Performance Evaluation of Local Descriptors for Affine Invariant Region Detector

Man Hee Lee and In Kyu Park

Department of Information and Communication Engineering
Inha University 100 Inha-ro, Incheon 402-751, Korea
maninara@hotmail.com and pik@inha.ac.kr

Abstract. Local feature descriptors are widely used in many computer vision applications. Over the past couple of decades, several local feature descriptors have been proposed which are robust to challenging conditions. Since they show different characteristics in different environment, it is necessary to evaluate their performance in an intensive and consistent manner. However, there has been no relevant work that addresses this problem, especially for the affine invariant region detectors which are popularly used in object recognition and classification. In this paper, we present a useful and rigorous performance evaluation of local descriptors for affine invariant region detector, in which MSER (maximally stable extremal regions) detector is employed. We intensively evaluate local patch based descriptors as well as binary descriptors, including SIFT (scale invariant feature transform), SURF (speeded up robust features), BRIEF (binary robust independent elementary features), FREAK (fast retina keypoint), Shape descriptor, and LIOP (local intensity order pattern). Intensive evaluation on standard dataset shows that LIOP outperforms the other descriptors in terms of precision and recall metric.

1 Introduction

Visual feature detection and description are widely used in most computer vision algorithms including visual SLAM (simultaneous localization and mapping) [1], structure from motion [2], object recognition [3], object tracking [4], and scene classification [5]. Various feature detectors and descriptors have been developed such as SIFT (scale invariant feature transform) [6] and SURF (speeded up robust features) [7]. SIFT is the one of the state-of-the-art algorithms with good repeatability and matching accuracy. SIFT detects local features using scale space extrema of DoG (difference of Gaussians) and describes feature point using HOG (histogram of oriented gradients). In addition to them, a considerable number of previous work have done to describe local features effectively. A rigorous survey of performance evaluation of local descriptor can be found in Mikolajczyk and Schmid's work [8].

On the other hand, robust region detectors have been developed such as Harris affine [9], Hessian affine [10], and MSER (maximally stable extremal regions) [11] detectors. Many computer vision applications utilize these detectors

because affine invariant region is robust to affine transformation and has better repeatability than local feature detector under significant viewpoint changes. Conventional affine invariant region detectors do not have their own inherent descriptors. Consequently, traditional feature descriptors have to be utilized to describe and match detected regions. Note that customized descriptors for describing affine invariant regions have been introduced, *e.g.* shape descriptor [12].

In this paper, we evaluate the performance of local descriptors for affine invariant region detectors. To the best knowledge of the authors, there has been no previous work that addresses this problem in recent years. While employing MSER detector for the affine invariant region detection, we compare standard local patch based descriptors (SIFT and SURF) as well as the state-of-the-art binary descriptors including BRIEF (binary robust independent elementary features) [13], FREAK (fast retina keypoint) [14], and LIOP (local intensity order pattern) [15]. The performance of those descriptors is evaluated and compared in various scenes with different zooming, rotation, large viewpoint changes, object deformation, and large depth variation.

This paper is organized as follows. In Section 2, the existing performance evaluation of feature descriptors is introduced. Section 3 describes the evaluation framework and criteria with brief summary of the evaluated region detectors and descriptors. The experimental result and discussion are presented in Section 4. Finally, we give a conclusive remark in Section 5.

2 Related Work

Table 1 shows the summarization of the previous performance evaluation of feature descriptors.

Mikolajczyk and Schmid [8] evaluated the performance of local feature descriptors in various geometric and photometric transformations, which is known to be the most exhaustive work. In addition, they proposed GLOH (gradient location and orientation histogram) descriptor which was the extension of SIFT descriptor using log-polar location grid. They concluded that GLOH and SIFT obtained the best performance to handle image rotation, zoom, viewpoint change, image blur, image compression, and illumination change.

Moreels and Perona [18] compared the feature detectors and descriptors for diverse 3D objects. They generated database which consists of 144 different objects with viewpoint and illumination changes. Several combinations of feature detectors and descriptors were evaluated, which shows that Hessian-affine detector combined with SIFT descriptor demonstrated the best performance.

Dickscheid *et al.* [21] measured the completeness of local features for image coding. They proposed the qualitative metric for evaluating the completeness of feature detection using feature density and entropy density. In their experiment, MSER detector achieved the best performance.

Dahl *et al.* [19] compared different pairs of local feature detectors and descriptors on the multi-view dataset. It was observed that MSER and DoG detectors with SIFT descriptor obtained the best performance.

Table 1: Previous works on performance evaluation of feature detectors/descriptors.

