image is presented as 3D smoothed histogram of gradient locations and orientations. The height of bin represents the weighted sum of gradient magnitude in each area and in each gradient orientation. To encode more spatial information, a square image patch around an interest point is subdivided into 16 smaller squares. And then, the gradient orientation is quantized into 8 bins in each smaller square. For each smaller square, an orientation based histogram is formed. All these orientation histograms over all smaller squares are concatenated together to construct the SIFT descriptor. Therefore, the SIFT is a 128D feature vector. Although the gradient orientation histogram provides stability against deformations of the image pattern, the SIFT still requires an accurate dominant orientation as a reference orientation for local image rotation alignment. The dominant orientation is perhaps more than one for each interest point.

In order to further improve its efficiency and effectiveness, a number of extensions of SIFT have been developed. PCA-SIFT descriptor introduced by Ken and Sukthankar [10] is a simplified version of the SIFT, which further processes gradient orientation values in SIFT through PCA sub-space analysis. GLOH descriptor is another extension of the SIFT. Instead of computing the histogram in the square grid, GLOH is calculated along log-polar location grids and the dimension of descriptor is reduced by PCA. PPD descriptor [12] reduces the complexity of standard SIFT and improves its discriminability. PPD adopts region-wise gradient statistics in phase space and also saves interpolation normally used by common SIFT.

To achieve the rotational invariance, SIFT and its variants rely on the estimation of dominant orientation. However, as discussed in [13], the dominant orientation estimation tends to be unreliable and thus it affects the performance of the local descriptor.

Directly designing rotation-invariant descriptor is another category of methods. The intensity-domain spin image [14] is a 2D histogram of distances between the pixel in the normalized patch and the center point of the normalized patch and intensities. Since the distance and the intensity value are invariant to orthogonal transformations of image neighborhood, the intensity-domain spin image itself is rotation-invariant. RIFT [14] (rotation-invariant feature transform) is another descriptor that achieves rotation invariance without requiring a reference orientation. A circular patch is built at an interest point and then is subdivided into 4 rings of equal width. To maintain rotation invariance, the relative orientation between the gradient orientation and the direction which points outward from the center of the circular patch is computed at each pixel. For each ring, 8 orientation histogram is computed. Thus, RIFT is a 32-dimension feature vector.

A good method for constructing local feature descriptor [13] is proposed and two descriptors, MROGH (multi-support region order-based gradient histogram) and MRRID (multi-support region rotation and intensity monotonic invariant descriptor), are presented. Pooling local features based on intensity order is the key idea of the construction method

which is just to aggregate the relevant pixels together. The intensity order and the feature pooling scheme are both rotation invariant, thus no reference orientation is required in both MROGH and MRRID descriptors. More importantly, both MROGH and MRRID achieve better performance than state-of-the-art descriptors. In addition, to improve the discriminative ability of the descriptor, the scheme of multiple support regions is used in [13, 17].

3. Proposed Descriptor

The key idea of our method is to directly segment the affine covariant region according to ellipse orientation. According to ellipse orientation, the affine covariant region is firstly divided into 4 sub-regions with equal angles. Thus, it is unnecessary to explicitly estimate the dominant orientation as other method such as SIFT. Meanwhile, the affine covariant regions are normalized to a fixed circular region to obtain affine invariance. Then, the gradients of all pixels in the normalized region are computed. In the end, the descriptor is calculated by using the gradients according to the divided sub-regions.

3.1 Sub-region Construction

Given the canonical equation of an arbitrary ellipse with center X_C .

$$(X - X_C)^T A (X - X_C) = 1 \quad (1)$$

where

$$A = \begin{pmatrix} d & e \\ e & f \end{pmatrix} \quad (2)$$

A is a real symmetric matrix. Then, the orientation of sub-major axis θ of the ellipse can be determined as:

$$\theta = \frac{1}{2} \tan^{-1} \frac{2e}{d-f} \quad (3)$$

From ellipse orientation angle θ , an ellipse is segmented into N sub-regions, where N is

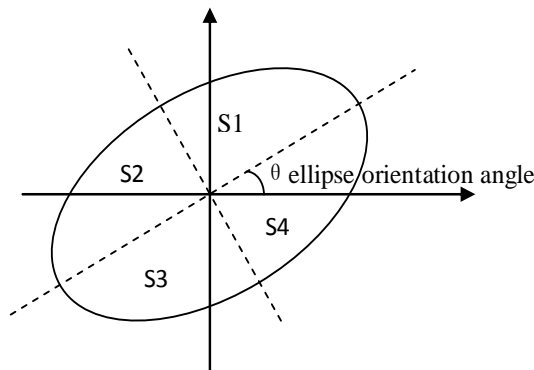


Fig. 1. An ellipse divided into 4 sub-regions

set to 4 in this paper. Each sub-region is one-fourth of the ellipse. Since the partition of ellipse is carried out from ellipse orientation, the size of each sub-region is always one fourth of original ellipse regardless the rotation, if any, of ellipse. An example of the divided sub-regions is shown in Fig. 1.

