Detecting Intrinsically Two-dimensional Image Structures Using Local Phase

Di Zang* and Gerald Sommer

Cognitive Systems Group
Institute of Computer Science and Applied Mathematics
Christian Albrechts University of Kiel, 24118 Kiel, Germany
{zd,gs}@ks.informatik.uni-kiel.de

Abstract. This paper presents a novel approach towards detecting intrinsically two-dimensional (i2D) image structures using local phase information. The local phase of the i2D structure can be derived from a curvature tensor and its conjugate part in a rotation-invariant manner. By employing damped 2D spherical harmonics as basis functions, the local phase is unified with a scale concept. The i2D structures can be detected as points of stationary phases in this scale-space by means of the so call phase congruency. As a dimensionless quantity, phase congruency has the advantage of being invariant to illumination change. Experiments demonstrate that our approach outperforms Harris and Susan detectors under the illumination change and noise contamination.

1 Introduction

Local image structures play important roles in many computer vision tasks. They can be associated with the term intrinsic dimensionality [1], which, as a local property of multidimensional signal, expresses the number of degrees of freedom necessary to describe local structures. For 2D images, there exist three type of structures. The intrinsically zero dimensional (i0D) structures are constant signals. Intrinsically one dimensional (i1D) structures represent lines and edges. Corners, junctions, line ends, etc. are all intrinsically two dimensional (i2D) structures which all have certain degree of curvature. It is well know that these i2D structures are of high significance in object recognition, motion estimation, image retrieval, etc. Consequently, correct detection of i2D structures under image deformations is very important.

There exist a lot of work concerning the detection of i2D structures based on intensity information, see [2–6]. These intensity based approaches are sensitive to variations in image illumination. Hence, it is necessary to find some features of local structures which are invariant with respect to image brightness change for a robust and reliable detection. Phase is such a good candidate, which carries most essential structure information of the original signal and has the advantage

* This work was supported by German Research Association (DFG) Graduiertenkolleg No. 357.

of being invariant to illumination variation [7]. Detecting local structures can be realized by means of the phase congruency. Using phase congruency to detect edges has been reported in [8, 9]. However, i2D structure detection based on its local phase has not yet been well investigated, although Kovesi proposed to use i1D local phase to detect i2D points by constructing the phase moments [10].

In this paper, we present a novel approach to detect i2D image structures using local phase information. The local phase of the i2D structure is derived from a curvature tensor and its conjugate part in a rotationally invariant way. By employing damped spherical harmonics as basis functions, the local phase is unified with a scale concept. The i2D structures can be detected as points of stationary phases in this scale-space by means of the so called phase congruency. Experimental results illustrate that our approach outperforms Harris and Susan detectors under illumination change and noise contamination.

2 Phase Estimation of Intrinsically Two-dimensional Image Structures

The local phase of an i2D structure can be derived from a tensor pair, namely, the curvature tensor and its conjugate part. By employing damped 2D spherical harmonics [11, 12] as basis functions, the local phase is unified with a scale-space framework. An nth order damped 2D spherical harmonic H_n has a much simpler representation in the spectral domain than that of the spatial domain. It takes the following form

$$H_n(\rho, \alpha; s) = \exp(in\alpha)\exp(-2\pi\rho s) = [\cos(n\alpha) + i\sin(n\alpha)]\exp(-2\pi\rho s) \quad (1)$$

where ρ and α denote the polar coordinates in the Fourier domain, s refers to the scale parameter. The damped 2D spherical harmonics are actually 2D spherical harmonics $\exp(in\alpha)$ combined with the Poisson kernel $\exp(-2\pi\rho s)$ [9]. The first order damped 2D spherical harmonic is basically identical to the conjugate Poisson kernel [9]. When the scale parameter is zero, it is exactly the Riesz transform [13]. In order to evaluate the local phase information, the curvature tensor and its conjugate part are designed to capture the even and odd information of 2D image structures. Designing the curvature tensor is motivated by the second order fundamental theorem of the differential geometry, that is the second derivatives or Hessian matrix which contains curvature information of the original signal. Let f be a 2D signal, its Hessian matrix is correspondingly given by