Author	Type	Environment	Best result
Mikolajczyk [8]	local descriptor	geometric + photometric transform	GLOH, SIFT
Miksik [16]	local descriptor	accuracy and speed	LIOP, BRIEF
Restrepo [17]	shape descriptor	object classification	FPFH
Moreels [18]	detector + descriptor	3D object	Hessian-affine + SIFT
Dahl [19]	detector + descriptor	multi-view dataset	MSER + SIFT
Mikolajczyk [10]	affine region detector	geometric + photometric transform	MSER
Haja [20]	region detector	texture + structure	MSER
Dickscheid [21]	local detector	image coding	MSER
Canclini [22]	local detector	image retrieval	BRISK

The performance of local shape descriptors for object classification task was evaluated by Restrepo and Mundy [17]. The local shape descriptors were extracted from the probabilistic volumetric model. They compared several shape descriptors to classify object categories using *Bag of Words* model from large scale urban scenes. FPFH (fast point feature histogram) obtained good performance in their experiments.

Miksik and Mikolajczyk [16] evaluated the trade off between accuracy and speed of local feature detectors and descriptors. They evaluated the performance of several binary descriptors and local intensity order descriptors. It was shown that binary descriptors outperformed other descriptors in time-constrained applications with low memory requirement.

Canclini *et al.* [22] evaluated the performance of feature detectors and descriptors for image retrieval application. They compared several low-complexity feature detectors and descriptors, which concluded that binary descriptors achieved better performance than non-binary descriptors in terms of matching accuracy and computational complexity.

Although the previous works have addressed the problem of performance evaluation of different feature detectors and descriptors, they are out of dated and do not consider the recently proposed descriptors. Recently computer vision applications constantly need the performance evaluation of contemporary state-of-the-art algorithms, which is the main motivation of this paper. In this paper, rather than providing too general performance evaluation which can be vague in practical point of view, we narrow the focus down to performance evaluation of descriptors combined with affine invariant region detection. This combination

has not been addressed in the previous literatures. Furthermore, recent state-of-the-art descriptors are fully covered in this paper.

3 Performance Evaluation Framework

3.1 Affine Invariant Region Detector

Several techniques have been proposed on affine invariant region detector such as Harris affine, Hessian affine, and MSER detectors. Harris affine region detector [9] detects interest points using multi-scale Harris detector, which is invariant to scale and affine transformation. The extremum in the scale space of Laplacian of Gaussian is selected as the proper scale of an interest point. The elliptical region at the interest point is estimated iteratively using second moment matrix.

Similar to Harris affine region detector, Hessian affine region detector [10] is also known to be invariant to scale and affine transformation. The interest points are detected using Hessian matrices which have strong response on blobs and ridges. The scale is estimated using the Laplacian over scale space. To estimate the elliptical affine shape, the second moment matrix is used too.

MSER [11] are the regions defined as connected components which are obtained by thresholding. The detected extremal regions are either darker or brighter than surrounding region. In addition, MSER can extract important regions regardless of the threshold. MSER has the following desirable properties.

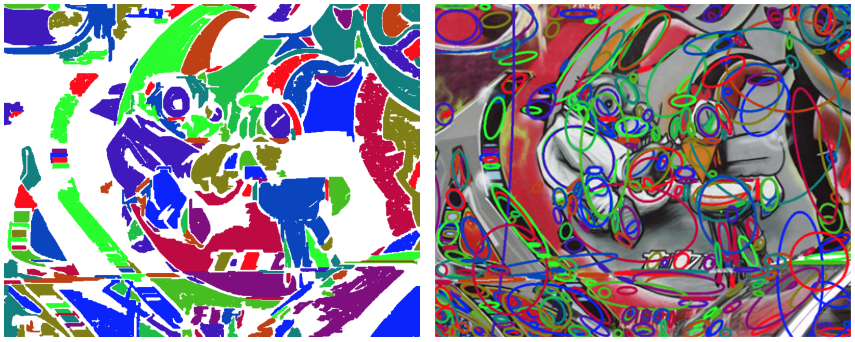
- Invariance to affine transformation and image intensity changes
- Adjacency of neighboring components is preserved in continuous geometric transformation
- Multi-scale detection
- Approximately linear complexity

The performance of various affine region detectors was compared by Mikolajczyk *et al.* [10], which shows that MSER detector has better repeatability than others in many cases. In addition, Haja *et al.* [20] compared the performance of different region detectors in terms of shape and position accuracy. They showed that MSER obtained the best accuracy than other detectors.

Based on the conclusion of those literatures, MSER is employed for the affine invariant region detector in this paper. Fig. 1 shows the typical result of affine invariant region detection using MSER detector.

3.2 Selected Descriptors to be Evaluated

Mikolajczyk and Schmid [8] evaluated the performance of several local descriptors including SIFT, GLOH (gradient location and orientation histogram), shape context, PCA-SIFT, spin images, steerable filters, differential invariants, complex filters, moment invariants, and cross-correlation of sampled pixel values. Their evaluation showed that SIFT outperformed other descriptors. The proposed evaluation framework is designed to provide valuable performance comparison of recent descriptors while avoiding duplicated evaluation with previous



(a) (b)

Fig. 1: Example of MSER detection. (a) Affine invariant regions detected by MSER. Each region is visualized as different color. (b) Each region is fit to elliptical shape.

literatures. Therefore, among the descriptors that Mikolajczyk and Schmid evaluated, only SIFT is selected in the proposed evaluation framework. In addition to SIFT, recent state-of-the-art feature descriptors, *i.e.* SURF [7], BRIEF [13], FREAK [14], Shape descriptor [12], and LIOP [15], which were published after Mikolajczyk and Schmid's work, have been included in the proposed evaluation framework. Therefore, a total of six descriptors are evaluated in this paper.

