Supplementary material for paper *Tree-Based Morse Regions: A Topological Approach to Local Feature Detection*

Yongchao Xu, Pascal Monasse, Thierry Géraud, and Laurent Najman

I. Introduction

In this supplementary material document, we show some complementary experimental results for the paper *Tree-Based Morse Regions: A Topological Approach to Local Feature Detection*. They are listed as bellow:

- In Section II, we show the result of the repeatability test (including DoG) on "Graffiti" sequence.
- In Section III, we show the comparison of features (along with the scales) extracted by different methods on a low contrast image and its contrast-enhanced one. This pair of images are shown in Fig. 6 of the paper.
- In Section IV, we show the location of detected points of images of CD-covers used in Fig. 11 and Fig. 13 These points are used in image registration experiments. The pairs of matched points used to estimate the homography are also illustrated.
- In Section V, we show some supplementary results of quantitative benchmark of registration experiments on Mikolajczyk dataset.

II. Repeatability test on "Graffiti" sequence

We include the repeatability test on the "Graffiti" sequence in Mikolajczyk's dataset. First of all, the number of detected points on each image of the sequence is given in Table I. TBMR extracts more points than MSER, but slightly less than Harris-Affine and Hessian-Affine, and DoG detects many more points than the others. The result of repeatability test (including DoG for which the round disks are used instead of squares) is shown in Fig. 1. The same observation is obtained as for the "Wall" sequence (with also the viewpoint change) shown in Fig. 9 in the paper. For the "Graffiti" sequence, in general, TBMR achieves better repeatability score than Harris-Affine and Hessian-Affine. Compared to MSER, TBMR obtains many more correspondences with a small loss of repeatability score. Note that we use the Linux executable available in http://www.robots.ox.ac.uk/~vgg/research/affine/detectors.html for the MSER. The public implementation of MSER in VLFeat [36] http:/www.vlfeat.org/ and OpenCV achieves results worse than the ones of the binary. DoG detects many points with a good repeatability score for small change of viewpoint. When the viewpoint change is large, the repeatability score decreases significantly, and from 50 degrees of viewpoint change. DoG does not give any correspondence. This explains the better performance in the homography experiments shown in Fig. 10 in the paper.

Methods	img1	img2	img3	img4	img5	img6
TBMR	1200	1285	1384	1588	1670	1886
MSER	547	633	700	707	798	925
Harris-Affine	1753	2059	2215	2065	2195	1888
Hessian-Affine	2510	2848	2782	2451	2385	1898
DoG	3198	3586	3911	3952	4299	4920

TABLE I. Number of points detected by different methods for each image of the "Graffiti" sequence. TBMR extracts more points than MSER, less than Harris-Affine and Hessian-Affine. DoG extracts many more points than the other methods.

Methods	Original image	contrast-enhanced image
TBMR	2978	2347
MSER	1025	1019
Harris-Affine	3119	1874
Hessian-Affine	2204	1474
DoG	7042	6768

TABLE II. Number of points detected by different methods on the low contrast image and its contrast-enhanced one. TBMR extracts more points than MSER, and comparable number of points with Harris-Affine and Hessian-Affine. DoG extracts many more points than the other methods.

III. Comparison of extracted local features on a low contrast image

We show the detected points with scales for the low contrast image and its contrast-enhanced one (as shown in Fig. 6 (a) of the paper). This original low contrast image and a contrast-enhanced one with significant increasing change of contrast are also shown in Fig. 2 (a). The number of extracted points by different methods are shown in Table II. The extracted points with scales of different methods are respectively shown in Fig. 2 (b), Fig. 3, and Fig. 4. As shown in Fig. 2, the points in mostly uniform regions shown in Fig. 5 (f) of the paper are actually points with a different large scale (the ellipses). MSER, DoG, Harris-Affine, and Hessian-Affine detect very few points in the area of low contrast (e.g., the body of the deer sculpture). By increasing the contrast, they detect some points. TBMR is perfectly insensitive to contrast change, up to quantization effects.

IV. Local feature detection for two images of CD-covers

We show in Fig. 5 and Fig. 6 the locations of the interest points extracted by TBMR and DoG for the two CD-covers images: 007 and 010 used respectively in Fig. 11 and Fig. 13 of the paper. The number of interest points detected by these two methods is shown in Table III. For the reference images, TBMR detects fewer points than DoG. For the target images