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \quad (2)$$

where x and y are the Cartesian coordinates. According to the derivative theorem of the Fourier theory [14, 15], the Hessian matrix in the spectral domain reads

$$\mathcal{F}\{H\} = \begin{bmatrix} -4\pi^2\rho^2\frac{1+\cos(2\alpha)}{2}F & -4\pi^2\rho^2\frac{\sin(2\alpha)}{2}F \\ -4\pi^2\rho^2\frac{\sin(2\alpha)}{2}F & -4\pi^2\rho^2\frac{1-\cos(2\alpha)}{2}F \end{bmatrix} \quad (3)$$

where F is the Fourier transform of the original signal f. It is obvious that angular parts of the second order derivatives in the Fourier domain are related to 2D spherical harmonics of even order 0 and 2. Hence, these harmonics represent the even information of 2D structures. Therefore, we are motivated to construct a tensor T_e, which is related to the Hessian matrix. This tensor is called a curvature tensor, because it is similar to the curvature tensor of the second fundamental form of the differential geometry. This curvature tensor T_e indicates the even information of 2D image structures and can be obtained from a tensor-valued filter H_e in the frequency domain, i.e. $T_e = \mathcal{F}^{-1}\{FH_e\}$, where $\mathcal{F}^{-1}$ means the inverse Fourier transform. Hence, the tensor-valued filter H_e, called the even filter reads

$$\begin{aligned} H_e &= \begin{bmatrix} \frac{H_0+\mathrm{real}(H_2)}{2} & \frac{\mathrm{imag}(H_2)}{2} \\ \frac{\mathrm{imag}(H_2)}{2} & \frac{H_0-\mathrm{real}(H_2)}{2} \end{bmatrix} = \begin{bmatrix} \frac{1+\cos(2\alpha)}{2} & \frac{\sin(2\alpha)}{2} \\ \frac{\sin(2\alpha)}{2} & \frac{1-\cos(2\alpha)}{2} \end{bmatrix} \exp(-2\pi\rho s) \quad (4) \\ &= \begin{bmatrix} \cos^2(\alpha) & \frac{1}{2}\sin(2\alpha) \\ \frac{1}{2}\sin(2\alpha) & \sin^2(\alpha) \end{bmatrix} \exp(-2\pi\rho s) \end{aligned}$$

where real$(\cdot)$ and imag$(\cdot)$ indicate the real and imaginary parts of the expression.

In this filter, two components $\cos^2(\alpha)$ and $\sin^2(\alpha)$ can be considered as two angular windowing functions. These angular windowing functions provide a measure of the angular distance, which is the same as the one of the structure tensor in [16]. From them, two perpendicular i1D components of the 2D image, oriented along the x and y coordinates, can be obtained. The other component of the filter is also the combination of two angular windowing functions, i.e. $\frac{1}{2}\sin(2\alpha) = \frac{1}{2}(\cos^2(\alpha - \frac{\pi}{4}) - \sin^2(\alpha - \frac{\pi}{4}))$. These two angular windowing functions yield again two i1D components of the 2D image, which are oriented along the diagonals. These four angular windowing functions can also be considered as four differently oriented filters, which are basis functions to steer a filter [17]. They make sure that i1D components along different orientations are extracted. Consequently, the even filter H_e enables the extraction of differently oriented i1D components of the 2D image.

The conjugate Poisson kernel, which evaluates the corresponding odd information of the i1D signal, is in quadrature phase relation with the i1D signal. Therefore, the odd representation of the curvature tensor, called the conjugate curvature tensor T_o, is obtained by employing the conjugate Poisson kernel to elements of T_e. Besides, the conjugate curvature tensor T_o results also from a tensor-valued odd filter H_o, i.e. $T_o = h_1 * T_e = \mathcal{F}^{-1}\{H_1 H_e F\} = \mathcal{F}^{-1}\{H_o F\}$, where h_1 denotes the conjugate Poisson kernel in the spatial domain. Hence, the odd filter H_o in the spectral domain is given by

$$H_o = \frac{1}{2}\begin{bmatrix} H_1(H_0 + \mathrm{real}(H_2)) & H_1(\mathrm{imag}(H_2)) \\ H_1(\mathrm{imag}(H_2)) & H_1(H_0 - \mathrm{real}(H_2)) \end{bmatrix} \quad (5)$$