Given a pixel X denoted by $(x, y)^T$ in the ellipse, its polar angle β is:

$$\beta = \tan^{-1}(y / x) \quad (4)$$

Depending on the location of the pixel in affine covariant region, the relation between the polar angle of each pixel and the ellipse orientation is summarized in Table 1.

Table 1. Relation between the polar angle of each pixel and the ellipse orientation

Sub-region: S1	$\theta < \beta \leq \theta + \frac{\pi}{2}$	Sub-region: S3	$\theta + \pi < \beta \leq \theta + \frac{3\pi}{2}$
Sub-region: S2	$\theta + \frac{\pi}{2} < \beta \leq \theta + \pi$	Sub-region: S4	$\theta + 3\pi / 2 < \beta \leq 2\pi$ or $0 < \beta \leq \theta$

3.2 Affine Invariant Regions

The detected affine covariant region may have different size and different orientation by using various region detectors, such as the Hessian-Affine/Harris-affine detector. Similar to many other local descriptors, we also normalized the affine covariant regions to the circular regions in order to obtain affine invariance, which is a circular region of radius 20.5 pixels [8, 13]. If the detected region is larger than the normalized region, a Gaussian kernel should be used to smooth the image of the detected region. The standard derivation of Gaussian kernel is set to be the size ratio of the detected region and the normalized region [8].

Given the center of ellipse as the origin of coordinate, any pixel X in the affine covariant region satisfies:

$$X^T A X \leq 1 \quad (5)$$

Pixel X' in the normalized circular region satisfies:

$$X'^T X' \leq r^2 \quad (6)$$

where r is the radius of normalized circular region. From Eq. (5) and Eq. (6), we have:

$$X = \frac{1}{r} A^{-\frac{1}{2}} X' = T^{-1} X' \quad (7)$$

Therefore, the intensity value of each pixel X' in normalized circular region can be calculated based on its corresponding pixel X in affine covariant region. Eq. (7) cannot guarantee there is always an X in the discrete coordinate corresponding to X' . If that is the case, bilinear interpolation is applied to obtain the approximate location. Example of normalized circular region is shown in Fig. 2.



(a) Affine covariant region (b) Normalized circular region

Fig. 2. Affine covariant region and its normalized circular region

3.3 Local Image Descriptor

In order to obtain the rotation invariant gradient measurement, a rotation invariant local coordinate system is constructed for each pixel. Given the center of the circular region ' P ' and any one of the pixels inside the region ' X_i ', a local coordinate system for X_i can be constructed, as shown in **Fig. 3**. Apparently, gradient measurement based on such local coordinate system will be rotation invariant. The gradient is calculated accordingly as Eq. (8) and (9).

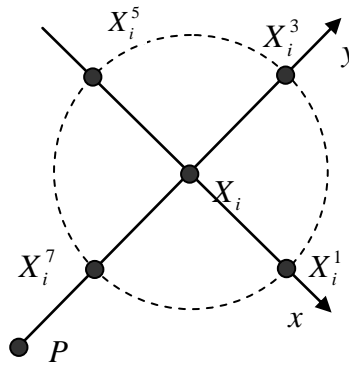


Fig. 3. Rotation invariant local coordinate system

$$D_x(X_i) = I(X_i^1) - I(X_i^5) \quad (8)$$

$$D_y(X_i) = I(X_i^3) - I(X_i^7) \quad (9)$$

where X_i^j , $j = 1, 3, 5, 7$ are neighbor pixels of X_i along x-axis and y-axis, and $I(X_i^j)$ denotes the intensity of pixel X_i^j . The gradient magnitude $m(X_i)$ and orientation $\phi(X_i)$ are obtained by:

$$m(X_i) = \sqrt{D_x(X_i)^2 + D_y(X_i)^2} \quad (10)$$

$$\phi(X_i) = \tan^{-1}(D_y(X_i) / D_x(X_i)) \quad (11)$$

$\phi(X_i) \in [0, 2\pi)$ which is quantized into 8 bins, $dir_i = (2\pi / d) \times (i - 1)$, $i = 1, 2, \dots, d$ ($d=8$ in this paper). And then, the gradient of X_i is transformed into a d -dimensional vector, denoted as $F_G(X_i) = (f_1^G, f_2^G, \dots, f_d^G)$, where,

$$f_j^G = \begin{cases} m(X_i) \frac{(2\pi / d - \alpha(\phi(X_i), dir_j))}{2\pi / d}, & \text{if } \alpha(\phi(X_i), dir_j) < 2\pi / d \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

$\alpha(\phi(X_i), dir_j)$ is the difference between $\phi(X_i)$ and dir_j .