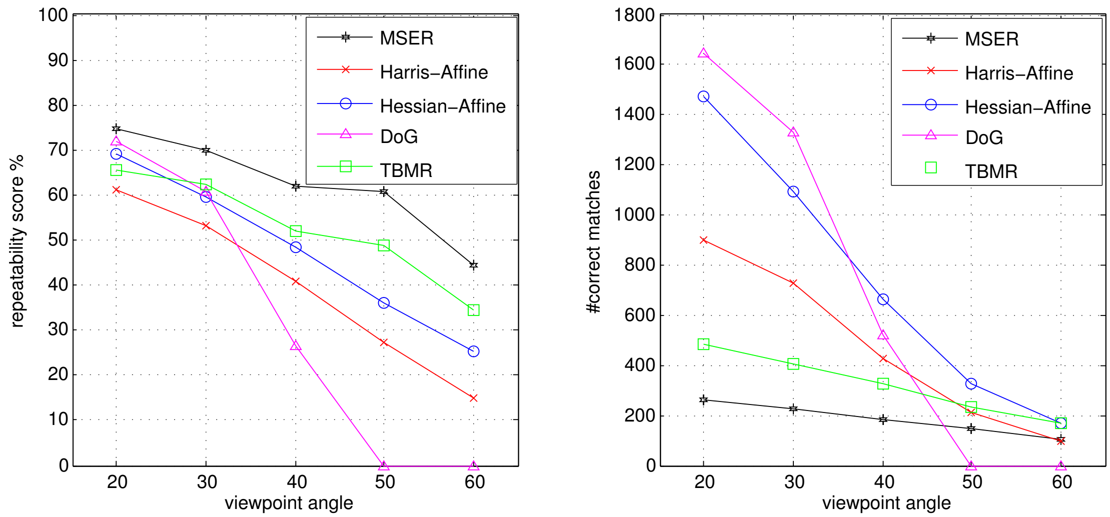


Fig. 1. Repeatability score (left) and number of correspondences (right) for the “Graffiti” sequence. TBMR is more robust with respect to viewpoint change, especially for strong viewpoint change.

Methods	Reference 007	Canon 007	Reference 010	Palm 010
TBMR	262	746	475	1581
DoG	1038	132	1017	537

TABLE III. Number of interest points detected by TBMR and DoG on two pairs of CD-covers images used respectively in Fig. 11 and Fig. 13.

taken with Canon and Palm cameras, TBMR extracts more points, but many of them lie in the area outside the cd covers. When we compute the homography between the two pairs of images, only the points extracted in reference images (inside the cd covers) are used. The registration based on TBMR makes use of fewer points than DoG. For the pair of images 007 of the CD-covers, both TBMR and DoG gives a registration result. The inlier matched pairs of points are given in Fig. 7. In fact, all the matched pairs of points given by TBMR are used as inliers to estimate the homography. One pair of matched points given by DoG is considered as outlier, which is not shown here. For the pair of images 010 of the CD-covers, all the matched pairs of points given by TBMR (shown in Fig. 8) are used as inliers to estimate the homography. DoG fails to estimate the homography, all the matched pairs of points are considered as outliers (see Fig. 8). Actually, four of the six DoG correspondences look correct. Although this is exactly the bare minimum required for a homography, there is no correct other correspondence to check its adequacy.

V. Image registration on Mikolajczyk dataset

We show some extra distributions of distances in image registration by homography on the dataset of Mikolajczyk *et al.* [6]. We compare mainly to MSER and DoG.

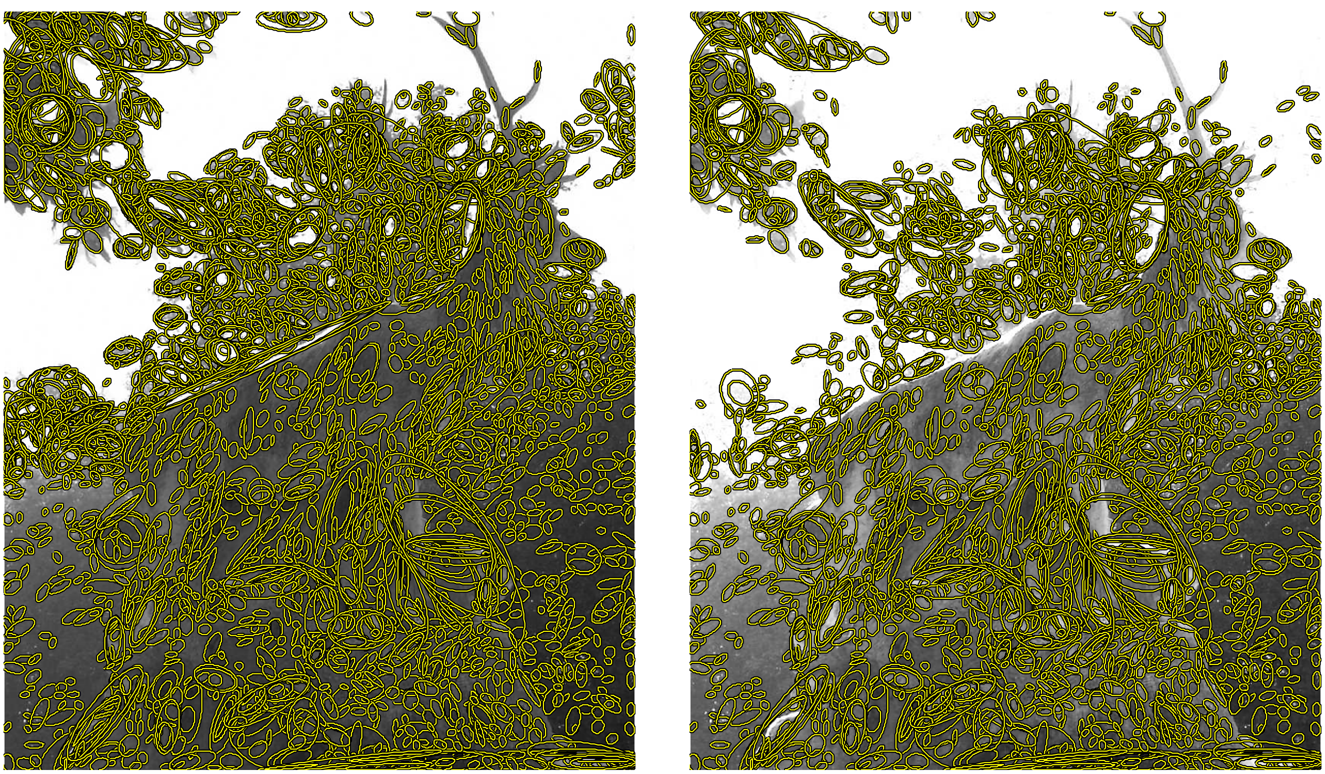
The results for the “Wall” sequence with viewpoint changes are shown in Fig. 9. Harris-Affine, Hessian-Affine, and DoG fail registering pairs (1,6); TBMR performs similar with MSER on all the pairs and similar with Harris-Affine, and Hessian on all other pairs than (1,6). In general, TBMR performs better than DoG on all the other pairs.