Similar as the Hessian matrix, we are able to compute the determinant of T_e and T_o for knowing the existence of the i2D structure. Combing the determinants of T_e and T_o results in a novel model for the i2D structure, which is called the

monogenic curvature scale-space $\mathbf{f}_{i2D}(\mathbf{x}; s)$,

$$\mathbf{f}_{i2D}(\mathbf{x}; s) = \det(T_e(\mathbf{x}; s)) + \det(T_o(\mathbf{x}; s)) \tag{6}$$

From it, the local amplitude for the i2D structure is given by

$$a(\mathbf{x}; s) = \sqrt{\det^2(T_e(\mathbf{x}; s)) + \det^2(T_o(\mathbf{x}; s))} \tag{7}$$

and the local phase can be obtained as

$$\varphi(\mathbf{x}; s) = \frac{\det(T_o(\mathbf{x}; s))}{|\det(T_o(\mathbf{x}; s))|} \operatorname{atan}\left(\frac{|\det(T_o(\mathbf{x}; s))|}{\det(T_e(\mathbf{x}; s))}\right) \tag{8}$$

where $\frac{\det(T_o(\mathbf{x};s))}{|\det(T_o(\mathbf{x};s))|}$ decides the local main orientation of the i2D structure. Hence, the local phase information of the i2D structure contains not only phase information but also the local main orientation. Therefore, the evaluation of the i2D structure can be realized in a rotation-invariant way.

3 Phase Congruency

Since the local phase is independent of the local amplitude, it thus has the advantage of being not sensitive to illumination change. Hence, detecting i2D image structures can be done by looking for points of stationary phase in the scale-space. This approach is commonly called phase congruency and is based on comparisons of the local phase at certain distinct scales [18, 8]. In this paper, we take a similar idea as those reported in [8, 10]. However, there are some differences. First, our local phase information can be evaluated in a rotation-invariant manner. Therefore, no orientation sampling is required. Second, the local phase directly indicates the phase information of the i2D structure. Thus, there is no need to construct principal moments of the phase congruency to determine i2D structures.

Morrone and Owens [19] define the phase congruency function in terms of the Fourier series expansion of a signal at a local position $\mathbf{x}$ as

$$PC = \max_{\overline{\phi}\in(0,2\pi]} \frac{\sum_n A_n \cos(\phi_n - \overline{\phi})}{\sum_n A_n} \tag{9}$$

where A_n represents the amplitude of the nth Fourier component, ϕ_n denotes the local phase of the Fourier component at position $\mathbf{x}$ and $\overline{\phi}$ is the amplitude weighted mean local phase angle of all the Fourier terms at the position being considered. The measure has a value between zero and one. A phase congruency of value one means that there is an edge or a line, zero phase congruency indicates there is no structure. However, this measure results in poor localization and is also sensitive to noise. Hence, Kovesi [8] developed a modified version of the phase congruency. In this measure, the local phase is obtained from the logarithmic Gabor wavelet. Due to its lack of rotation invariance, orientation sampling must

be employed to make sure that features at all possible orientations are treated equally. Hence, the new measure of phase congruency reads

$$PC = \frac{\sum_o \sum_n W_o \lfloor A_{no}(\cos(\phi_{no} - \overline{\phi}_o) - |\sin(\phi_{no} - \overline{\phi}_o)|) - T_o \rfloor}{\sum_o \sum_n A_{no} + \varepsilon} \tag{10}$$

where n and o refer to the scale parameter and the index over orientations, respectively. And W_o denotes a factor that weights for frequency spread along certain orientation and ε is added to avoid division by zero. The terms A_{no} and ϕ_{no} are the local amplitude and local phase at a certain scale and orientation, respectively. The mean local phase at a certain orientation is represented as $\overline{\phi}_o$. Only energy values that exceed the estimated noise influence T_o can be taken into consideration. The symbols $\lfloor$ and $\rfloor$ indicate that the enclosed entity equals itself when its value is positive and zero otherwise. This new phase congruency measure produces a more localized response and it also incorporates noise compensation. However, the estimated local phase is only valid for the i1D signal. Hence, using phase congruency to detect i2D structures requires the construction of principal moments of the phase congruency, see [10].