Gradient of the pixels in each sub-region are summed together to become feature vector for every sub-region respectively. Then, the feature vectors of all sub-regions are concatenated together to represent this normalized region,

$$D(R) = (F(S1), F(S2), F(S3), F(S4)) \quad (13)$$

where $F(Si)$ is the feature vector of each sub-region Si , i.e.,

$$F(Si) = \sum_{X \in R_i} F_G(X) \quad (14)$$

In order to enhance the discriminability of descriptor, the method of utilizing multi-support region as [13, 17] is also followed to construct the proposed descriptor. The number of the multi-support region is 4 in this paper.

Feature vectors calculated from all support regions are concatenated together to form our final descriptor vector: $\{D_1, D_2, D_3, D_4\}$.

4. Experiments

4.1 Data Set

Benchmark dataset [18] is employed in this paper. These images are either from planar scene or captured by a fixed position camera during acquisition. Therefore, the relationship between transformed images and reference image can be modeled by a 2D homography matrix. Moreover, the dataset also provides homography for provided images. Fig. 4 shows the data set. The types of transformations include changing viewing angle, scaling, rotation, image blurring, changing illumination, and image compression. The reference image is the 1st column image (1st), the transformed images are these from the 2nd to 5th column image (2nd to 5th).

4.2 Region Detector

The relative performance among different descriptors is consistent with different feature detectors [13]. Since Hessian-affine detector can detect blob-like points, which are less likely at the positions of depth-difference pixel points and favor local planarity and smoothness assumption [12], it is selected in our experiments.

4.3 Similarity Measurement

For region matching, three strategies proposed in [8] are nearest neighbor (NN), nearest neighbor distance ratio (NNDR) and threshold. Although these three matching methods are functionally different, their ranking results of the performances of the various descriptors are virtually the same [8]. The strategy of NN was adopted in this paper.



(a) Image blur (tree)



(b) Image blur (bikes)



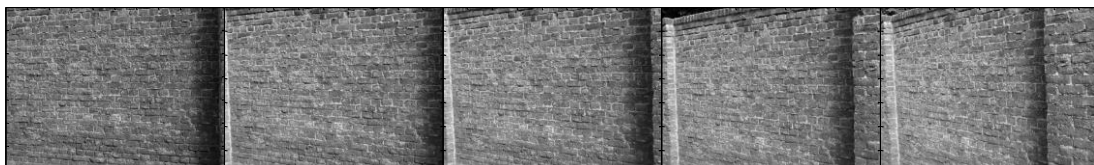
(c) Illumination changes (leuven)



(d) JPEG compression (ubc)



(e) Viewpoint changes (graf)



(f) Viewpoint changes (wall)



(g) Rotation and scale changes (boat)

Fig. 4. Data set

4.4 Evaluation

The metric proposed by Krystian Mikolajczyk and Cordelia Schmid is adopted to evaluate the performance of MEGH. Two values, recall and 1-precision, are employed as the evaluation criteria which are based on the numbers of correct matches and false matches by checking the reference image and transformed images [8]. Recall is the ratio of the correctly matched number to the number of corresponding regions. 1-precision is the ratio of the number of false matches to the total number of matches. The curve of Recall v.s. 1-Precision (i.e. Precision-Recall) [8] is drawn by changing the similarity measurement threshold.

4.5 Experimental Results

In our experiments, the proposed method is compared against SIFT, GLOH, PCA-SIFT, spin-image (denoted as SPIN), and MROGH. The implements of these benchmark descriptors follow the codes provided in [18, 19].

For the case of image blurring, Fig. 5-6 show the results. The blur degree increases gradually from the 2nd to 5th images from Fig. 4 (a, b). From Fig. 5, it is seen that MEGH achieves the best performance. From Fig. 6, it is shown that the proposed descriptor along with MROGH achieves better performance than other descriptors while it is slightly worse than MROGH.

