Junction classification by multiple orientation detection

M. Michaelis[1] and G. Sommer[2]

[1] GSF-MEDIS-Institut, Neuherberg, D-85758 Oberschleißheim, Germany
[2] Institut für Informatik, Christian-Albrechts-Universität, D-24105 Kiel, Germany

Abstract. Junctions of lines or edges are important visual cues in various fields of computer vision. They are characterized by the existence of more than one orientation at one single point, the so called keypoint. In this work we investigate the performance of highly orientation selective functions to detect multiple orientations and to characterize junctions. A quadrature pair of functions is used to detect lines as well as edges and to distinguish between them. An associated one-sided function with an angular periodicity of 360° can distinguish between terminating and non-terminating lines and edges which constitute the junctions. To calculate the response of these functions in a continuum of orientations and scales a method is used that was introduced recently by P. Perona [8].

1 Introduction

Junctions of lines and edges are important visual cues in various fields of computer vision. They have a complex 2D structure with many degrees of freedom. Hence, there will be no computationally cheap solution for the classification and quantitative description of junctions. Therefore, one should drop the requirement of fast algorithms that can process the whole image as in simple edge detection. Instead, junction classification might only be sensible in computer vision systems with attentional mechanisms that become more and more popular in recent time. There, only some small regions of interest are investigated in depth what justifies computationally involved methods.

In this contribution we propose a method to classify and describe junctions that is based on multiple orientation detection. To limit interference effects of lines or edges of different orientations we use a function with a high orientation selectivity. The point where the lines or edges intersect is called the keypoint. It is assumed that the keypoint already has been detected and all functions will be centered at the keypoint. To detect the events independently of their profiles (line, edge, mixed) we use a (pseudo) quadrature pair of functions ([7]). The complex response of the quadrature pair can be transformed to an energy/phase representation, where the energy detects the events and the phase reveals the type (line or edge). We define the energy to be the modulus of the complex response. The interpretation of these signals requires the response in a continuum of orientations and scales. For this we use the steerability method of Perona [8]. Perona introduces a set of basis functions that can be steered easily and that

Lecture Notes in Computer Science, Vol. 800
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

can approximate the original function by linear combinations. A discussion of this method is given in Michaelis and Sommer [5].

Related work: Noble [6] outlines a detection and classification method based on the evaluation of groups of elliptic and hyperbolic points. This method is quite sensitive to noise. Guiducci [4] estimates the parameters of corners (amplitude, aperture and smoothness of edges) from second order Taylor approximations. More complicated junctions are not dealt with. Rohr [9] detects and classifies junctions with two or more edges by fitting wedge models to the image. This method gives a precise description of edge junctions but it has difficulties with situations that are not in the range of the models of the fitting algorithm, e.g. line junctions or a non uniform brightness between the edges. Brunnström et al. [2] evaluate the local histogram of the junction to derive a hypothesis about the number of edges. The hypothesis is verified using an edge image provided by a Canny-Deriche edge detector. This method cannot handle line junctions, and the evaluation of the histogram might be hard for real world junctions. Andersson [1] developed a method similar to ours in evaluating the prominent events in the orientational responses of filters of a two- and one-sided type. But the design of the filters, the interpolation method and the evaluation of the responses are different. Perona [8] hints very briefly to the detection and classification of junctions using steerable filters.

2 The double Hermite function

The function we use for orientation analysis consists of two elongated 2D-Gaussians and a second/first derivative in the y-direction for the real/imaginary part:

$$\mathrm{F}_\sigma(\mathrm{x},\mathrm{y}) \;=\; \mathrm{N}\ (1-(\mathrm{y}/\sigma)^2+\mathrm{j}\mathrm{y}/\sigma)\ \mathrm{e}^{-\frac{\mathrm{y}^2}{2\sigma^2}}\ (\mathrm{e}^{-\frac{(\mathrm{x}-2\varepsilon\sigma)^2}{2(\varepsilon\sigma)^2}}+\mathrm{e}^{-\frac{(\mathrm{x}+2\varepsilon\sigma)^2}{2(\varepsilon\sigma)^2}})$$

$$F_{\sigma,\theta}(x,y) \;:=\; F_\sigma(x\cos\theta-y\sin\theta, x\sin\theta+y\cos\theta)$$