SIFT [6] descriptor is a distinctive local descriptor which is invariant to the scale and illumination changes. In our implementation, gradient magnitude and orientation is sampled in a 16×16 region around the keypoint. Then, orientation histograms (quantized to 8 directions) are generated over 4×4 subregion of the original sampling region. To increase the robustness to small location changes, the magnitude of each sample is weighted by Gaussian weighting function. Since there are 4×4 histograms with 8 orientation bins, the descriptor is represented by 128 dimensional feature vector at each keypoint.

SURF [7] descriptor is the speeded up version of SIFT descriptor. SURF descriptor is used widely in the feature matching as well as SIFT descriptor. The integral image and binary approximated integer Gaussian filter are utilized to approximate SIFT descriptor with significantly low computation. Because the gradient values within a subpatch are integrated to generate a SURF descriptor, it is more robust to image noise than SIFT.

BRIEF [13] is the notable binary descriptor that is computed by pairwise intensity comparison. To reduce the influence of noise and therefore to increase stability and the repeatability, local patches are first smoothed using Gaussian filter. Then, binary test samples are selected from an isotropic Gaussian distribution. Hamming distance is used to compute the distance between BRIEF descriptors.

FREAK [14] is the up-to-date binary descriptor that is biologically motivated by human retinal structure. The sequence of binary string is computed by

pairwise comparison of image intensities over a retinal sampling pattern which are obtained by training data. Each sample point is smoothed with different size of Gaussian kernel so that it becomes less sensitive to noise. In the matching procedure, most of the outliers are removed by comparing first 16 bytes which represents coarse information.

Shape descriptor [12] is an affine invariant descriptor designed for MSER, which uses the shape of the detected MSER itself as the descriptor. In each local patch, the detected region and the background are converted to the white and black, respectively. The gradient histogram based descriptor is constructed similar with SIFT descriptor.

LIOP [15] descriptor is known as the state-of-the-art feature descriptor. LIOP is also invariant to image rotation and intensity change by encoding local ordinal information of each pixel. A patch is divided to subregions using the intensity order based region division method. In each subregion, intensity relation between each pixel is mapped to the appointed value. The histogram of these values is used as the descriptor of the subregion. The overall descriptor is constructed by accumulating the histogram.

The original implementations of these algorithms are provided by the authors directly or indirectly via their contribution in OpenCV[1] and VLFeat[2].

3.3 Descriptor Matching

In this paper, each region is fit to an elliptical shape which is subsequently warped to a square patch. The orientation of normalized patch is estimated from the histogram of gradient directions. Then, nearest neighbor thresholding method is utilized to match each descriptor, in which matching pair is identified if the distance is smaller than a threshold. Note that, thresholding based matching is commonly used to evaluate the performance of descriptor, because it can explain well how many descriptors are similar to each other. However, the distinctiveness of each descriptor has been already shown in their original papers. Since this paper is intended to investigate the performance itself of each descriptor, we utilize the high performance matching method, i.e. nearest neighbor thresholding method. Correct matching is determined by the overlapping ratio of the reprojected region [9].

3.4 Dataset

In the proposed evaluation framework, test images are selected from three popular dataset, including Mikolajczyk and Schmid's dataset [8], Salzmann's dataset[3] [23], and Moreels's dataset[4] [18].

[1] http://www.opencv.org
[2] http://www.vlfeat.org
[3] http://cvlab.epfl.ch/data/dsr
[4] http://vision.caltech.edu/pmoreels/Datasets/TurntableObjects/

Mikolajczyk and Schmid's dataset [8] is used for measuring the performance under viewpoint change, image zoom/rotation. The images in the dataset have other imaging conditions such as illumination change, image blur, and JPEG compression. However, they are not included in our experiments since we focus on the affine invariant property of MSER. For in-plane image rotation and scale change, *boat* and *bark* datasets are used. And we utilize *graffiti* and *bricks* datasets for viewpoint change. Salzmann's dataset has useful test images with deformable objects. For deformable objects, we test *bed sheet* and *cushion* datasets. In each dataset, we select several frames to evaluate the matching performance. Moreels's dataset consists of 144 different 3D objects with calibrated viewpoints under 3 different lighting conditions. We test *potato* and *volley ball* dataset for 3D objects. In each dataset, we select one image pair with 45 degree viewpoint change.

For quantitative evaluation of descriptor matching, the standard metric (recall vs. 1-precision plot) is utilized which was proposed in Mikolajczyk and Schmid's work [8].