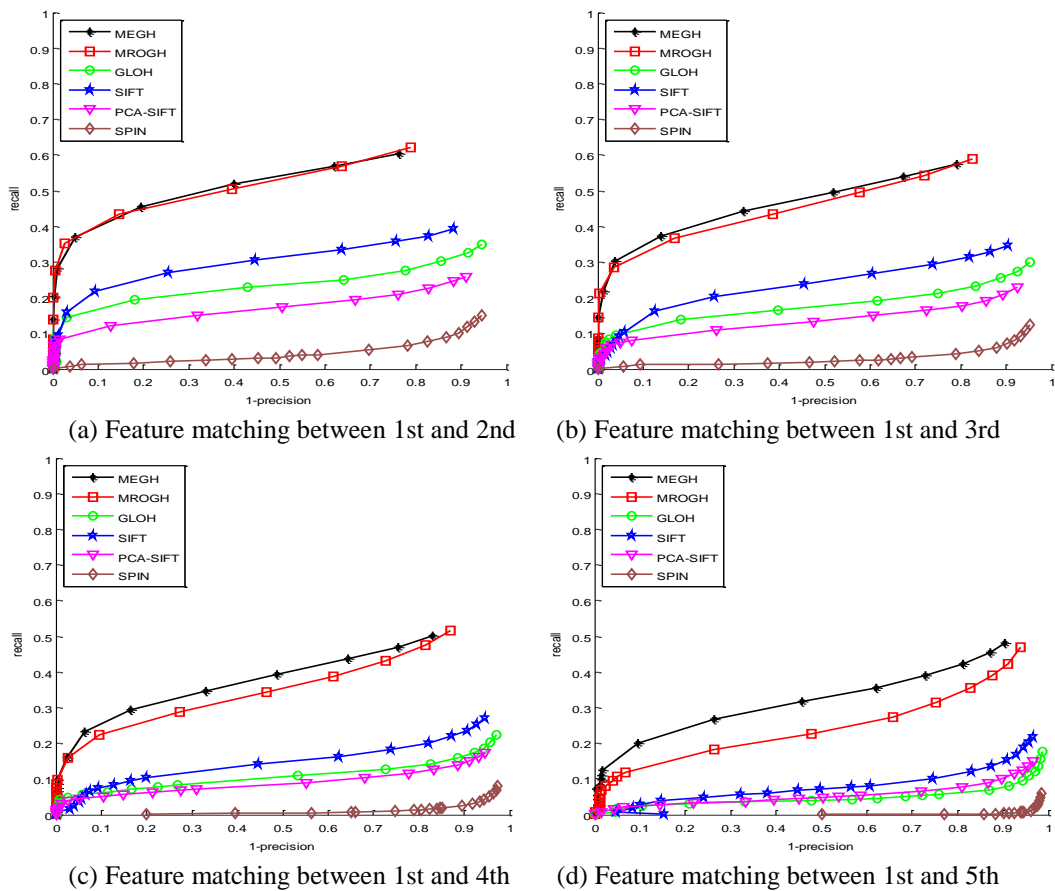


Fig. 5. Performance for the case of image blurring - tree

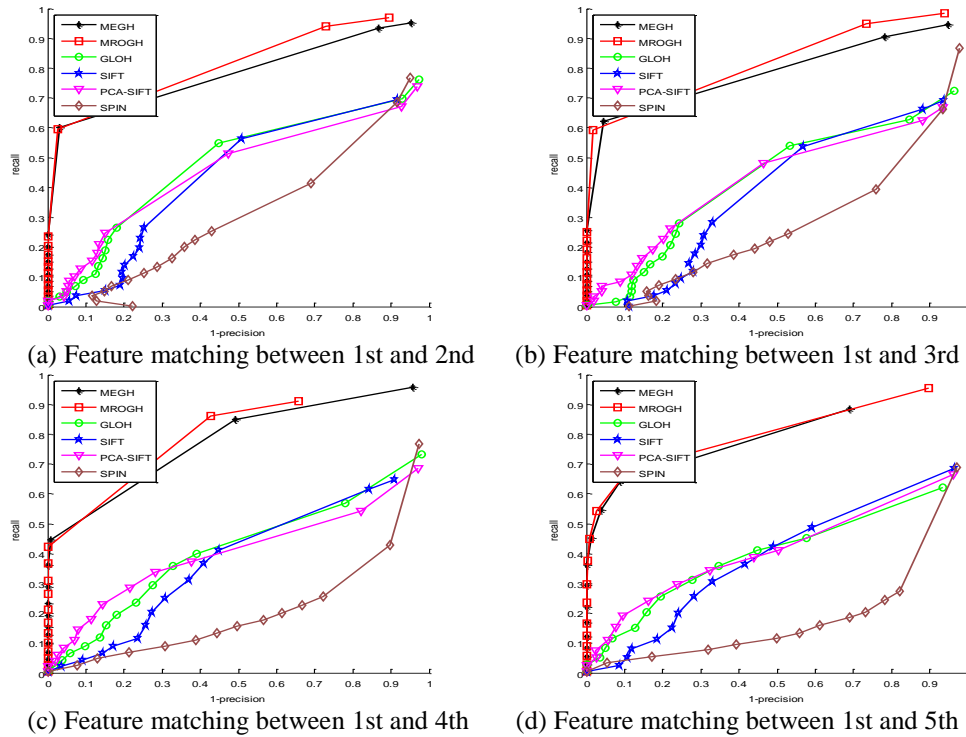


Fig. 6. Performance for the case of image blurring - bikes

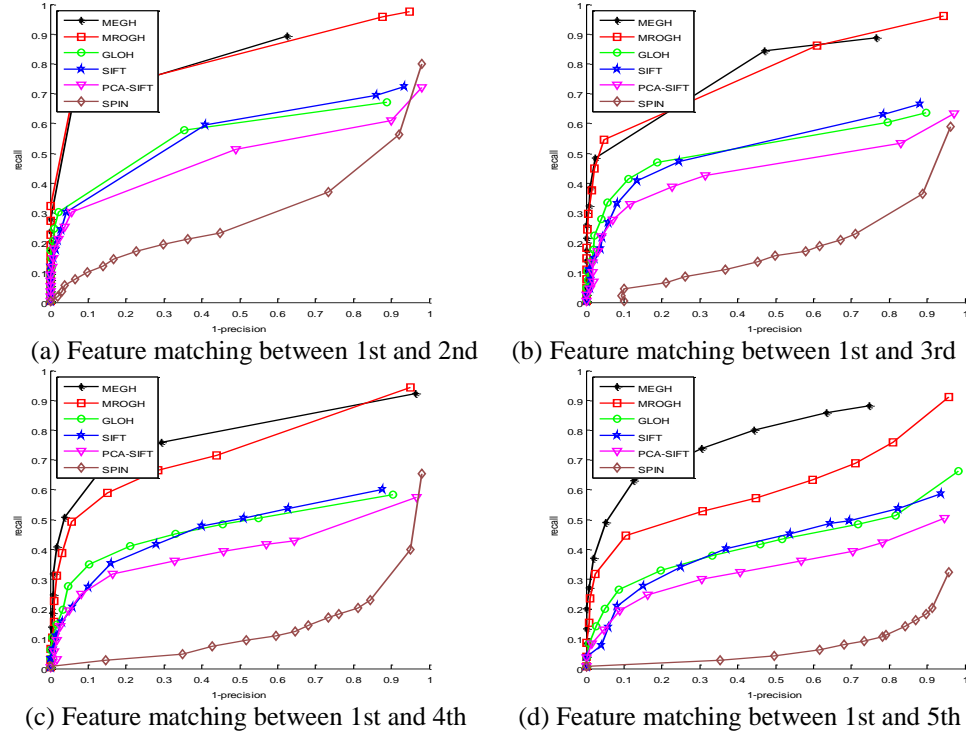


Fig. 7. Performance for the case of illumination change – leuven

Fig. 7 shows the performance for the case of illumination change. The level of illumination change gradually becomes larger the 2nd to 5th images from **Fig. 4(c)**. It is seen that overall MEGH performs better than other descriptors.

Fig. 8 shows the performance for the case of JPEG image compression. The compression ratio increases gradually from the 2nd to 5th images from **Fig. 4(d)**. It is shown that MEGH achieves the best performance.