ε is the aspect ratio of the 2D-Gaussians, N is a normalization factor, θ and σ are the orientation and the scale. The real and imaginary y-cross-sections are approximately in quadrature. To give it a name we called this function the double Hermite function. Fig.1 shows this function and the associated one-sided function [8]. The latter consists essentially (but not exact) of the right lobe of the double Hermite function. The one-sided function has a periodicity in orientation of 360° and can therefore distinguish between terminating and non-terminating events (T- or X-junction). The same basis functions that steer the double Hermite function can steer the one-sided function as well. Hence, there is no extra computational effort. This is one advantage of the steerability method of Perona compared to approaches that use rotated copies of the original function as basis filters for the interpolation (Freeman and Adelson [3]).

Commonly the functions used for estimating the orientation have the bulk of their energy at the center ([3],[8]). There are three reasons to choose a function with almost no energy at the center: (1) the discretization of the orientation is worst at the center, (2) frequently the immediate neighborhood of real world keypoints has a confusing structure (e.g. Fig.4), (3) the associated one-sided function performs better.

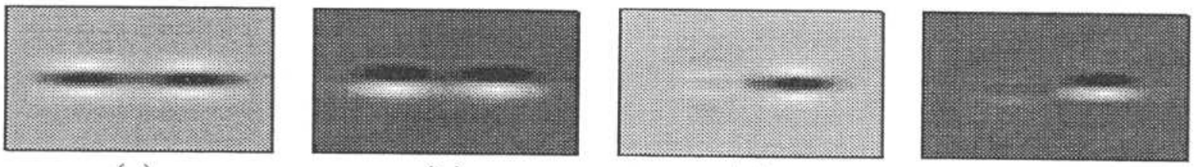


Fig. 1. Double Hermite function (a) Re, (b) Im and the associated one-sided function (c) Re, (d) Im.

3 Junction classification by multiple orientation detection

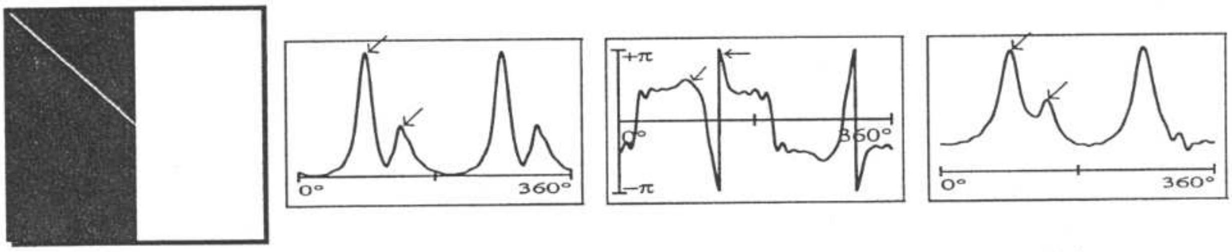
Our model for junctions consists of straight edges and lines that intersect in one point. The junctions will be classified by the number of events, their orientation, their type (line or edge) and whether the events are terminating or non-terminating. For the presentations of this section 23 basis functions in orientation and 4 in scale are used to steer the double Hermite function. The aspect ratio ε is 3. The orientation $0°$ is from the keypoint to the right, positive rotations are anti-clockwise. All responses in the figures are normalized to the same amplitude because we are mainly interested in their shape and not in the absolute values. Hence, the relative strength of the response between the figures and sub-figures cannot be compared.

3.1 Principle of junction classification

Fig.2 demonstrates the principle of the analysis with a synthetic junction of events without an intrinsic scale (step edge, thin line). Hence, we can use any scale to analyse the junction. The prominent maxima in the energy of the double Hermite response give the number and orientations of the events. For this, in general the double Hermite function is preferable to the one-sided function because it has the better orientation selectivity. The phase at these orientations reveals the type of the events. A phase of $\pm\pi/2$ indicates an odd event (dark/bright or brigth/dark edge), for $0, \pm\pi$ the event is even (dark or bright line). The phase has a periodicity of 2π, i.e. $+\pi$ and $-\pi$ have to be identified. The double Hermite function cannot distinguish between terminating and non-terminating events. This information is contained in the one-sided response that has a periodicity of $360°$.