4 Performance Evaluation Results

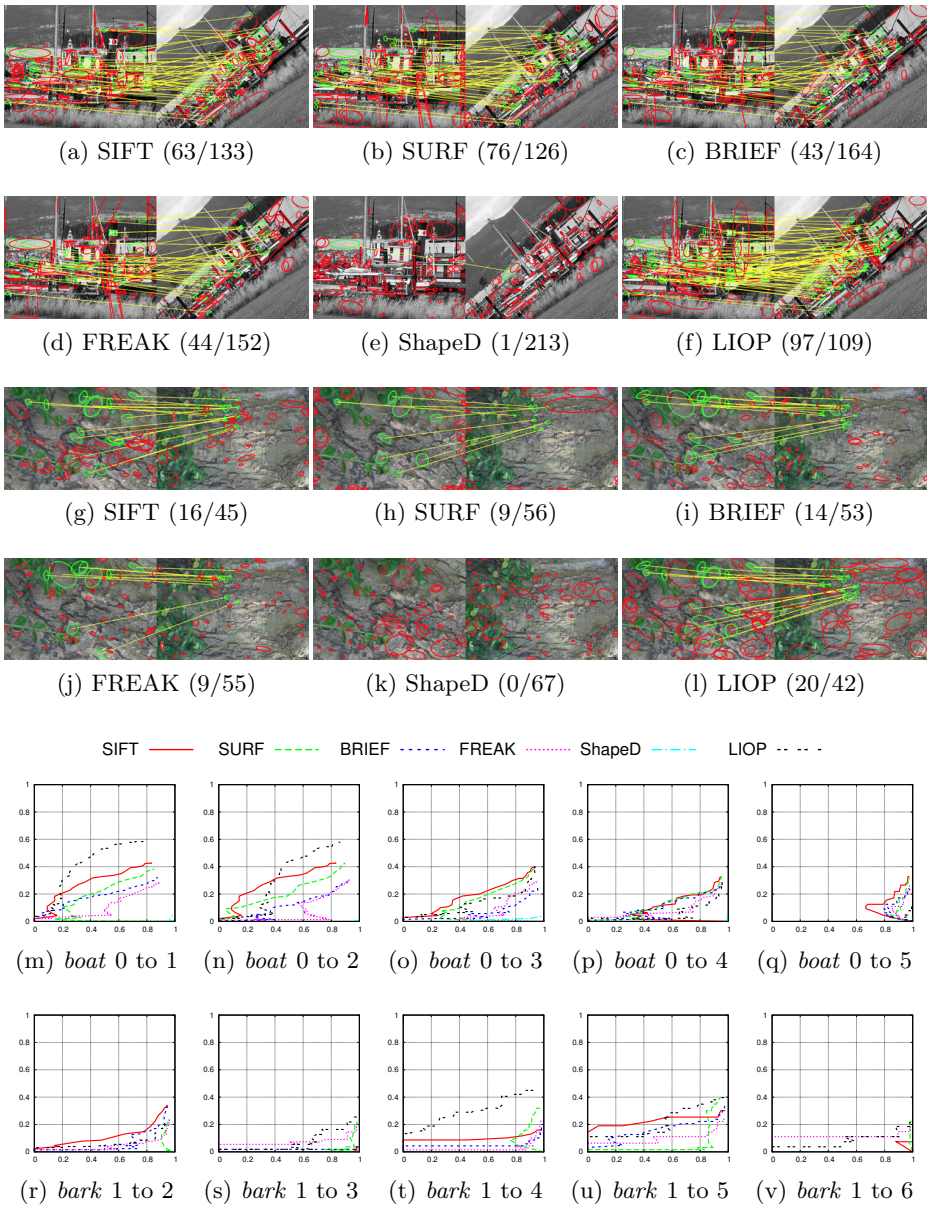
In this section, we summarize and compare the performance of SIFT, SURF, BRIEF, FREAK, Shape descriptor, and LIOP descriptors combined with MSER detector. Each warped patch is organized 144×144 pixels including 16 border pixels. The size of patch for BRIEF is set to 82×82 with same border. The scale of each scale invariant descriptor is fit to the patch size. Matching performance is measured by (number of inlier / number of outlier). In the following figures, green and red ellipses denote the correct and incorrect correspondences, respectively. Also in the following plots, horizontal and vertical axes represent 1-precision and recall, respectively. The experiment is carried out on Intel Core i7 2.7 GHz processor with 16GB memory.

4.1 Image Rotation and Scale Changes

boat and *bark* datasets [8] are used for in-plane image rotation and scale change. The correct match is determined by reprojecting each region using ground truth homography matrix. Fig. 2 shows the visual comparison of the matching results and the recall vs. 1-precision plot for *boat* and *bark* dataset. In each descriptor, we change the distance threshold for the nearest neighbor matching to measure the variation of the performance. As shown in Fig. 2, LIOP descriptor outperforms other descriptors in image zoom and rotation. Also, it is observed that SIFT descriptor achieves better performance than others including SURF descriptor. For binary descriptors only, BRIEF outperforms FREAK descriptor.

