Negative Determinant of Hessian Features

Ruan Lakemond, Clinton Fookes, Sridha Sridharan
Image and Video Research Laboratory, Queensland University of Technology
GPO Box 2434, 2 George St, Brisbane, Queensland 4001
{*r.lakemond, c.fookes, s.sridharan*}@qut.edu.au

***Abstract*—Local image feature extractors that select local maxima of the determinant of Hessian function have been shown to perform well and are widely used. This paper introduces the negative local minima of the determinant of Hessian function for local feature extraction. The properties and scale-space behaviour of these features are examined and found to be desirable for feature extraction. It is shown how this new feature type can be implemented along with the existing local maxima approach at negligible extra processing cost.**

Applications to affine covariant feature extraction and sub-pixel precise corner extraction are demonstrated. Experimental results indicate that the new corner detector is more robust to image blur and noise than existing methods. It is also accurate for a broader range of corner geometries.

An affine covariant feature extractor is implemented by combining the minima of the determinant of Hessian with existing scale and shape adaptation methods. This extractor can be implemented along side the existing Hessian maxima extractor simply by finding both minima and maxima during the initial extraction stage. The minima features increase the number of correspondences by two to four fold. The additional minima features are very distinct from the maxima features in descriptor space and do not make the matching process more ambiguous.

I. Introduction

Local image features are patterns in an image that are defined in limited image areas and are distinguishable from the surrounding image in some way. Such features may be extracted from each view of a scene independently and then matched to find sets of correspondences between views. The correspondences are commonly used for a large variety of tasks, including automatic camera calibration [1], [2], 3D reconstruction [3], [4], mosaicing [5], object recognition and classification [6] and arranging image databases [7]. A comprehensive review of local image feature extractors may be found in [8].

Feature extractors based on the determinant of Hessian function [9] have been shown to perform well [10], [11] and are widely used. The determinant of Hessian function is used to compute a saliency map, which yields a strong response in image regions of high curvature. Features are extracted by finding the local maxima in the saliency map.

This paper demonstrates that the negative minima of the determinant of Hessian function are reliable local features. They can be extracted in parallel to the maxima with negligible additional image processing cost. In some cases the minima features perform significantly better than the maxima features. In all test cases, combining both methods produced more than double the correspondences of using Hessian maxima alone while slightly improving repeatability and matching scores. The minima features are shown to be very effective for corner detection. A method is demonstrated that can extract corner and junction vertex locations to sub-pixel accuracy, even when corners are blurred or rounded.

II. Background

Saliency map-based local image feature extractors compute a saliency map from an image using a function of local gradients (for example [9], [12]) or local information content [13]. The local maxima of the saliency map are used to localise potential features. A positive threshold is typically applied to suppress maxima generated by noise.

This paper focuses on the determinant of Hessian function as saliency operator. The Hessian of an image is defined as the matrix of second order partial derivatives of the image intensity with respect to coordinates,

$$\frac{\partial^2 I(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^\top} = \begin{bmatrix} \frac{\partial^2}{\partial x^2} & \frac{\partial^2}{\partial xy} \\ \frac{\partial^2}{\partial xy} & \frac{\partial^2}{\partial y^2} \end{bmatrix} I(\mathbf{x}). \tag{1}$$

The determinant of this matrix gives large positive values where its eigenvalues are large and have the same sign. This occurs in the presence of strong edges at multiple orientations, such as corners. Its response is strongest in the presence of opposing edges and it is therefore a good blob detector.

The saliency map approach gives the locations of features in an image, but is not robust to changes in viewpoint or scale. Where features need to be robust to changes in scale, a multi-scale approach is used to select an appropriate scale for each feature [14]–[16]. Affine adaptation [16]–[18] can be used to estimate the local shape of features and make them more robust to changes in viewpoint.

III. Negative Minima of the Determinant of Hessian Function

The determinant of Hessian yields a negative response when its eigenvalues are of opposite sign, as in the case of a saddle point. A strong negative response indicates the presence of multiple edges, just like a strong positive response. It therefore stands to reason that the local minima of the determinant of Hessian may yield good features just like the maxima. This section examines the characteristics of these features.