Smooth edges and thick lines have an intrinsic scale and therefore these events are only classified correctly above a certain scale. At smaller scales the local structure looks completely different. Fig.3 shows an example of a blurred edge and a thick line.

Resolution: Two events cannot be distinguished if their separation in orientation is below the resolution of the analysing function. The exact possible resolution depends on the relative strength and the profile of the events. The



(a) (b) (c) (d)

Fig. 2. (a) junction, (b) energy, (c) phase for the double Hermite function. (d) energy for the one-sided function. The energy shows two events at 90° and 136° (modulo 180°). The phase at the orientation of the maxima of the energy (1.77 and 3.00) reveals that they originate from a line at 136° and an edge at 90°. Due to the influence of the dominating edge the maximum of the line is shifted by less than one degree. The missing second peak of the line at 316° in (d) shows that it terminates at the keypoint. The center of the double Hermite function matches the keypoint within 1 pixel.

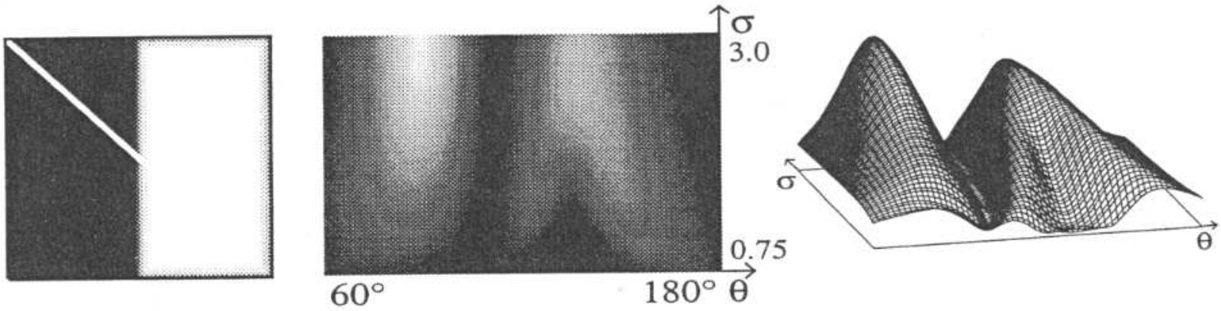


Fig. 3. Depicted are from left to right the junction and the energy of the double Hermite response as an intensity image and as a wire plot. The junction is the same as in Fig.2 but with a Gaussian blurred edge ($\sigma = 3$) and a thick line (5 pixels). At small scales the response to the edge is weak and it is interfering with the line. The line is seen as two events with a scale dependent orientation.

peaks in the real and imaginary response are better resolved than in the energy. However, the connection between the peaks and the underlying events is more complex. Furthermore the resolution can be improved by analysing the exact shape of the peaks in the response. A peak that results from two unresolved events will have a different shape than a single event peak (see [1]). However, such an analysis depends on the models of the junctions. When dealing with real world junctions it will be prone to give false results.

3.2 Junctions in real-world images

For real-world junctions there are difficulties that are not present in the previous synthetic examples: noise, the neighborhood of the junction, and not ideal events. The proposed method is designed to be robust with respect to these difficulties: (1) The responses are practically not affected by noise because the smoothing in the orientation of the double Hermite function $\varepsilon\sigma$ is large. (2) For more robust orientation estimations of terminating events the one-sided function instead of the double Hermite function can be used. It is more accurate if the neighborhood is not 'neutral' (other events, not constant gray values). (3) Both functions have the main part of their energy some distance away from the center. This corresponds to the fact that the immediate neighborhood of the keypoint frequently has a different structure than the junction (e.g. blob) and that it is not so important for the classification. (4) Another important part of the analysis of real junctions is the scale dependence of the responses. The events are frequently only visible in a certain range of scales. For the orientation estimation the scale with the strongest response is used. Important for the acceptance of an event is a stable orientation with respect to the scale. If the orientation is not stable the event might be not straight or the scale is not appropriate (e.g. small scales for the thick line in Fig.3). Another reason for an unstable orientation is an inaccurately detected keypoint. If the shift between the keypoint and the center of the functions is not too large the correct orientation is given by the larger scales where it is stable. Moreover, from the bending of the ridge in the one-sided energy the direction of the true keypoint can be estimated.