4.2 Viewpoint Changes

To evaluate the performance for viewpoint change, we utilize *graffiti* and *bricks* datasets [8] which varies their viewpoint approximately 50 degrees apart. The



(a) SIFT (63/133) (b) SURF (76/126) (c) BRIEF (43/164)

(d) FREAK (44/152) (e) ShapeD (1/213) (f) LIOP (97/109)

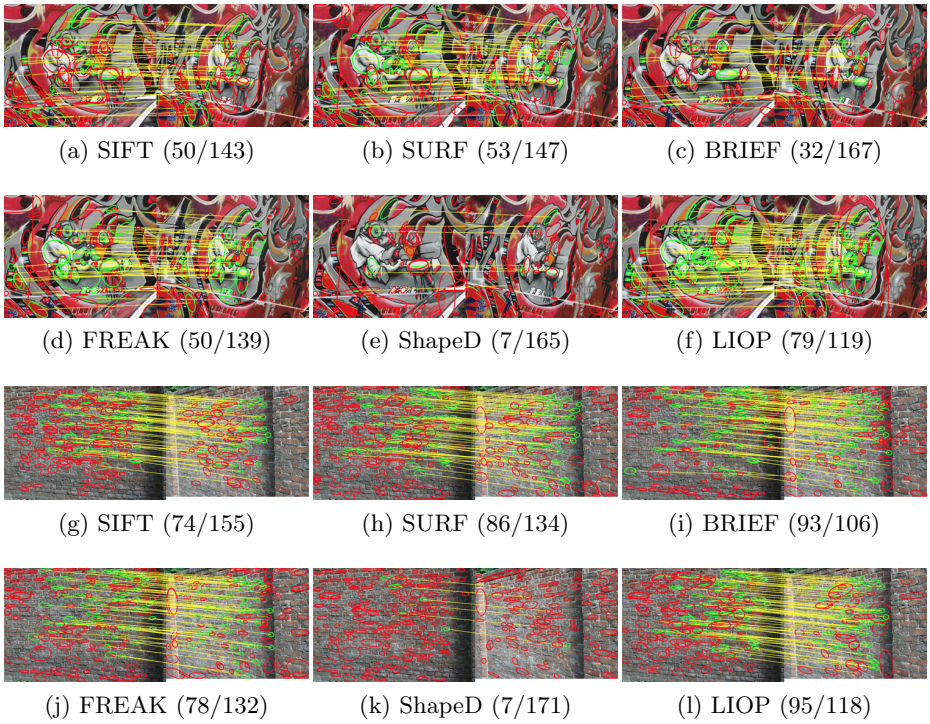
(g) SIFT (16/45) (h) SURF (9/56) (i) BRIEF (14/53)

(j) FREAK (9/55) (k) ShapeD (0/67) (l) LIOP (20/42)

(m) *boat* 0 to 1 (n) *boat* 0 to 2 (o) *boat* 0 to 3 (p) *boat* 0 to 4 (q) *boat* 0 to 5

(r) *bark* 1 to 2 (s) *bark* 1 to 3 (t) *bark* 1 to 4 (u) *bark* 1 to 5 (v) *bark* 1 to 6

Fig. 2: Matching performance evaluation for image rotation and scale change (*boat* and *bark* dataset). (a)∼(l) (number of inlier / number of outlier) of evaluated descriptors. (m)∼(v) recall vs. 1-precision plots.

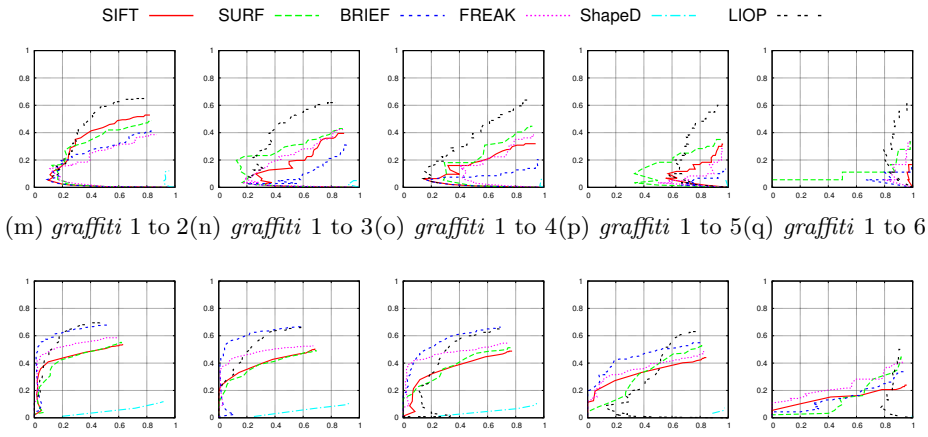


(a) SIFT (50/143) (b) SURF (53/147) (c) BRIEF (32/167)