The top row of Figure 1 shows examples of multi-scale determinant of Hessian features (both maxima and minima) generated by simple structures. Determinant of Hessian minima have been observed to have the following properties:

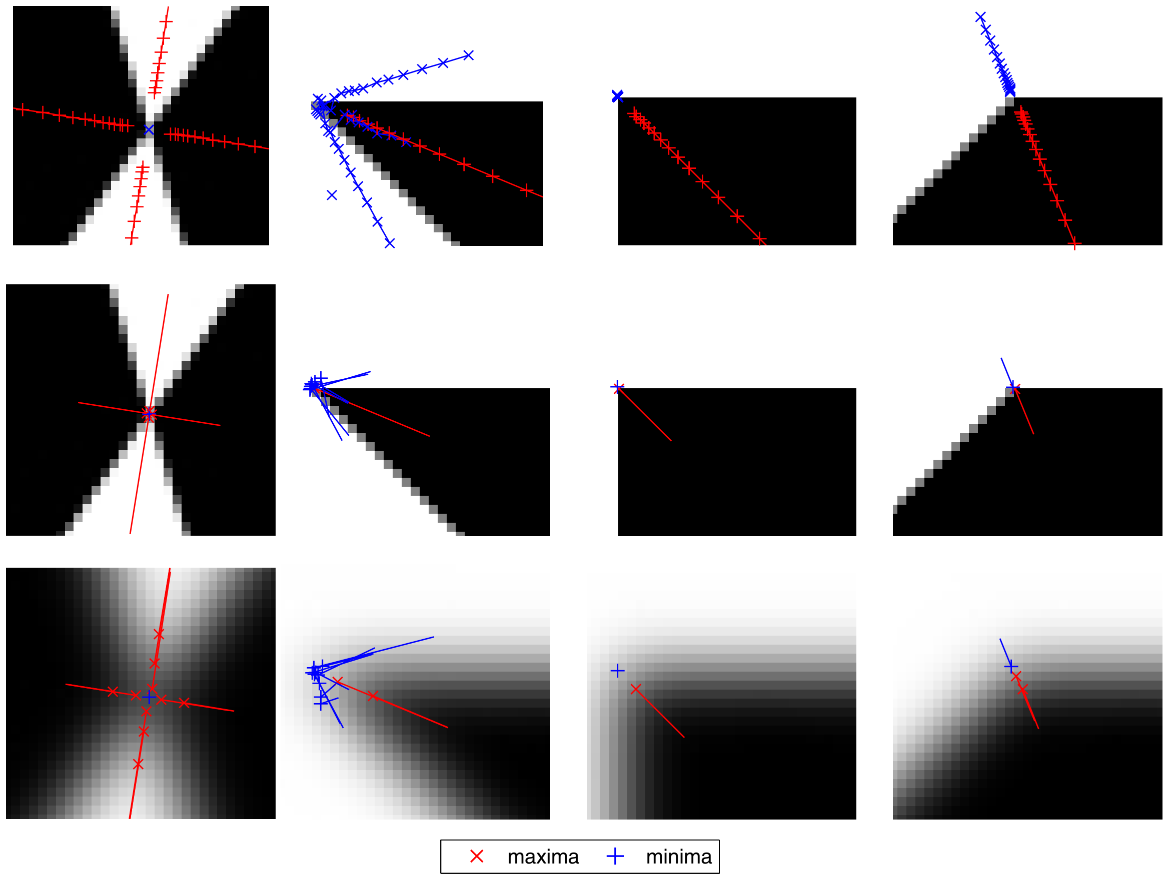

Fig. 1. Example corner structures and resulting determinant of Hessian features. The first row shows multi-scale minimal (+) and maximal (×) features. The lines joining features indicate the scale space locus. Row two and three show the corner locations estimated by fitting lines to the multi-scale minimal and maximal feature loci. The lines extending from the corner centres indicate the directions of the lines fit to the multi-scale loci. In the third row the images were blurred by Gaussian convolution to show the effect of image blur on the corner estimates.

1) Location: The minima occur at saddle points, which occur at the intersection of edges in the image.