In contract to this, we have now a rotationally invariant evaluation of the local phase for the i2D structure, no orientation sampling is needed. Hence, the computation of phase congruency can be simplified as the following

$$PC = \frac{\sum_n W \lfloor A_n(\cos(\phi_n - \overline{\phi}) - |\sin(\phi_n - \overline{\phi})| - T) \rfloor}{\sum_n A_n + \varepsilon} \tag{11}$$

where n denotes the scale parameter, W is also a factor weighting for frequency spread, A_n and ϕ_n represent the local amplitude and local phase of the i2D structure point, respectively. This new measure can be directly applied to detect i2D image structures. Any point with a phase congruency value higher than a certain threshold can be considered as an i2D point.

4 Performance Evaluation Criteria

In the literature, many detectors are designed for detecting i2D image structures. However, most of them show only qualitative experimental results. Because computer vision tasks require more robust and reliable detection results, there has been an increasing emphasis on quantitative performance evaluation. There also exists a number of research for assessing the detector performance. The measure suggested by Schmid et al. [20] is based on the idea of repeatability. Rockett [21] and Martinez-Fonte et al. [22] proposed a more empirical method for accessing. In their research, examples of true corners and non-corners are provided. For each threshold level, the corner detection probability and the false alarm rate are estimated to plot an ROC curve. In [23], Carneiro et al. assessed the detector performance by two measures, namely, the precision and recall rates.

The repeatability evaluation delivers the number of points repeated between two images with respect to the total number of detected points. However, this

measure does not consider those correctly or wrongly detected points which do not repeat at all. The ROC curve plots the relation between the detection rate and false alarm rate with respect to the threshold variation, but it is not easy to show the detection performance with respect to image deformations like illumination change, rotation change and so on. In this paper, we follow the measures in [23].

The recall rate measures the probability of finding an i2D point in a deformed image given that it is detected in the reference image. The definition of the recall rate is given by

$$R = \frac{TP}{TP + FN} \tag{12}$$

where TP denotes the true positive and FN is the false negative. Since it is not easy to identify the ground truth, in this case, the true positive means the number of correctly matched points. Given a point $\mathbf{x}_i$ in the reference image and a point $\mathbf{x}_j$ in the deformed image, let $M(\cdot)$ represent the deformation transform, if the Euclidean norm condition is satisfied, i.e. $\|M(\mathbf{x}_i) - \mathbf{x}_j\| < 1.5$, then these two points are correctly matched. False negative is the number of points in the reference image which cannot be matched with any points in the deformed image.

The precision rate indicates the probability that an i2D point detected in a deformed image is actually an i2D point in the reference image. Its definition reads

$$P = \frac{TP}{TP + FP} \tag{13}$$

where FP is false positive, it means the number of points in the deformed image which cannot be matched with any points in the reference image. Both the recall and precision rates have values between zero and one. If the rate is higher, the detection performance is better.

5 Experimental Results

In this section, we present some experimental results. As shown in Fig. 1, two test images and one image sequence are employed for the experiments. The first

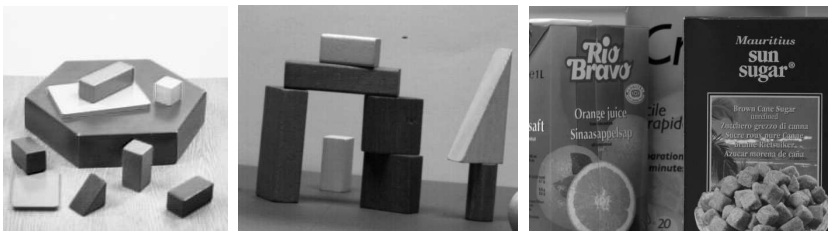

Fig. 1. Two test images (blox and blocks) and one frame of a boxes image sequence.

experiment aims to illustrate some qualitative comparison results between our

approach and the well-known Harris detector. The blox image is used for detection under the rotation change and the additive Gaussian noise contamination (standard derivation is 10). For the illumination change, we use the blocks image to show the detection difference. Fig. 2 demonstrates the detection results of our approach and the Harris detector under the rotation change, the noise contamination and the illumination change. According to the false positives and false negatives, it can be shown that our approach performs better than the Harris detector when the illumination changes and the noise is added to some degree.