Fig.4 shows the responses of the one-sided function for a junction of poor quality.

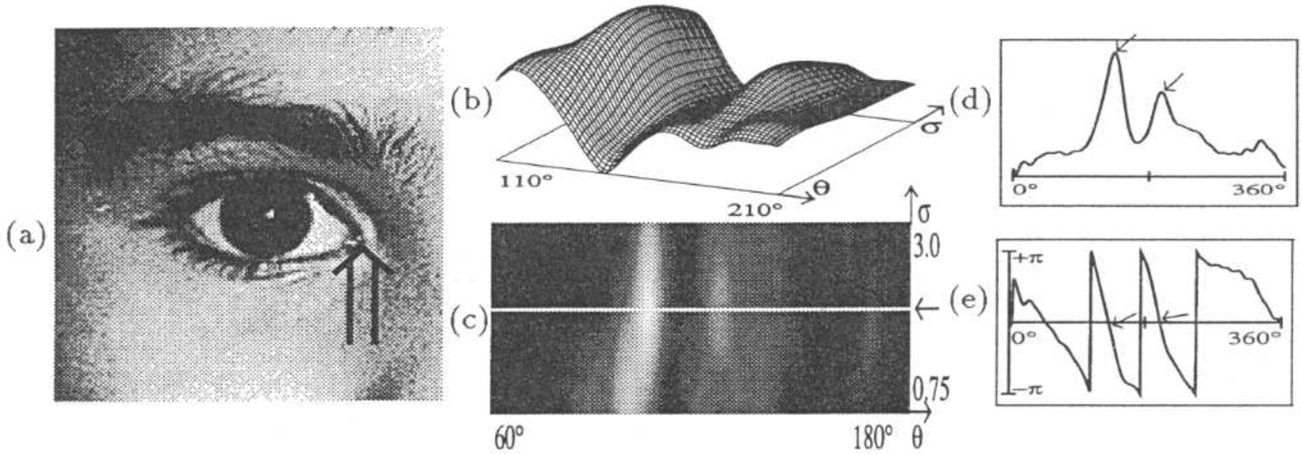


Fig. 4. (a) L-junction of the dark lines at the border of the eye. (b),(c) one-sided energy as wire plot and intensity image. (d),(e) energy and phase at the optimal scale that is marked in (c). The lower line of the junction is only visible in a certain range of scales. The one-sided response is more reliable, because there are only terminating lines.

4 Steerable functions for multiple orientation detection

In this section we discuss the double Hermite function in some more detail and we compare it to the function used by Perona [7],[8]. The latter is one elongated Gaussian with the second derivative (Re) and its true Hilbert-partner (Im) as the y-cross-section (Fig.5a,b).

Double-lobed/single-lobed: The function used by Perona consists of one elongated Gaussian with the main part of its energy around the center (Fig.5a,b). The same is true for the associated one-sided function (Fig.5c,d). We argued before that this is disadvantageous because the immediate neighborhood of a real world keypoint might be confusing. The weights the function gives to the neighborhood are not in agreement with the 'usual' interpretation of the local structure. In contrast the double Hermite function and its associated one-sided function have almost no energy at the center (Fig.1). As a consequence Fig.8a,c shows that e.g. for an endstopping the response for the double Hermite function falls almost to zero. This is not the case for the function used by Perona.

Quadrature: Both functions are complex with an even real part and an odd imaginary part. This allows to detect lines as well as edges and events with mixed profiles ([7]). For both functions the real part is the second derivative of a Gaussian whereas the imaginary part is the first derivative in case of the double Hermite function and the true Hilbert partner in case of Perona's function. The y-cross-sections, the energy, and the phase of the cross-sections are depicted in Fig.5e-h and 8a-d. For the exact quadrature pair the energy is monomodal and the phase is linear around the center. For the pseudo quadrature pair the phase is not exactly linear and the energy is at the borderline of having several peaks. In return the central peak is very sharp. The same is true for the energy of the responses of the two functions.