The results for the “Bark” sequence with scale changes are shown in Fig. 10. TBMR performs better than MSER for all the pairs. Except for the pair (1,3) where TBMR is on par with DoG, TBMR performs better than DoG. Harris-Affine fails registering pair (1,6). In general, TBMR performs better than Hessian-Affine.

The results for the “Trees” sequence with different blur effects are shown in Fig. 11. TBMR performs better than MSER for all the pairs. Besides, MSER fails registering the pair (1,6). In general, TBMR also performs better than Harris-Affine and Hessian-Affine. When the blur is not very important (the pairs 1-2, 1-3, and 1-4), TBMR achieves similar results with DoG. But when the blur is very important, DoG performs better. This also explains the better performance of DoG in the registration experiments for “Palm” in Table I. However, as shown in [45], the multi-resolution detection improves the performance of MSER under blur. We would expect the same improvements by applying a multi-resolution analysis.

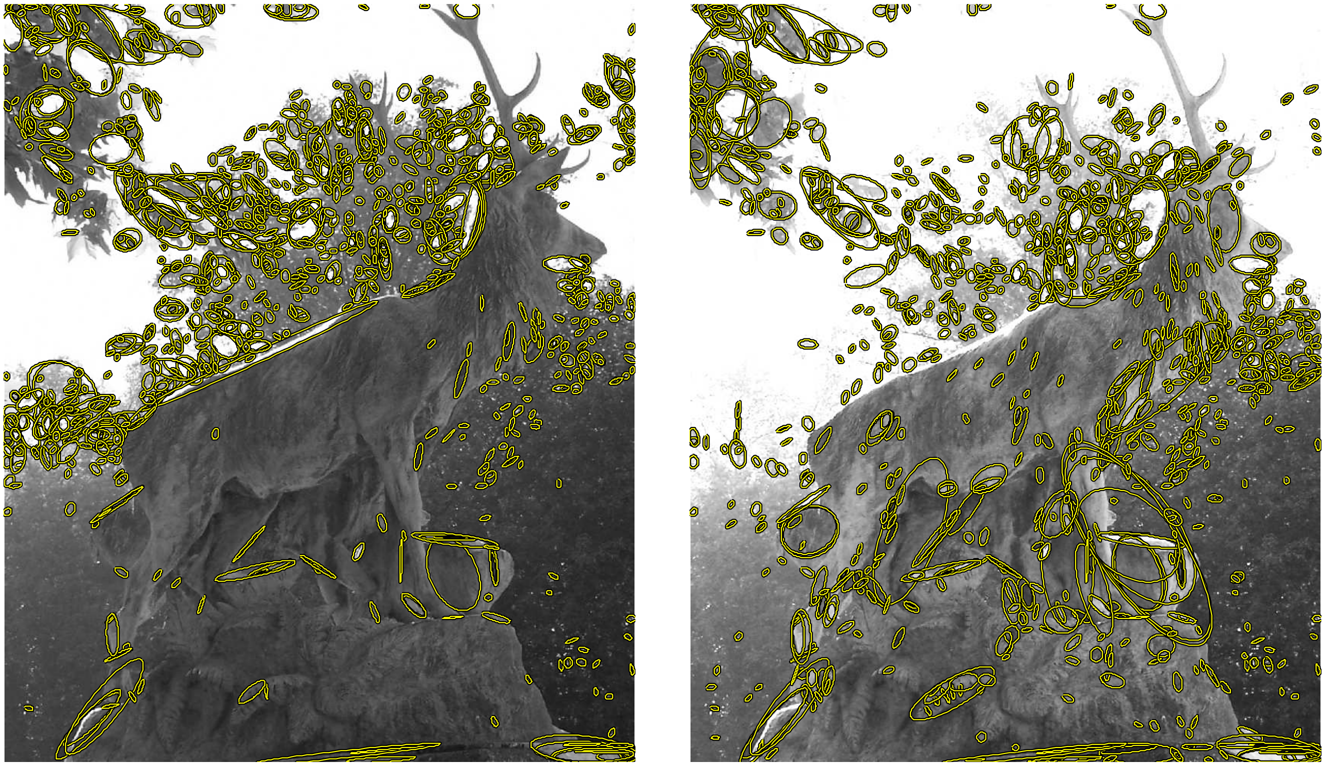
The results for the “Leuven” sequence with different blur amounts are shown in Fig. 12. In general, TBMR performs better than MSER, Harris-Affine, and Hessian-Affine. However, DoG performs better than TBMR: As all pairs of images cover almost the same scene and there are always some parts of the image having good contrast, even for the last image (img6) of the sequence, the change of contrast does not have a strong impact on the homography estimation for DoG.