The second experiment is to show some quantitative comparison results. We follow the evaluation criteria of recall and precision rates to compare the performances of our approach, the Harris detector and also the well-known Susan detector. Ten frames of the boxes image sequence are employed for this experiment. Image deformations of rotation change, additive Gaussian noise contamination and the illumination variation are considered. For each deformation, the averaged values of ten frames are recorded to plot the recall and precision rates. Fig. 3 demonstrates comparison results between our approach, the Harris detector and the Susan detector according to the performance assessment criteria of recall and precision rates. Note that recall and precision rates have different scales for different image deformations. The top row shows detection results under the rotation change. Our approach has a comparable result with the Harris detector, and the Susan detector performs worse than these two approaches. The second row are recall and precision rates for the illumination change. The phase congruency is a dimensionless quantity which is in theory invariant to the illumination change, although it is not absolutely invariant to brightness variation in practice, it is still less sensitive to the illumination variation than those insensity based approaches. Results indicate that our approach performs much better than the Harris and Susan detectors especially in the case of higher illumination change. Bottom row shows the additive Gaussian noise contaminated results. Since the phase congruency takes several scales into consideration and it also incorporates noise compensation, our approach demonstrates a better performance than that of the Harris detector. And the Harris detector is less sensitive to the noise when compared with that of the Susan detector due to the Gaussian smoothing in the local neighborhood.

6 Conclusions

We present a novel approach towards detecting i2D image structures using local phase information. The local phase of the i2D structure can be derived from a curvature tensor and its conjugate part in a rotation invariant manner. The i2D image structures are detected as those points with stationary phases in the scale-space by means of phase congruency. The recall and precision rates are employed as detection performance assessment criteria. Experimental results illustrate that our approach outperforms the Harris and Susan detectors when the illumination changes and the images are contaminated by the additive Gaussian noise. For