(d) FREAK (50/139) (e) ShapeD (7/165) (f) LIOP (79/119)

(g) SIFT (74/155) (h) SURF (86/134) (i) BRIEF (93/106)

(j) FREAK (78/132) (k) ShapeD (7/171) (l) LIOP (95/118)



(m) *graffiti* 1 to 2(n) *graffiti* 1 to 3(o) *graffiti* 1 to 4(p) *graffiti* 1 to 5(q) *graffiti* 1 to 6

(r) *bricks* 6592 to 6593(s) *bricks* 6592 to 6594(t) *bricks* 6592 to 6595(u) *bricks* 6592 to 6596(v) *bricks* 6592 to 6597

Fig. 3: Matching performance evaluation for viewpoint change (*graffiti* and *bricks* dataset). (a)∼(l) (number of inlier / number of outlier) of evaluated descriptors. (m)∼(v) recall vs. 1-precision plots.

visual comparison of the matching result is shown in Fig. 3 (a)∼(l). In the scene with significant viewpoint change, LIOP descriptor also outperforms other descriptors. Fig. 3 (m)∼(v) shows the recall vs. 1-precision comparison of *graffiti* and *bricks* datasets. As shown in Fig. 3, the best performance is archived again by LIOP descriptor. Note that, since *bricks* image contains repeated patterns, MSER detector extracts uniform regions. In this case, all descriptors show higher performance than other dataset with nonuniform MSER regions.

4.3 Deformable Objects

To evaluate the performance for deformable objects, we test *bed sheet* and *cushion* dataset [23]. They provide 3D coordinates of ground truth 3D mesh and corresponding 2D coordinates of mesh vertices in images. Therefore, true correspondence can be estimated using 2D coordinate pairs. Fig. 4 shows the visual comparison of the matching results. As shown in Fig. 4, all descriptors show poor performance, which can be explained as follows. MSER is an affine invariant region detector which detects maximal or minimal connected regions in the image. If the detected region is changed to different elliptical shape due to the scene deformation, the normalized patch of each region is also changed. In that case, local descriptor matching is not going to be accurate. Nevertheless, SIFT and LIOP achieve relatively better performance than others.

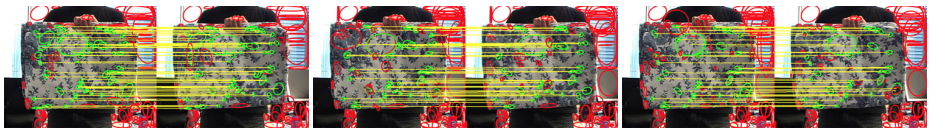
4.4 3D Objects

In Moreels's dataset [18], we test *potato* and *volley ball* models which have heavily textured objects. Fig. 5 shows the visual comparison of the matching results, which shows that, a few correspondences are matched with few inliers. Since the test image has significant viewpoint change, the matching performance decreases even though MSER is robust to the affine transform. Therefore, it is difficult to match features using MSER in the scene with significant viewpoint change and the deformable object.

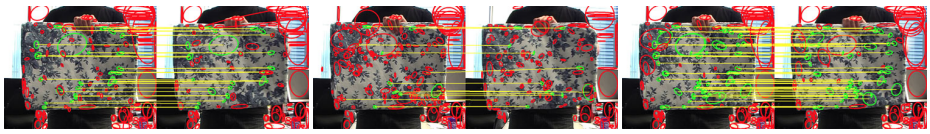
4.5 Processing Time

Table 2 presents the computational time for descriptor generation and nearest neighbor matching. In our MSER implementation, approximately 500 regions are detected from two images. Our observation can be summarized as follows.

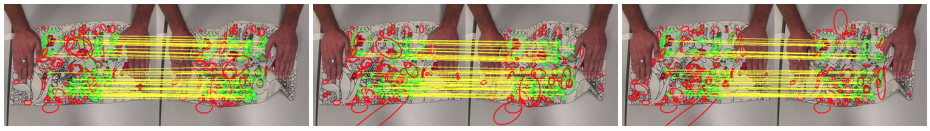
- Binary descriptors show the fastest description and matching speed. Binary descriptors are described with 256 and 512 bits and the difference of descriptors is calculated using Hamming distance.
- As is observed in other literatures, SURF is 5.6× faster than SIFT.
- Shape descriptor is slower than SIFT descriptor because SIFT can be constructed from gray scale image directly. On the other hand, in Shape descritpor generation, binary image of MSER regions has to be computed.