2) Scale Behaviour: Minima features drift linearly away from corners as scale increases. The drift direction and speed depends on the corner geometry (see Figure 1) – obtuse angles result in a drift towards the outside of the corner; acute angles result in a divergence and drift on either side of the corner; right angles and opposing corners result in the minimum remaining stationary over a large range of scales. Minima features can diverge as scale increases, but cannot converge. This behaviour is the opposite of the maxima features.

3) Characteristic Scale: A characteristic scale is typically not well defined. The Laplacian is typically of small magnitude and unstable, while the determinant of Hessian function does not yield a stable and repeatable minimum over scales. Neither of these functions provide a reliable method for scale selection.

4) Characteristic Shape: The local shape of structures detected by the minima can be ambiguous. All the structures in Figure 1, for example, can be normalised such that their edges are perpendicular, but there remains a relative scaling ambiguity along the direction of each edge.

While the minima of the determinant of Hessian are not ideal features, they do have the potential to be useful in practice. Two potential applications were investigated. In Section IV a corner detector based on Hessian minima is presented and tested using a calibration experiment. In Section V, Hessian minima features are used to construct affine covariant features using the same methods used for maxima. The affine features are evaluated against a standard database.

IV. Corner Detection

From Figure 1 it can be seen that the minima of the determinant of Hessian function are potentially more suitable for finding the vertices of corners than the maxima. For some corners the minima features are located exactly on the corner vertex over a range of scales, but for others they diverge linearly away from the vertex over increasing scale. A method is proposed here that attempts to find the true corner vertex from a multi-scale analysis of determinant of Hessian features.

Determinant of Hessian features (and most other gradient-based features) move away from a corner's vertex as scale increases. The corner vertex position can be recovered by fitting a straight line to the feature locus in scale space coordinates and finding the point where this line intersects the scale zero plane. Row two and three of Figure 1 shows the corner positions estimated using this process on the images and features, as well as on blurred images. Multiple corner estimates are produced where the feature locus is not straight and multiple line segments are fitted (as with heavily blurred images) or where the features diverge in multiple directions (as with minimal features in the presence of sharp corners).

In the sharp, high resolution images, the both feature types give highly accurate estimates of the corner vertex position. Where the images are blurred, the corner estimates produced by the maxima features are significantly biased towards the inside of the corner, while the minima features produce some accurate and some outlying estimates. The line fit to maxima features gives a good estimate of the corner bisector line in all cases. A better estimate of the corner position may be found by taking the average position of the estimates produced by the minimal features that are in close proximity to the bisector line produced by maximal features. This approach rejects outlier corner estimates produced by minima features and corrects for the drift of maxima corner estimates.

A. Evaluation

The proposed corner extractor is compared by means of a calibration experiment to the popular Harris corner extractor and to the iterative corner refinement algorithm implemented in OpenCV *cv::cornerSubPix* [19]. The test data consists of images of two calibration patterns acquired by six different cameras. One pattern is a checkerboard printed on paper. The second pattern is a grid of squares pattern designed for simultaneous visible and thermal image calibration. It consists of a white thermally insulating mask with square holes placed over a black heated object. This pattern makes the same corners visible in both visible and thermal domain.

Cameras 1 and 2 feature a small aperture and automatic focus adjustment, resulting in consistently sharp image sequences. Cameras 3 and 4 have a narrow depth of field and fixed focus. Parts of the calibration patterns were significantly blurred in some frames. Camera 5 features a wide angle lens with significant distortion, fixed focal distance and could only be used under poor lighting conditions. The sixth camera is a thermal infra-red camera. This camera has severe lens distortion and a particularly shallow depth of field, resulting in severe blurring in much of the image sequence.

The calibration algorithm tracks the calibration pattern through a video sequence. A set of 100 frames are uniformly sampled from each sequence. Corners are detected in each frame using one of the test detectors and the tracking information is used to register the corners to the calibration pattern. The camera is then calibrated using [19] and the reprojection error of the corners is computed.

Table I reports the test results of the squares mask pattern in the form of the mean squared reprojection error and detector miss rate. Table II lists the results for the checkerboard pattern. Miss rates are not reported for the OpenCV method, since it requires reasonably accurate initialisation and does not serve as a raw corner detector itself.