(a) Original low contrast image (left) and contrast-enhanced image (right).

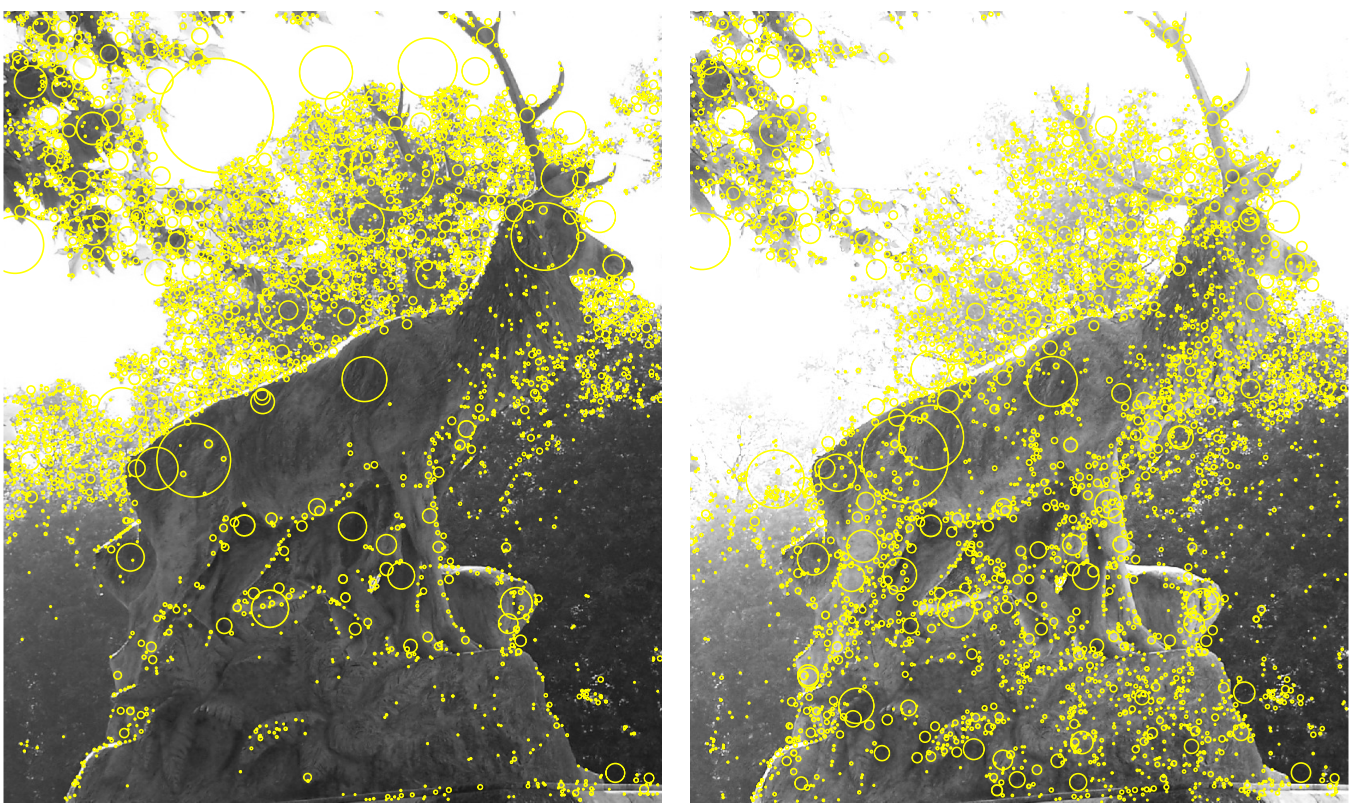


(b) Local features extracted by TBMR on original images (left) and contrast-enhanced image (right).

Fig. 2. Local features extracted by TBMR on low contrast image and contrast-enhanced image. The points in mostly uniform regions shown in Fig. 5 (f) of the paper are actually points with a different large scale (the ellipses). TBMRs are perfectly insensitive to contrast change, up to quantization effects.