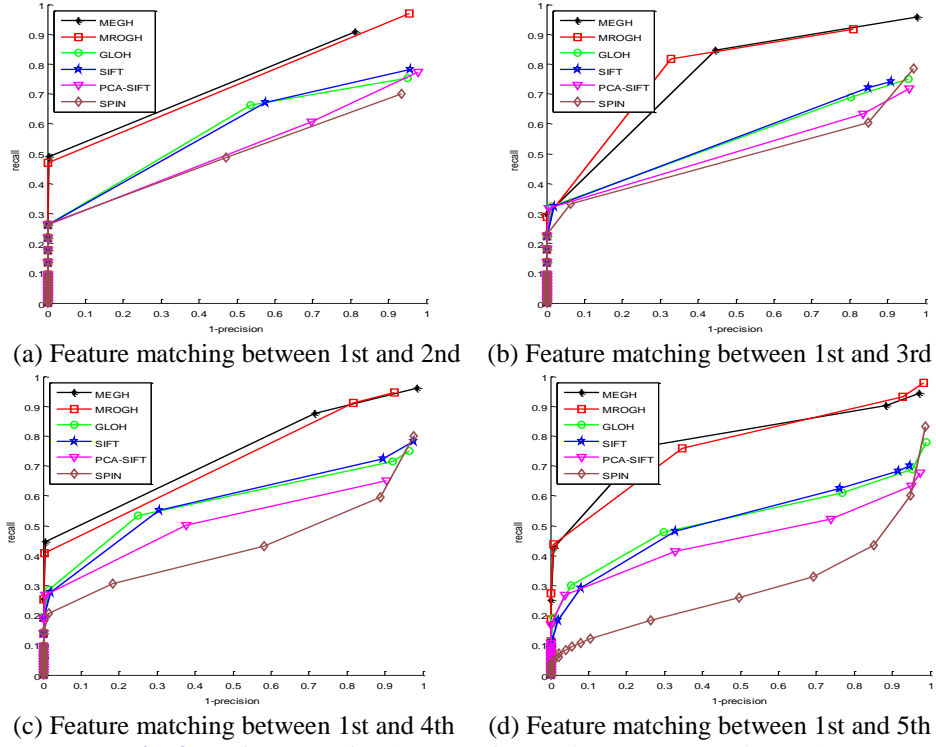


Fig. 8. Performance for the case of JPEG image compression - ubc

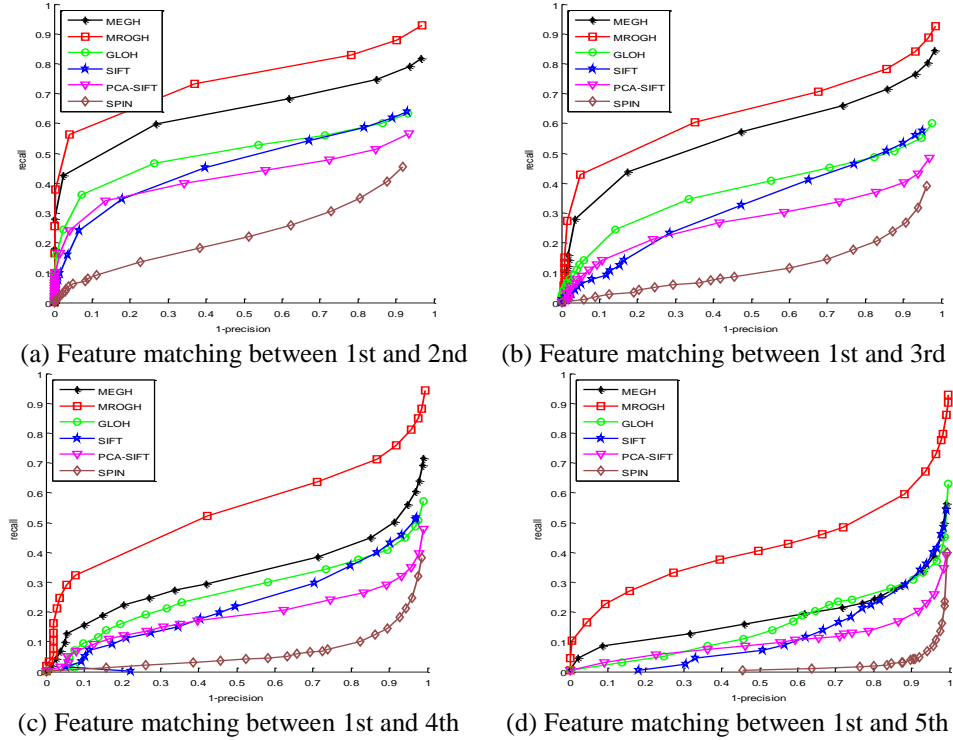


Fig. 9. Performance for the case of viewing angle change - graft

Fig. 9-10 show the performance for the case of viewing angle change. It is shown that the performance of the proposed method is lower than MROGH and higher than any other descriptors.