Cam	Harris		OpenCV	Hessian	
	MSE	miss rate	MSE	MSE	miss rate
1	7.952	0.94%	0.453	0.504	0.35%
2	8.882	1.29%	3.612	7.448	7.43%
3	12.997	3.51%	1.603	0.715	1.85%
4	11.037	0.81%	0.906	0.697	0.08%
5	8.671	11.40%	247.177	3.154	6.70%
6	119.460	27.13%	1.304	0.491	16.77%

TABLE I

MEAN SQUARED REPROJECTION ERROR (MSE, IN PIXELS SQUARED) AND DETECTOR MISS RATE FOR CALIBRATION EXPERIMENTS USING THE SQUARES MASK PATTERN. NOTE THAT MISS RATE IS NOT RELEVANT TO THE OPENCV METHOD.

Cam	Harris		OpenCV	Hessian	
	MSE	miss rate	MSE	MSE	miss rate
1	0.184	0.00%	0.082	0.093	0.00%
2	1.763	0.56%	4.162	1.289	2.02%
3	0.238	0.00%	0.223	0.225	0.02%
4	0.371	0.00%	0.133	0.152	0.00%

TABLE II

MEAN SQUARED REPROJECTION ERROR (MSE, IN PIXELS SQUARED) AND DETECTOR MISS RATE FOR CALIBRATION EXPERIMENTS USING THE CHECKER BOARD PATTERN. NOTE THAT MISS RATE IS NOT RELEVANT TO THE OPENCV METHOD.

As expected, the Harris detector consistently performs the worst when using the Squares pattern, due to corner drift. The Hessian-based detector consistently achieves superior results in the more difficult cases (cameras 3 - 6) and only performs marginally worse than the OpenCV method with the clearer image sequences (cameras 1 and 2). The OpenCV method fails in the case of camera 5 due to the interference of shadows cast by the square mask in bad lighting. The Harris detector fails in the case of the thermal camera (camera 6) due to heavy blurring. The Hessian-based detector produces good results in both these difficult cases. There is very little difference between the methods when using a checkerboard pattern, since the Harris and OpenCV methods were specifically designed for this task.

V. AFFINE COVARIANT FEATURES

A. Method

An affine covariant feature extractor can be constructed from Hessian minima features in the same way as for Hessian maxima. An experiment was set up using the evaluation system of [10] to evaluate whether this is effective. Both determinant of Hessian minima and maxima were extracted from each image. The scale selection method of [14], [15] and the affine adaptation method of [18] were used to find affine covariant features. The Hessian minima features are compared against Hessian maxima, and to a combination of both.

The evaluation dataset of [10] consists of image sequences, each featuring a change in viewpoint throughout the sequence. The *bark* and *boat* sequences involve a change in scale (zoom or distance to subject) and rotation. The *graf* and *wall* sequences involve a change in view angle relative to a planar object. The *bikes* and *trees* sequences involve increasing