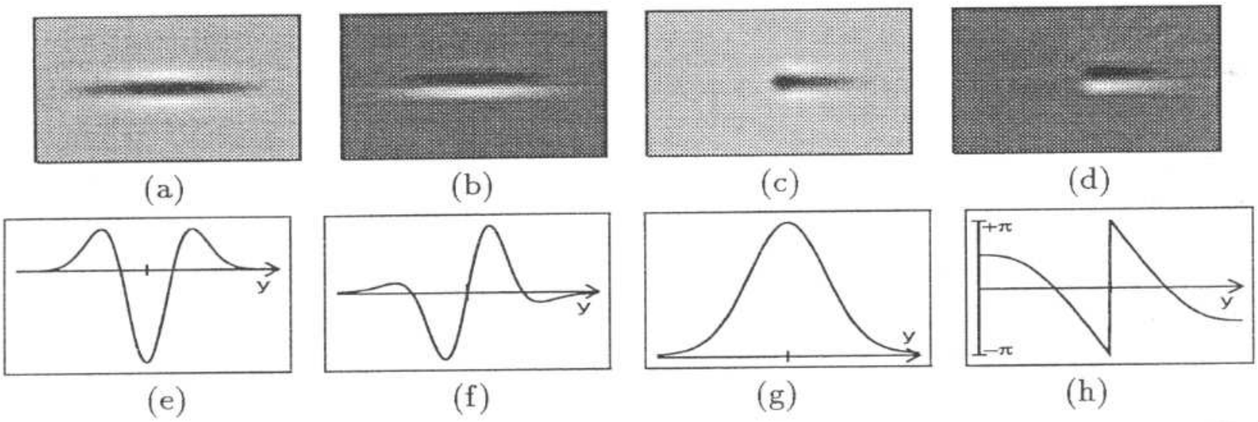


Fig. 5. The function used by Perona (a) Re, (b) Im and the associated one-sided function (c) Re, (d) Im. The aspect ratio of the Gaussian is 5. (e) and (f) are the y-cross-sections of (a) and (b). (g) and (h) are the energy and phase of the y-cross-section.

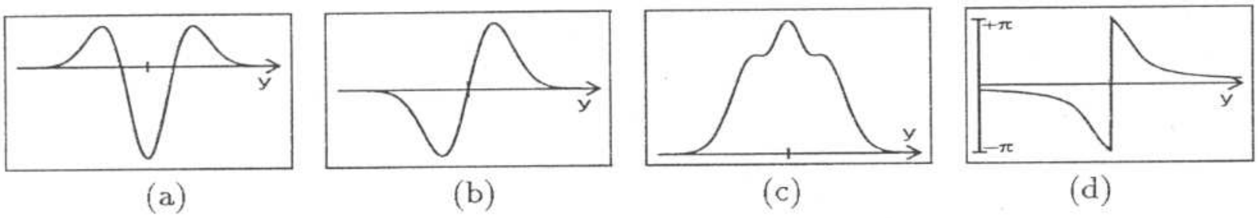
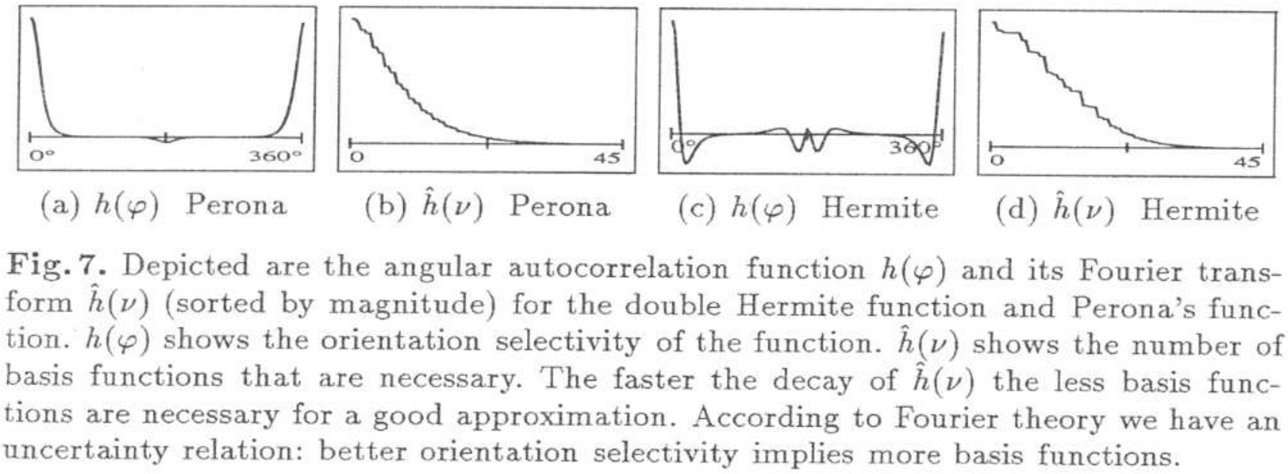