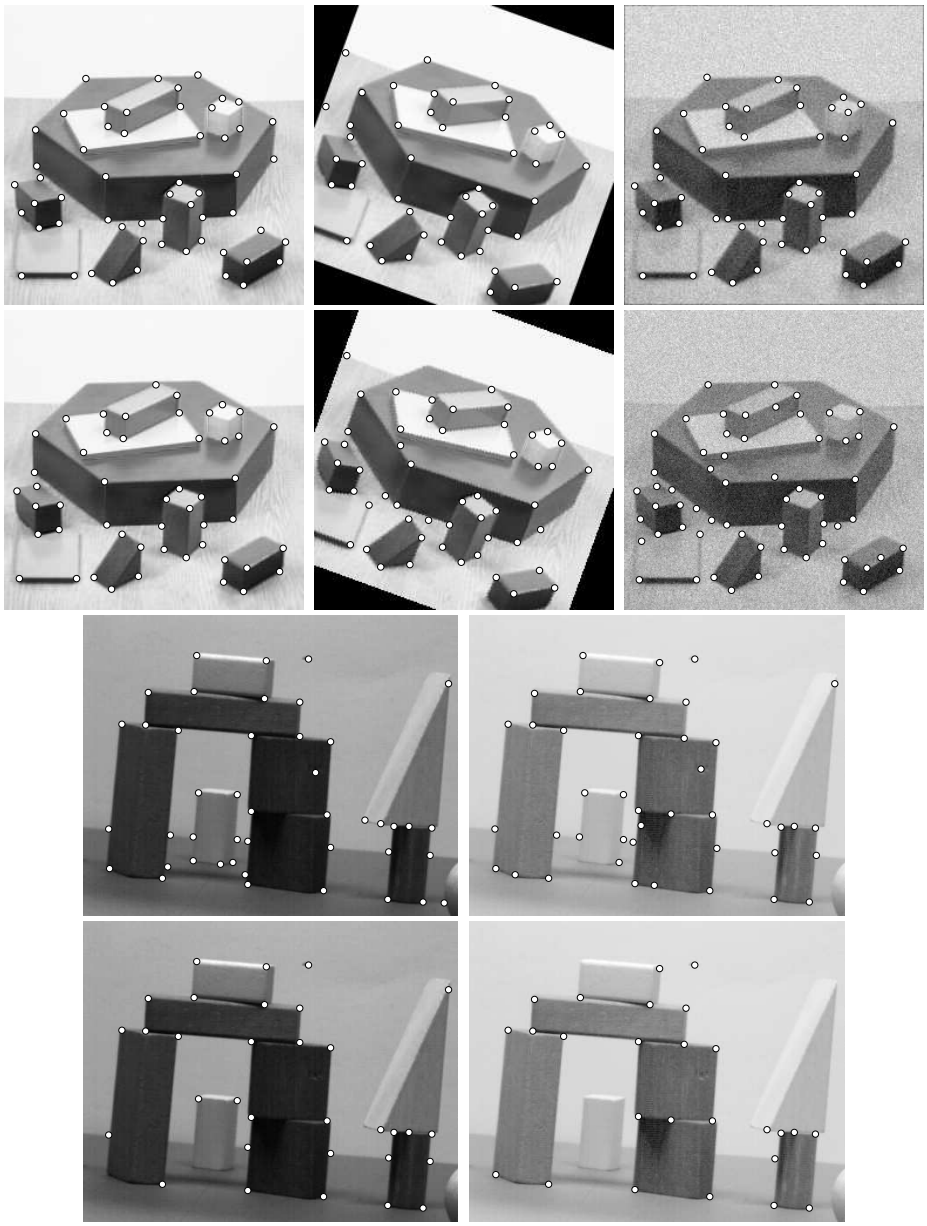

Fig. 2. Top row shows the detection results using our approach for the original image, the rotated image and the Gaussian noise contaminated image. The second row demonstrates the results from the Harris detector for the original image, the rotated one and the noise contaminated one. Results shown in the third row are detections for the original image and the illumination varied one by using our approach. Bottom row illustrates results from the Harris detector for the original image and the illumination changed one.