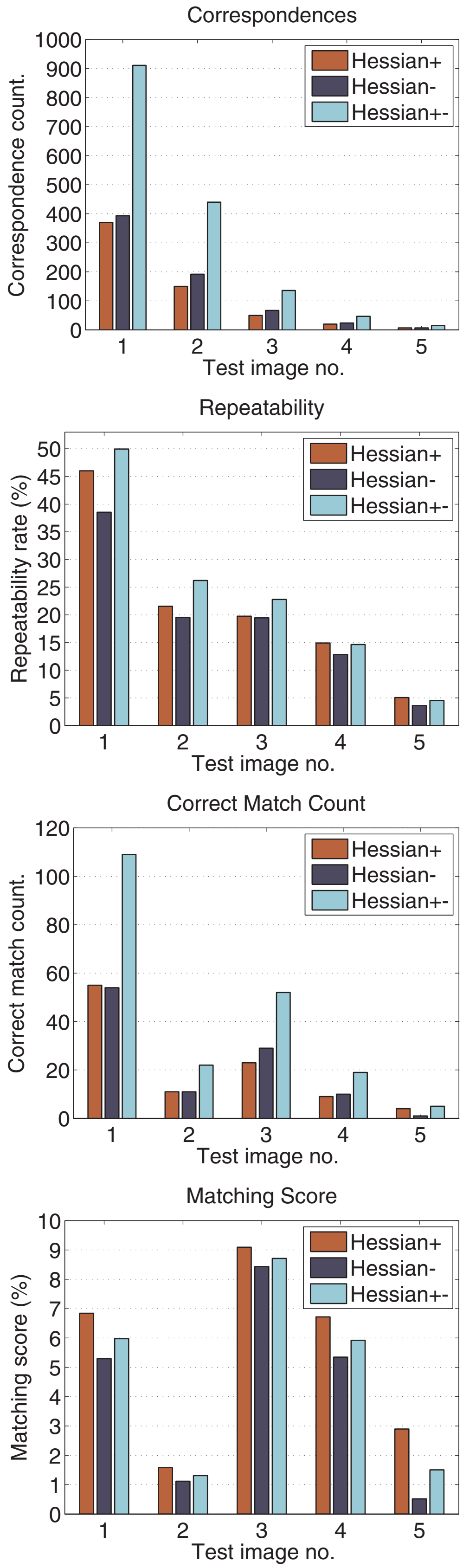

Fig. 2. Evaluation results for *bark* sequence.

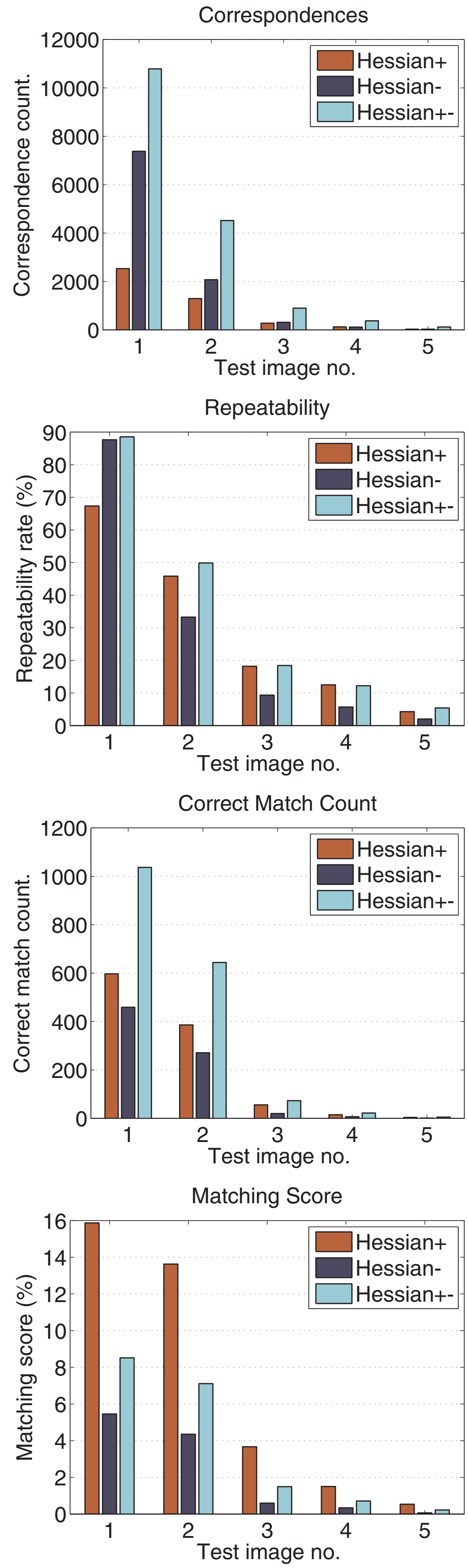

Fig. 3. Evaluation results for *boat* sequence.