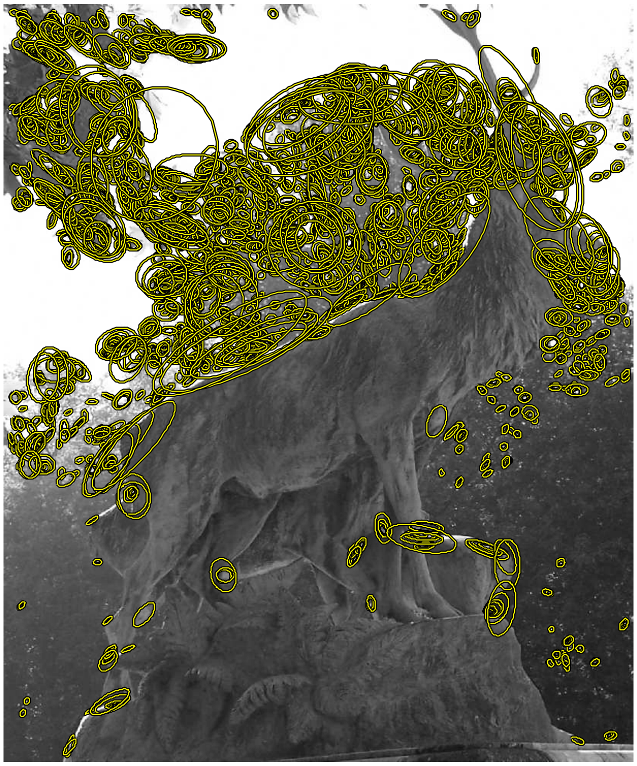
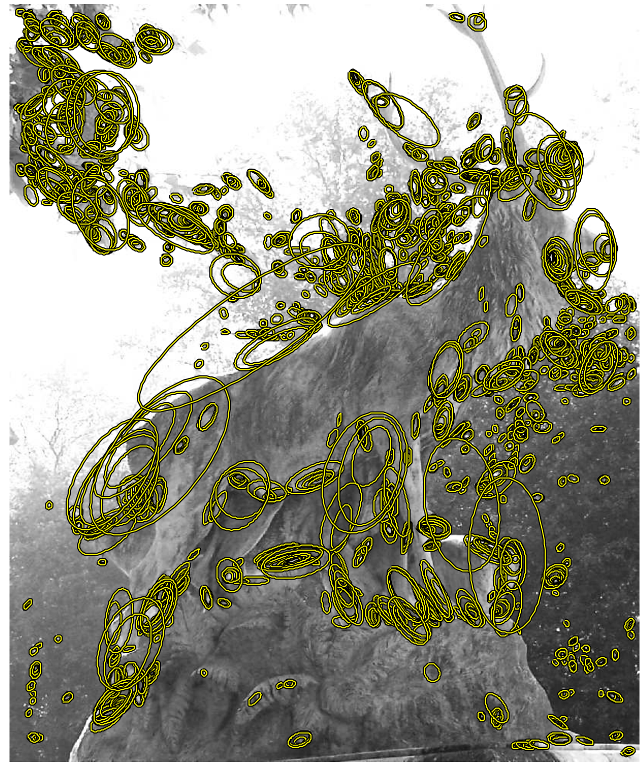


(a) Local features extracted by MSER on original images (left) and contrast-enhanced image (right).

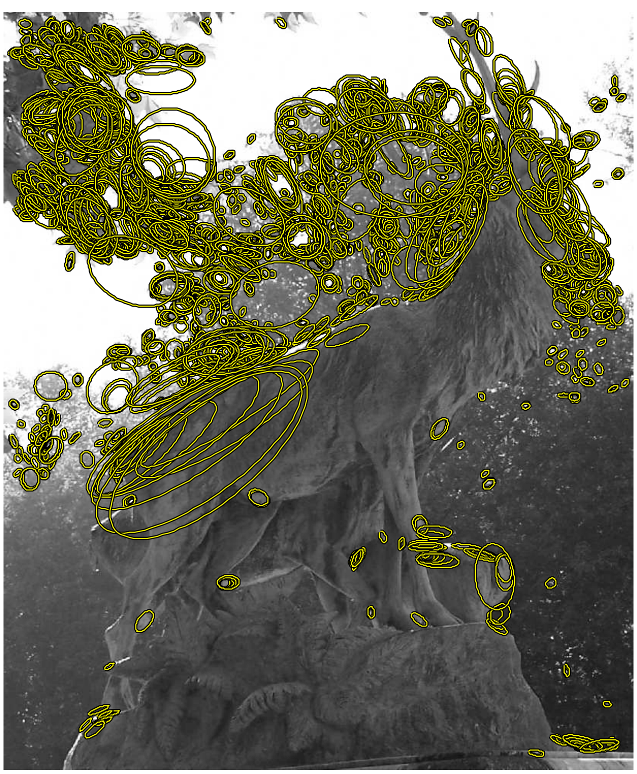
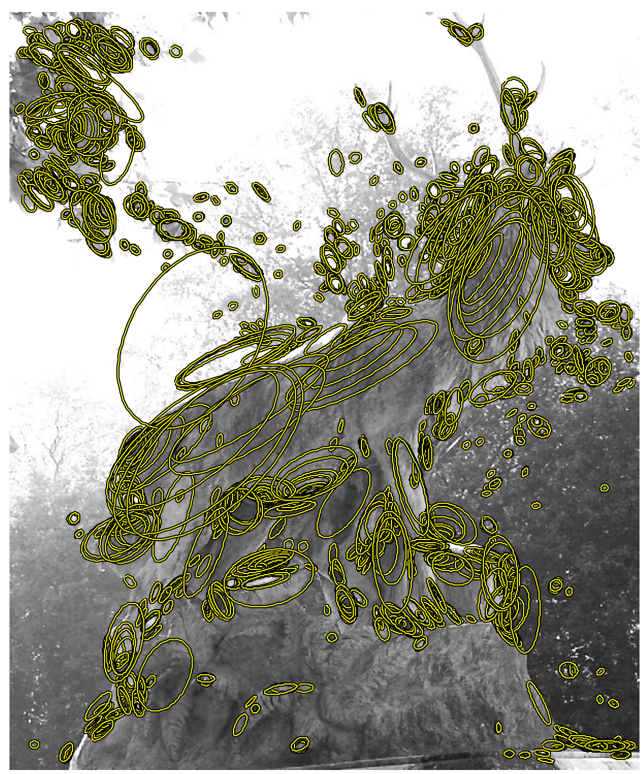


(b) Local features extracted by DoG on original images (left) and contrast-enhanced image (right).