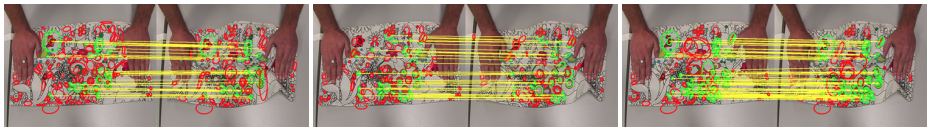
(a) SIFT (85/108) (b) SURF (66/138) (c) BRIEF (57/146)



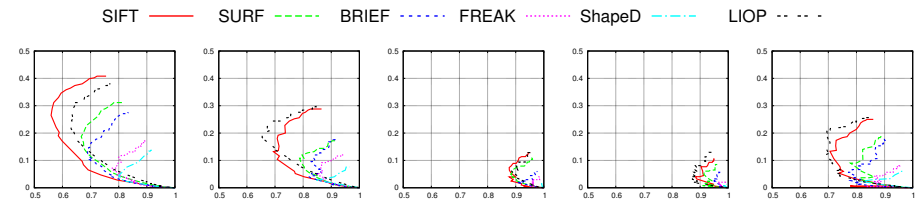
(d) FREAK (38/164) (e) ShapeD (18/223) (f) LIOP (77/132)



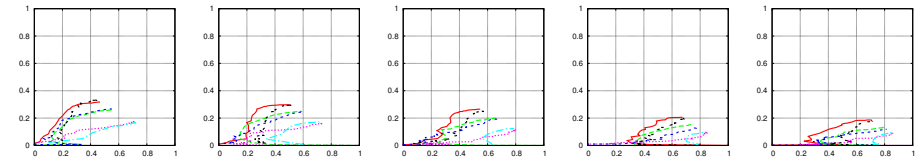
(g) SIFT (107/86) (h) SURF (93/114) (i) BRIEF (73/127)



(j) FREAK (45/160) (k) ShapeD (57/140) (l) LIOP (97/103)

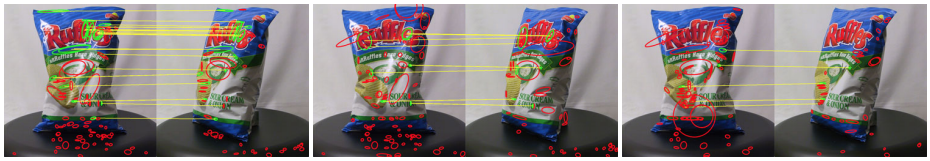


(m) *bed sheet* 122 to 129 (n) *bed sheet* 122 to 132 (o) *bed sheet* 122 to 148 (p) *bed sheet* 122 to 170 (q) *bed sheet* 122 to 180

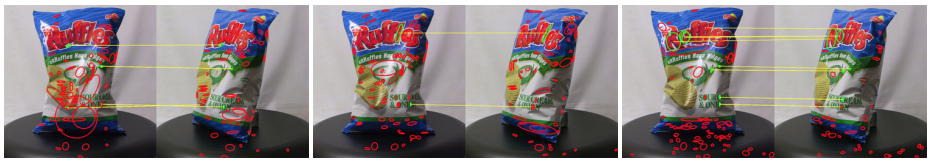


(r) *cushion* 144 to 160 (s) *cushion* 144 to 165 (t) *cushion* 144 to 170 (u) *cushoin* 144 to 177 (v) *cushoin* 144 to 184

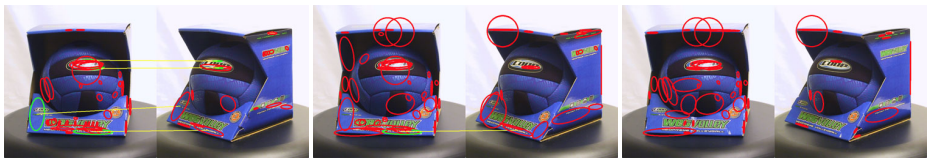
Fig. 4: Matching performance evaluation for deformable objects (*bed sheet* and *cushion* dataset). (a)∼(l) (number of inlier / number of outlier) of evaluated descriptors. (m)∼(v) recall vs. 1-precision plots.



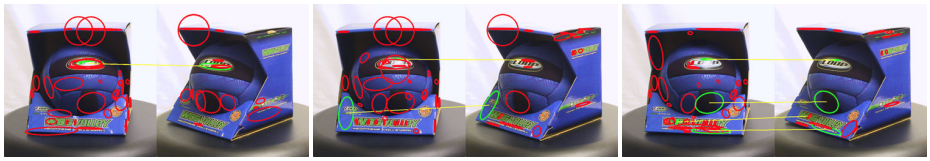
(a) SIFT (29/73) (b) SURF (12/88) (c) BRIEF (6/95)



(d) FREAK (7/89) (e) ShapeD (2/103) (f) LIOP (12/90)



(g) SIFT (4/47) (h) SURF (1/51) (i) BRIEF (0/50)



(j) FREAK (2/48) (k) ShapeD (2/51) (l) LIOP (4/48)

Fig. 5: Matching performance evaluation for 3D objects (*potato* and *volley ball* dataset) measured by (number of inlier / number of outlier).