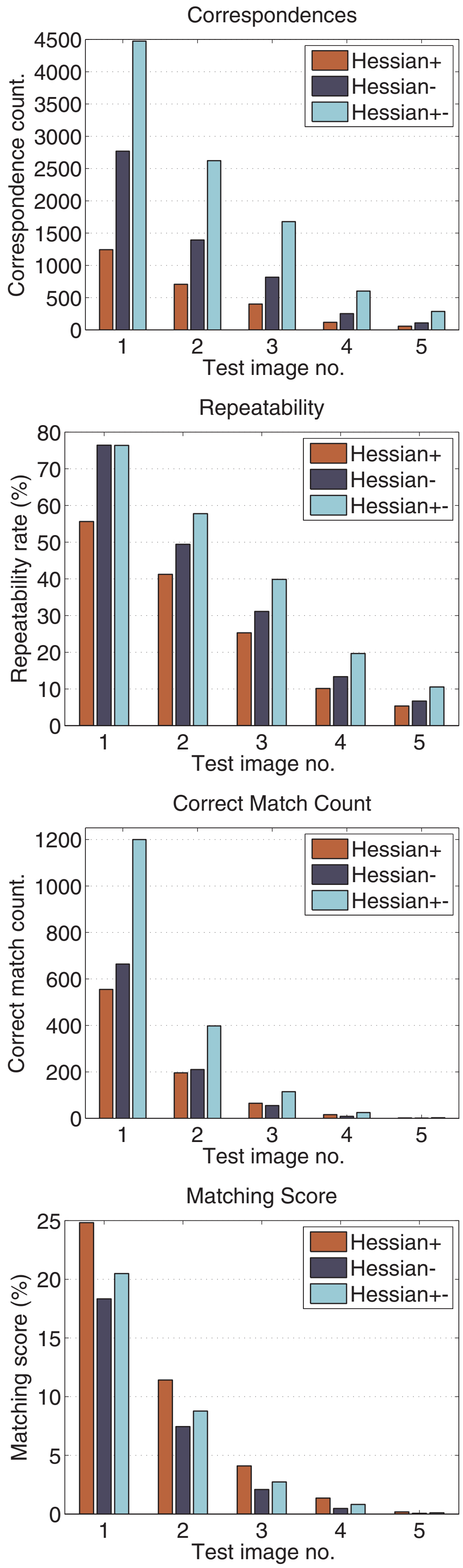

Fig. 4. Evaluation results for *graf* sequence.

image blur or defocusing. Lastly, the *UBC* sequence features JPEG compression and the *Leuven* sequence features lighting variation or camera exposure variation.

Ground truth homographies are provided that map each image to a reference image. The evaluation measures the following metrics: Correspondence count – the number of features that overlap corresponding image regions with sufficiently low error. Repeatability – the proportion of features in the common image area that overlap corresponding image regions with sufficiently low error. Correct match count – the number of matches made using the feature descriptors that are correct. Matching score – the proportion of matches in the common image area that are correct.

B. Results

Figures 2 - 4 show detailed results for selected interesting data sets (*bark*, *boat* and *graf*). The remaining datasets yielded trends consistent with the results presented in detail and are omitted for brevity. The *graf* and *wall* sequences produced very similar results.

The minima features produce a larger number of correspondences than the maxima features, ranging from a relatively small increase, to up to three times more. Combining the two simply yields the sum of both sets of correspondences. The test results show that the combined set of features has more correspondences than the sum of the two. This is due to a shortcoming in the testing procedure, refer to [10] for a discussion on the effects of feature density.

The repeatability of the minima features is significantly higher than the maxima features for the viewpoint change tests (*graf* and *wall*) and slightly lower for the scale change tests (*bark* and *boat*). The increase in repeatability is correlated with the increase in correspondences, indicating that it may be a side effect of the testing procedure and more correspondences, rather than improved performance. Combining both feature types results in a repeatability slightly higher than either alone. Again this is likely related to the increase in correspondences. Overall it appears that the repeatability of the minima features is at least on par with the maxima features.

In most cases the number of correctly matched features is approximately the same for both minima and maxima features. The combined set of features yielded a number of correct matches that is the sum of matches from both sets. This indicates that combining the feature sets does not result in interference between the sets when matching. Combining the minima and maxima features into one set does not result in matches becoming more ambiguous and it is therefore not necessary to match each set separately.

Matching scores are lower for the minima features than for the maxima features. The difference in matching score appears to be heavily dependent on the scene contents. For example, the *bark* and *boat* sequences both involve scale change, but the matching score performance of minima features is considerably worse in the *boat* sequence. Combining both sets yields a matching score that is a weighted average of the two, indicating again that there is little interference between feature types.