Fig. 3. Local features extracted by MSER (top) and DoG (Down) on low contrast image and contrast-enhanced image. To better visualize the patches extracted by DoG, the round disks with diameter being the size of square patches divided by 6 are shown. MSER and DoG detect very few points in the area of low contrast (e.g., the body of the deer sculpture). By increasing the contrast, MSER and DoG detect some points.

(a) Local features extracted by Harris-Affine on original images (left) and contrast-enhanced image (right).

(b) Local features extracted by Hessian-Affine on original images (left) and contrast-enhanced image (right).

Fig. 4. Local features extracted by Harris-Affine (top) and Hessian-Affine (Down) on low contrast image and contrast-enhanced image. They both detect very few points in the area of low contrast (e.g., the body of the deer sculpture). By increasing the contrast, they detect some points.

Fig. 5. Locations of extracted interest points detected by TBMR (left) and DoG (right) on the two reference images of CD-covers. TBMR detects fewer points than DoG.

Fig. 6. Locations of extracted interest points detected by TBMR (left) and DoG (right) on the two target images taken by respectively Canon (top) and Palm (down) cameras of CD-covers. TBMR detects more points than DoG. Many of them lie in the area outside of the cd covers. In the area of low contrast, TBMR extracts more points than DoG.

Fig. 7. Inlier matched pairs of points using respectively TBMR (top) and DoG (down). TBMR has more inlier matched pairs of points than DoG. In fact, all the matched pairs of points using TBMR are considered as the inliers to estimate the homography. One pair of matched points using DoG is considered as outliers in the homography estimation.

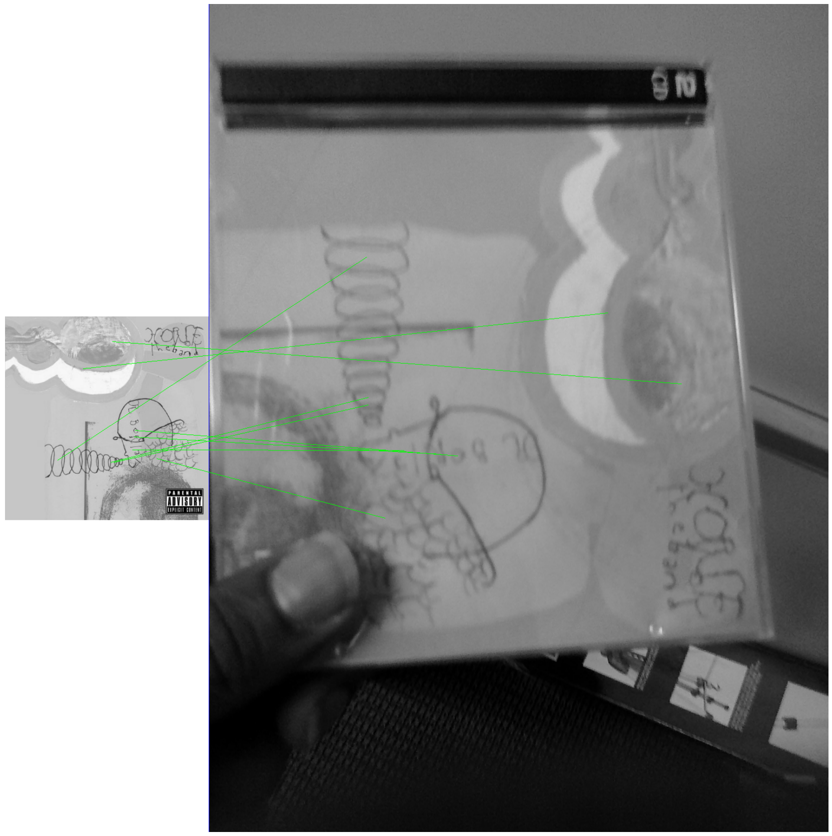
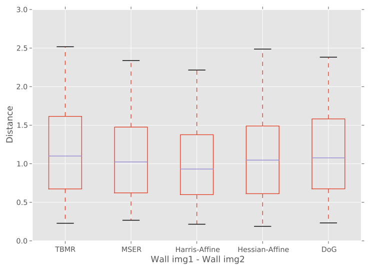
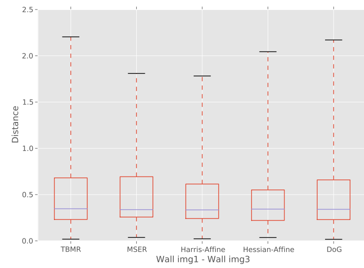


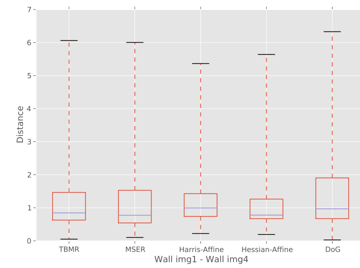
Fig. 8. Matched pairs of points using respectively TBMR (top) and DoG (down). All the matched pairs of points using TBMR are used as inliers to estimate the homography. Using DoG fails to estimate the homography: all the matched pairs of points are considered as outliers. Though four pairs are correct, there is no other to check automatically the resulting homography.