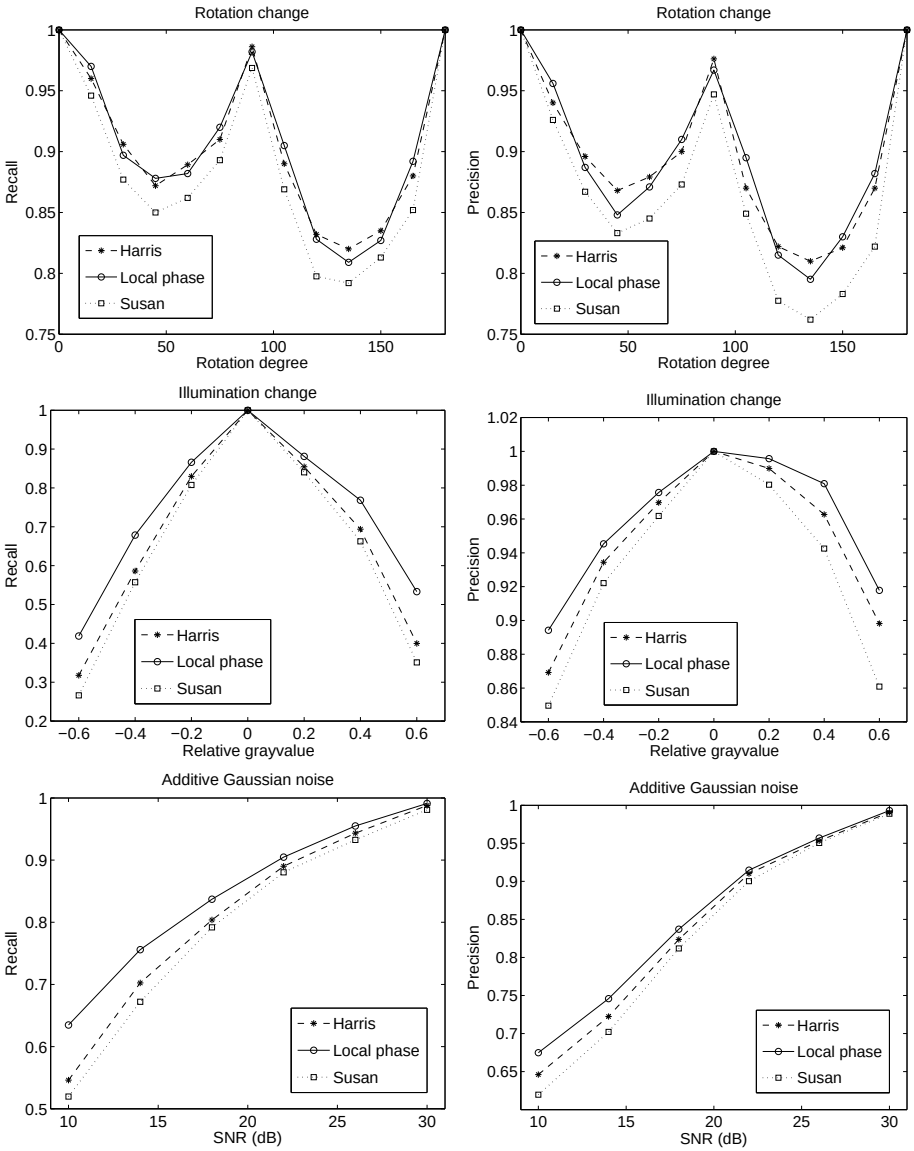

Fig. 3. First column: from top to bottom are recall rates under the rotation change, illumination variation and the additive Gaussian noise contamination. Second column: from top to bottom are precision rates under the rotation change, illumination variation and the additive Gaussian noise contamination.

the deformation of rotation change, our approach shows a comparable result with the Harris detector.

References

1. Zetzsche, C., Barth, E.: Fundamental limits of linear filters in the visual processing of two-dimensional signals. Vision Research **30** (1990) 1111–1117
2. Beaudet, P.: Rotationally invariant image operators. In: Proceedings of International Joint Conference on Artificial Intelligence. (1978) 579–583
3. Kitchen, L., Rosenfeld, A.: Grey-level corner detection. Pattern Recognition Letters (1982) 95–102
4. Förstner, W., Gülch, E.: A fast operator for detection and precise location of distinct points, corners and centers of circular features. In: Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland (1987) 281–305
5. Harris, C., Stephen, M.: A combined corner and edge detector. In: Proceedings of 4th Alvey Vision Conference, Manchester (1988) 147–151
6. Smith, S.M., Brady, J.M.: Susan – a new approach to low level image processing. International Journal of Computer Vision **23**(1) (1997) 45–78
7. Oppenheim, A.V., Lim, J.S.: The importance of phase in signals. IEEE Proceedings **69** (1981) 529–541
8. Kovesi, P.: Videre: A image features from phase congruency. Journal of Computer Vision Research **1**(3) (1999)
9. Felsberg, M., Sommer, G.: The monogenic scale-space: A unifying approach to phase-based image processing in scale-space. Journal of Mathematical Imaging and Vision **21** (2004) 5–26
10. Kovesi, P.: Phase congruency detects corners and edges. In: Proc. The Australian Pattern Recognition Society Conference. (2003) 309–318
11. Felsberg, M.: Low-level image processing with the structure multivector. Technical Report 2016, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik (2002)
12. Felsberg, M.: On the design of two-dimensional polar separable filters. In: 12th European Signal Processing Conference, Vienna, Austria, EURASIP (2004) 417–420
13. Felsberg, M., Sommer, G.: The monogenic signal. IEEE Transactions on Signal Processing **49**(12) (2001) 3136–3144
14. Papaulis, A.: The Fourier Integral and its Application. McGraw-Hill, New York (1962)
15. Bracewell, R.: Fourier Analysis and Imaging. Kluwer Academic / Plenum Publishers, New York (2003)
16. Granlund, G., Knutsson, H.: Signal Processing for Computer Vision. Kluwer Academic Publishers, Dordrecht (1995)
17. Freeman, W.T., Adelson, E.H.: The design and use of steerable filters. IEEE Trans. Pattern Analysis and Machine Intelligence **13**(9) (1991) 891–906
18. Reisfeld, D.: The constrained phase congruency feature detector: simultaneous localization, classification and scale determination. Pattern Recogn. Lett. **17**(11) (1996) 1161–1169
19. Morrone, M.C., Owens, R.A.: Feature detection from local energy. Pattern Recognition Letters **6** (1987) 303–313

20. Schmid, C., Mohr, R., Bauckhage, C.: Evaluation of interest point detectors. International Journal of Computer Vision **37**(2) (2000) 151–172
21. Rockett, P.: Performance assessment of feature detection algorithms: A methodology and case study on corner detectors. IEEE Transactions on Image Processing **12**(12) (2003) 1668–1676
22. Martinez-Fonte, L., Gautama, S., Philips, W.: An empirical study on corner detection to extract buildings in very high resolution satellite images. In: Proc. of IEEE-ProRisc. (2004) 288–293
23. Carneiro, G., Jepson, A.D.: Multi-scale phase-based local features. In: Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR03). (2003) 736–743