VI. Conclusion

This paper explores the properties of a novel feature, the minima of the determinant of Hessian operator. Two applications of this new feature are explored – corner extraction and affine feature extraction.

A corner detector is proposed that makes use of both maxima and minima of the determinant of Hessian function to accurately locate corners despite significant image blur and noise. Fitting line segments to the scale space loci of Hessian minima provides good corner estimates, while fitting line segments to the loci of Hessian maxima gives a good estimate of the corner bisector angle and assists in removing outlier minima corner estimates. A calibration experiment showed that this corner detector is superior to commonly used methods, especially when applied to heavily blurred and noisy images. The Hessian-based corner detector allows the use of new calibration patterns suitable for multi-spectral calibration.

It is shown that the minima of the determinant of Hessian function can easily be integrated into the well established determinant of Hessian (maxima) affine feature extractor to yield additional affine covariant features. The additional features can easily be extracted in parallel with the Hessian maxima and can be made scale and affine covariant using existing methods. An evaluation shows that the Hessian minima provides many additional correspondences. Matching experiments show that the new features do not confound the matching process and contributes additional matches without resulting in increased ambiguity for the existing maxima features. Minima and maxima features can therefore be managed as a single set of features and do not have to be matched separately.

Acknowledgment

This project was supported by Australian Research Council grant number LP0990135.

References

[1] C. Fookes, S. Denman, R. Lakemond, D. Ryan, S. Sridharan, and M. Piccardi. Semi-supervised intelligent surveillance system for secure environments. In *Industrial Electronics (ISIE), 2010 IEEE International Symposium on*, pages 2815–2820, 2010. 1

[2] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, 2 edition, 2003. 1

[3] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch. Visual modeling with a hand-held camera. *International Journal of Computer Vision*, 59(3):207–232, 2004. 1

[4] Y. Furukawa and J. Ponce. Accurate, dense, and robust multi-view stereopsis. In *Proc. Computer Vision and Pattern Recognition, IEEE Conference on*, pages 1–8, 2007. 1

[5] M. Brown and D.G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 74(1):59–73, 2007. 1

[6] P. Carbonetto, G. Dorkó, C. Schmid, H. Kück, and N. de Freitas. Learning to recognize objects with little supervision. *International Journal of Computer Vision*, 77(1):219–237, 2008. 1

[7] N. Snavely, S. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *International Journal of Computer Vision*, 80(2):189–210, 2008. 1

[8] T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: A survey. *Foundations and Trends in Computer Graphics and Vision*, 3(3):177–280, 2008. 1

[9] P.R. Beaudet. Rotationally invariant image operators. In *Proc. Pattern Recognition, International Joint Conference on*, pages 579–583, Kyoto, Japan, 1978. 1

[10] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1-2):43–72, 2005. 1, 3, 5

[11] P. Moreels and P. Perona. Evaluation of features detectors and descriptors based on 3d objects. *International Journal of Computer Vision*, 73(3):263–284, 2007. 1

[12] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conference*, pages 189–192, 1988. 1

[13] T. Kadir, A. Zisserman, and J.M. Brady. An affine invariant salient region detector. In *Proc. European Conference on Computer Vision*, pages 228 – 241, 2004. 1

[14] R. Lakemond, D.N.R. McKinnon, C. Fookes, and S. Sridharan. A feature clustering algorithm for scale-space analysis of image structures. In *Proc. Signal Processing and Communication Systems, International Conference on*, pages 186–192, 2007. 1, 3

[15] R. Lakemond. *Multiple Camera Management using Wide Baseline Matching*. Thesis, 2010. 1, 3

[16] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 60:63–86, 2004. 1

[17] R. Lakemond, C. Fookes, and S. Sridharan. Affine adaptation of local image features using the hessian matrix. In *Proc. Advanced Video and Signal Based Surveillance, IEEE conference on*, pages 496–501, Genoa, Italy, 2009. 1

[18] A. Baumberg. Reliable feature matching across widely separated views. In *Proc. Computer Vision and Pattern Recognition, IEEE Conference on*, volume 1, pages 774–781, 2000. 1, 3

[19] Opencv, http://opencv.willowgarage.com/, 22 February 2011. 3