- The slowest descriptor is LIOP. If we reduce the patch size to half, the computational time of LIOP descriptor decrease to 3,344 ms but it is still slow.
- Because SIFT, SURF, and Shape descriptor have same dimension of descriptor vector, the matching time is almost similar.
- The matching with LIOP descriptor is two times slower than SIFT because LIOP has twice size of descriptor than SIFT.

5 Conclusion

In this paper, we evaluated the performance of several local descriptors as well as binary descriptors for the affine invariant region detector, *i.e.* MESR. A total

Table 2: Computation time in milliseconds (500×500 matching).

Descriptor (dimension)	Descriptor generation	Nearest neighbor matching
SIFT (128)	2,189	142
SURF (128)	388	157
BRIEF (256)	43	84
FREAK (512)	84	163
Shape descriptor (128)	2,888	112
LIOP (255)	42,827	267

of six (SIFT, SURF, BRIEF, FREAK, Shape descriptor, and LIOP) descriptors were tested in different geometric transforms including large viewpoint changes, image zoom and rotation, deformable objects, and 3D objects. Under the evaluation framework, LIOP outperformed the other descriptors in image zoom and rotation, and viewpoint change. The binary descriptors archived the fastest description and matching with comparable performance with patch based descriptor. The experimental result indicated that MSER detector was not suitable for describing deformable object and 3D object.

Acknowledgement. This work was supported by the IT R&D program of MSIP/ KEIT. [10047078, 3D reconstruction technology development for scene of car accident using multi view black box image].

References

1. Karlsson, N., Bernardo, E.D., Ostrowski, J., Goncalves, L., Pirianian, P., Munich, M.E.: The vSLAM algorithm for robust localization and mapping. Proc. of IEEE International Conference on Robotics and Automation (2005) 24–29
2. Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S.M., Szeliski, R.: Building Rome in a day. Communications of the ACM **54** (2011) 105–112
3. Lowe, D.G.: Object recognition from local scale-invariant features. Proc. of IEEE International Conference on Computer Vision **2** (1999) 1150–1157
4. Zhou, H., Yuan, Y., Shi, C.: Object tracking using SIFT features and mean shift. Computer Vision and Image Understanding **113** (2009) 345–352
5. Serrano, N., Savakis, A.E., Luo, J.: Improved scene classification using efficient low-level features and semantic cues. Pattern Recognition **37** (2004) 1773–1784
6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60** (2004) 91–110
7. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features. Computer Vision and Image Understanding **110** (2008) 346–359
8. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. on Pattern Analysis and Machine Intelligence **27** (2005) 1615–1630
9. Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. International Journal of Computer Vision **60** (2004) 63–86

10. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Gool, L.V.: A comparison of affine region detectors. International Journal of Computer Vision **65** (2005) 43–72
11. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. Proc. of the British Machine Vision Conference **1** (2002) 384–393
12. Forssen, P.E., Lowe, D.G.: Shape descriptors for maximally stable extremal regions. Proc. of IEEE International Conference on Computer Vision (2007) 1–8
13. Calonder, M., Lepetit, V., Strecha, C., Fua, P.: BRIEF: Binary robust independent elementary features. Proc. of European Conference on Computer Vision (2010) 778–792
14. Alahi, A., Ortiz, R., Vandergheynst, P.: FREAK: Fast retina keypoint. Proc. of IEEE Conference on Computer Vision and Pattern Recognition (2012) 510–517
15. Wang, Z., Fan, B., Wu, F.: Local intensity order pattern for feature description. Proc. of IEEE International Conference on Computer Vision (2011) 603–610
16. Miksik, O., Mikolajczyk, K.: Evaluation of local detectors and descriptors for fast feature matching. Proc. of International Conference on Pattern Recognition (2012) 2681–2684
17. Restrepo, M.I., Mundy, J.L.: An evaluation of local shape descriptors in probabilistic volumetric scenes. Proc. of the British Machine Vision Conference (2012) 46.1–46.11
18. Moreels, P., Perona, P.: Evaluation of features detectors and descriptors based on 3D objects. International Journal of Computer Vision **73** (2007) 263–284
19. Dahl, A.L., Aanaes, H., Pedersen, K.S.: Finding the best feature detector-descriptor combination. Proc. of International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (2011) 318–325
20. Haja, A., Jahne, B., Abraham, S.: Localization accuracy of region detectors. Proc. of IEEE Conference on Computer Vision and Pattern Recognition (2008) 1–8
21. Dickscheid, T., Schindler, F., Forstner, W.: Coding images with local features. International Journal of Computer Vision **94** (2011) 154–174
22. Canclini, A., Cesana, M., Redondi, A., Tagliasacchi, M., Ascenso, J., Cilla, R.: Evaluation of low-complexity visual feature detectors and descriptors. Proc. of International Conference on Digital Signal Processing (2013) 1–7
23. Salzmann, M., Moreno-Noguer, F., Lepetit, V., Fua, P.: Closed-form solution to non-rigid 3D surface registration. Proc. of European Conference on Computer Vision (2008) 581–594