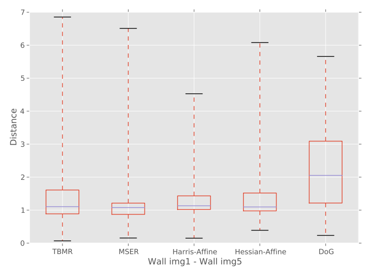
(a) Wall img1 - img2.



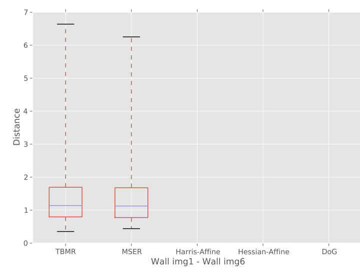
(b) Wall img1 - img3.



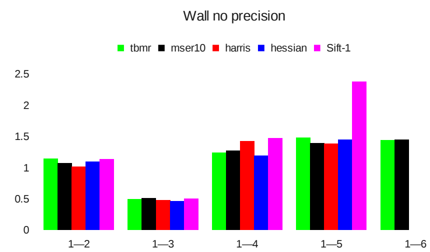
(c) Wall img1 - img4.



(d) Wall img1 - img5.

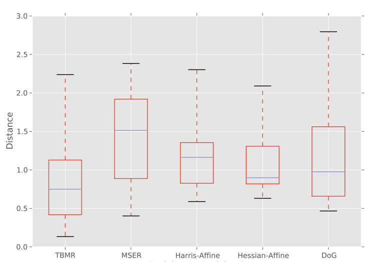


(e) Wall img1 - img6.

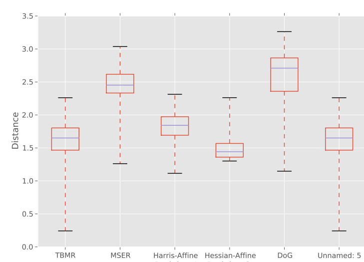


(f) Average of all pairs.

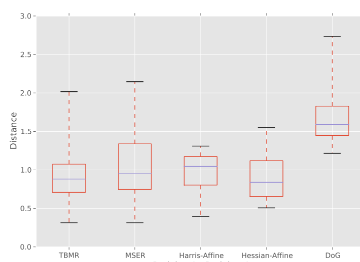
Fig. 9. Distributions of errors in image registration by homography on "Wall" image pairs, which comes with a ground truth. Harris-Affine, Hessian-Affine, and DoG fails registering pairs (1,6); TBMR performs similar with MSER on all the pairs and similar with Harris-Affine, and Hessian on all other pairs than (1,6). In general, TBMR performs better than DoG on all the other pairs.



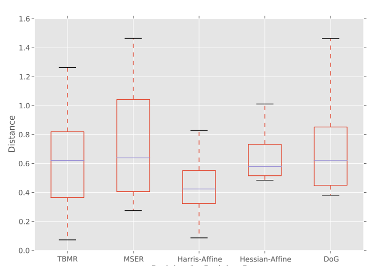
(a) Bark img1 - img2.



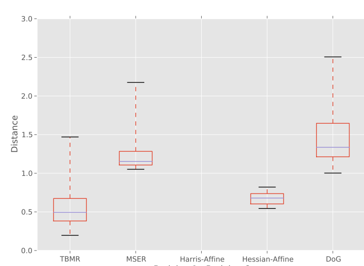
(b) Bark img1 - img3.



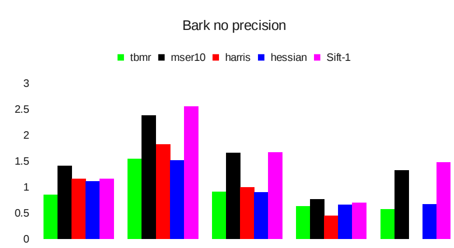
(c) Bark img1 - img4.



(d) Bark img1 - img5.

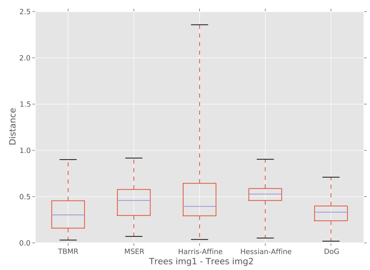


(e) Bark img1 - img6.

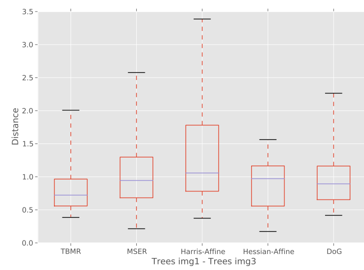


(f) Average of all pairs.