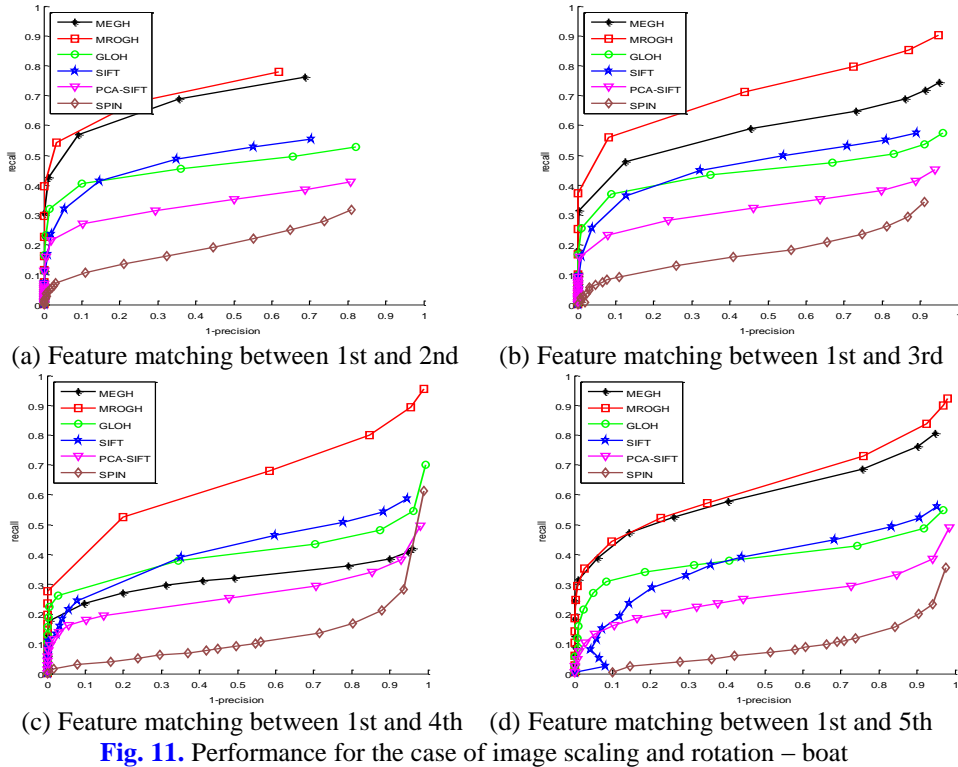
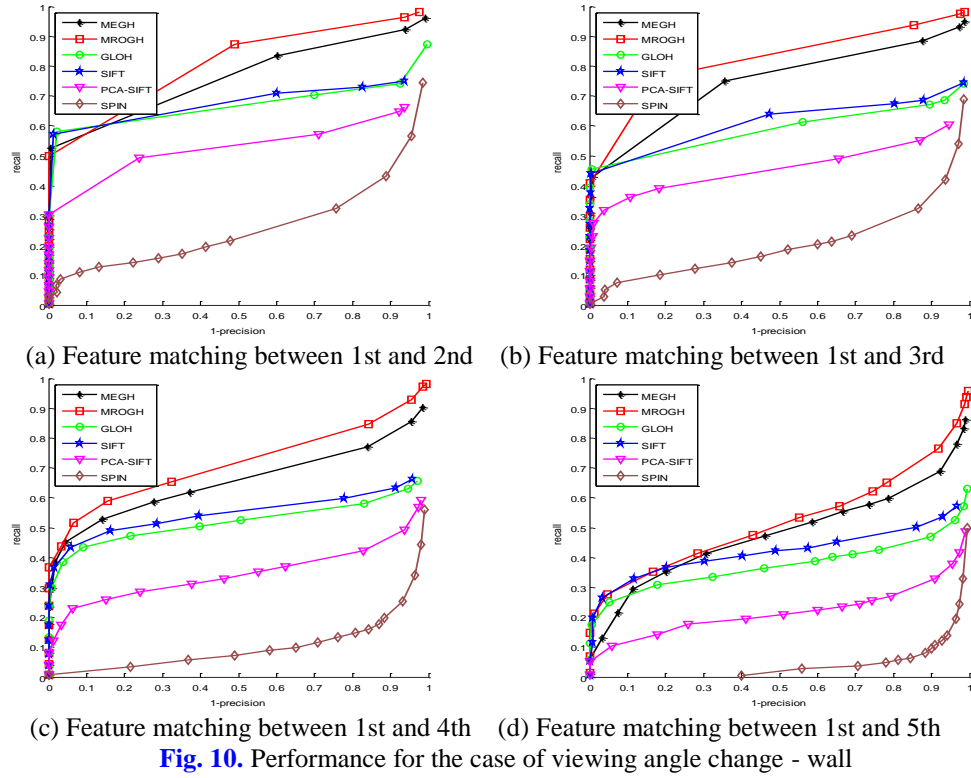


Fig. 11 shows the performance for the case of image scaling and rotation. It is shown that the performance of the proposed method is overall lower than MROGH and higher than any other descriptors.

From the experiments above, it is seen that there does not exist one descriptor which outperforms the other descriptors for all scene types and for all types of transformations. For the case of photometric transformations (**Fig. 5-8**), the proposed method achieves the best performance against all compared methods. For the case of geometric transformations, the performance of the proposed method is lower than that of MROGH and higher than that of any other descriptors. The performance difference between MEGH and MROGH generally increases with the severity of the transformations. This is because estimating ellipse affine covariant region is sensitive under geometric transformation, particularly, to viewpoint change and rotation.

5. Conclusions

In this paper, a special image segmentation based on affine covariant region is proposed. The proposed feature descriptor MEGH based on image segmentation is presented, which is robust to the photometric transformations and geometric transformations.

Compared with 5 existing descriptors, extensive experiments show that the feature descriptor based on ellipse orientation rather than an estimated reference orientation achieves better performance.

In the case of geometric transformations, the proposed segmentation process using ellipse orientation is still robust although the performance of MEGH is lower than that of MROGH. This is due to an estimation error of the affine invariant region itself.

Further investigation on the study of the affine covariant region and the incorporation of the proposed descriptor into various applications is currently underway.