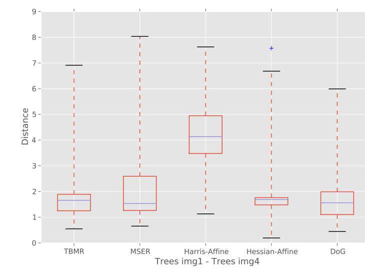
Fig. 10. Distributions of errors in image registration by homography on "Bark" image pairs, which comes with a ground truth. TBMR performs better than MSER for all the pairs. Except for the pair (1,3) where TBMR similar with DoG, TBMR performs better than DoG. Harris-Affine fails registering pair (1,6). And in general, TBMR performs better than Hessian-Affine.



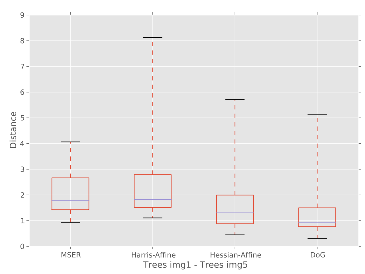
(a) Trees img1 - img2.



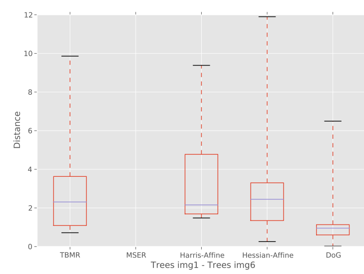
(b) Trees img1 - img3.



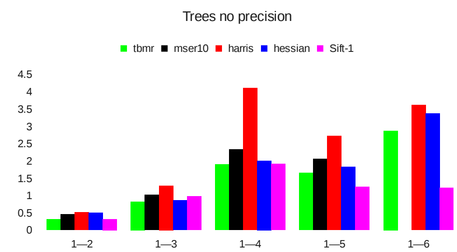
(c) Trees img1 - img4.



(d) Trees img1 - img5.

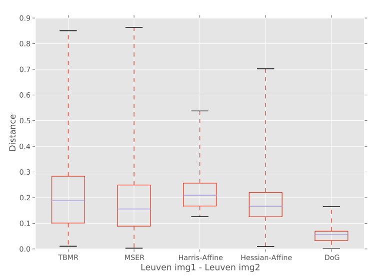


(e) Trees img1 - img6.

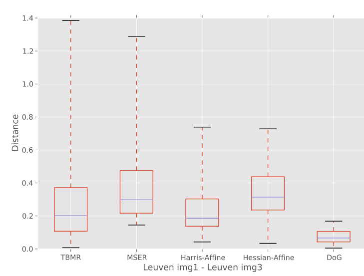


(f) Average of all pairs.

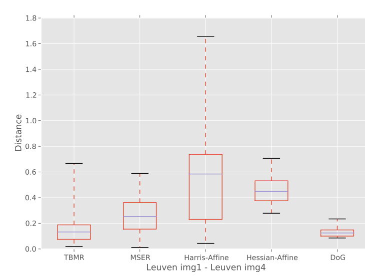
Fig. 11. Distributions of errors in image registration by homography on “Trees” image pairs, which comes with a ground truth. TBMR performs better than MSER for all the pairs. Besides, MSER fails registering the pair (1,6). In general, TBMR also performs better than Harris-Affine and Hessian-Affine. When the blur is not very important (the pairs 1-2, 1-3, and 1-4), TBMR achieves similar results with DoG. But when the blur is very important, DoG performs better.



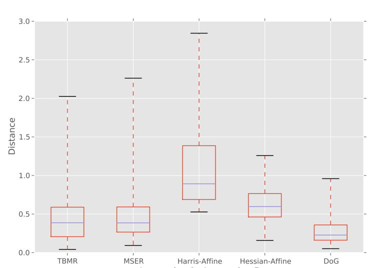
(a) Leuven img1 - img2.



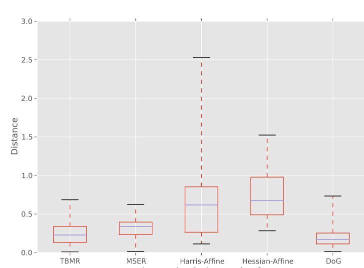
(b) Leuven img1 - img3.



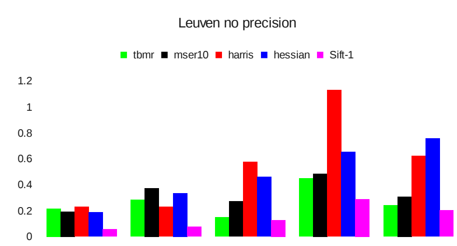
(c) Leuven img1 - img4.



(d) Leuven img1 - img5.



(e) Leuven img1 - img6.



(f) Average of all pairs.

Fig. 12. Distributions of errors in image registration by homography on “Leuven” image pairs, which comes with a ground truth. In general, TBMR performs better than MSER, Harris-Affine, and Hessian-Affine. However, DoG performs better than TBMR: As all pairs of images cover almost the same scene and there are always some parts of the image having good contrast, even for the last image (img6), the change of contrast does not have a strong impact on the homography estimation for DoG.