References

- [1] Zen Chen, and Shu Kuo Sun, "A Zernike moment phase-based descriptor for local image representation and matching," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 205-219, January, 2010. [Article \(CrossRef Link\)](#)
- [2] J. Matas, O. Chum, M. Urban and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761-767, 2004. [Article \(CrossRef Link\)](#)
- [3] Tinne Tuytelaars and Luc Van Gool, "Matching widely separated views based on affine invariant regions," *International journal of computer vision*, vol. 59, no. 1, pp. 61-85, 2004. [Article \(CrossRef Link\)](#)
- [4] Krystian Mikolajczyk and Cordelia Schmid. "Scale & affine invariant interest point detectors," *International journal of computer vision*, vol. 60, no. 1, pp. 63-86, 2004. [Article \(CrossRef Link\)](#)
- [5] Timor Kadir, Andrew Zisserman and Michael Brady. "An affine invariant salient region detector," in *Proc. of 8th European Conference on Computer Vision*, pp. 228-241, May 11-14, 2004. [Article \(CrossRef Link\)](#)
- [6] Krystian Mikolajczyk, T. Tuytelaars, C. Schmid, et al. "A comparison of affine region detectors," *International journal of computer vision*, vol. 65, no. 1-2, pp. 43-72, 2005. [Article \(CrossRef Link\)](#)
- [7] David G Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91-110, 2004. [Article \(CrossRef Link\)](#)
- [8] Krystian Mikolajczyk and Cordelia Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615-1630, 2005.

- [Article \(CrossRef Link\)](#)
- [9] Herbert Bay, Andreas Ess, Tinne Tuytelaars, et al, "Speeded-up robust features (SURF)," *Computer vision and image understanding*, vol.110, no. 3, pp. 346-359, 2008. [Article \(CrossRef Link\)](#)
 - [10] Yan Ke and Rahul Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 503-513, 2004. [Article \(CrossRef Link\)](#)
 - [11] Engin Tola, Vincent Lepetit and Pascal Fua, "Daisy: An efficient dense descriptor applied to wide-baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 815-830, 2010. [Article \(CrossRef Link\)](#)
 - [12] Congxin Liu, Jie Yang and Deying Feng, "PPD: A Robust Low-computation Local Descriptor for Mobile Image Retrieval," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 3, pp. 305-323, 2010. [Article \(CrossRef Link\)](#)
 - [13] Bin Fan, Fuchao Wu and Zhanyi Hu, "Rotationally Invariant Descriptors Using Intensity Order Pooling." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no.10, pp. 2031-2045, 2012. [Article \(CrossRef Link\)](#)
 - [14] Svetlana Lazebnik, Cordelia Schmid and Jean Ponce, "A sparse texture representation using local affine regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265-1278, 2005. [Article \(CrossRef Link\)](#)
 - [15] Canlin Li and Lizhuang Ma, "A new framework for feature descriptor based on SIFT," *Pattern Recognition Letters*, vol. 30, no. 5, pp. 544-557, 2009. [Article \(CrossRef Link\)](#)
 - [16] Adam Baumberg, "Reliable feature matching across widely separated views," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 774-781, 2000. [Article \(CrossRef Link\)](#)
 - [17] Hong Cheng, Zicheng Liu, Nanning Zheng, Jie Yang, "A deformable local image descriptor," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1-8, 2008. [Article \(CrossRef Link\)](#)
 - [18] <http://www.robots.ox.ac.uk/~vgg/research/affine/>
 - [19] <http://www.sigvc.org/bfan/>



Xiaojie Dong received the M.S. degree from Beijing University of Aeronautics & Astronautics. Now he is a Ph.D. candidate of Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University. His current research interests include image processing and image registration. E-mail: xjdong_2010@sjtu.edu.cn



Erqi Liu works in China Aerospace Science and Industry Corporation, and he is the adjunct professor of Shanghai Jiaotong University. His current research interests include precision guidance and engineering management. E-mail: everyth_ok@yahoo.com.cn



Jie Yang received his Ph.D. degree in computer science from University of Hamburg, Germany. Now he is a professor and the Director of Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University. His current research interests include image processing, pattern analysis, and computational intelligence. E-mail: jieyang@sjtu.edu.cn



Qiang Wu received the B.Eng. and M.Eng. degrees in electronic engineering from the Harbin Institute of Technology, Harbin, China, in 1996 and 1998, respectively, and the Ph.D. degree in computing science from the University of Technology Sydney, Sydney, Australia, in 2004. He is currently a Senior Lecturer with the School of Computing and Communications, University of Technology Sydney. He is the author of more than 70 refereed papers in these areas, including those published in prestigious journals and top international conferences. His major research interests include computer vision, image processing, pattern recognition, machine learning, and multimedia processing. Dr. Wu has been a Guest Editor of several international journals, such as the Pattern Recognition Letters (PRL) and the International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI). He has served as a Chair and/or a Program Committee Member for a number of international conferences. He has also served as a Reviewer for several international journals, such as PRL, IJPRAI, the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART B: CYBERNETICS, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (TCSVT), the Pattern Recognition, and the EURASIP Journal on Image and Video Processing.