Fig. 6. y-cross-section of the double Hermite function. (a) real part, (b) imaginary part, (c) energy, and (d) phase.

Orientational uncertainty relation: The better the orientation selectivity of a function the more basis functions are necessary to steer it. Both functions are not exactly steerable with a finite number of basis functions. According to the theory ([5],[8]) the orientation selectivity is given by the autocorrelation function $h(\varphi) = \int F_0(\mathbf{x}) F_\varphi(\mathbf{x}) d^2\mathbf{x}$ (Fig.7a,c). To steer a function in orientation one basis function is needed for every non zero component of the Fourier transform $\hat{h}(\nu)$ of h (Fig.7b,d). The L^2 contribution of a basis function to the steered function is given by the modulus of $\hat{h}$ for this component. Therefore the area below the curve in Fig.7b,d shows the quality of the approximation given a finite number of basis functions. For very accurate approximations both functions need about the same number of basis functions but in the beginning (only few basis functions) the decay is faster for Perona's function (Fig.7b). This does not mean that Perona's function has a better orientation selectivity when it is approximated with only a few basis functions. Instead the main approximation error are strong ripples (Fig.8b,d). The double Hermite function has stronger ripples (Fig.8d) consistend with the slower decay of $\hat{h}(\nu)$ but the peaks are sharper.

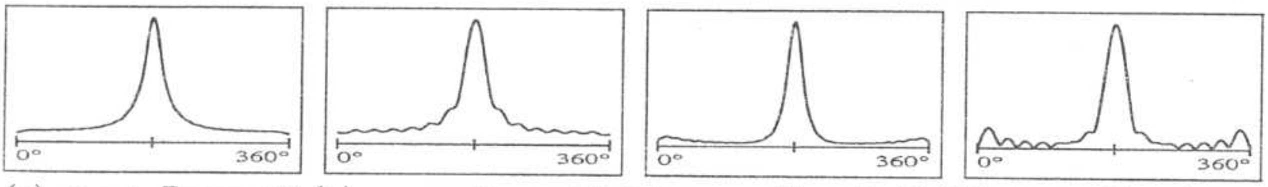
Fig.7a,c show that for the double Hermite function $h(\varphi)$ is sharper than for the function used by Perona but not monomodal. This is due partly to the 'double-lobedness' and partly to the **pseudo**-quadrature pair.

For short one could roughly say that there is a tradeoff between sharper peaks and stronger ripples between the functions. Sharper peaks imply a better orientation selectivity and more robustness against noise. On the other hand stronger ripples lead to false event detections.



(a) $h(\varphi)$ Perona (b) $\hat{h}(\nu)$ Perona (c) $h(\varphi)$ Hermite (d) $\hat{h}(\nu)$ Hermite

Fig. 7. Depicted are the angular autocorrelation function $h(\varphi)$ and its Fourier transform $\hat{h}(\nu)$ (sorted by magnitude) for the double Hermite function and Perona's function. $h(\varphi)$ shows the orientation selectivity of the function. $\hat{h}(\nu)$ shows the number of basis functions that are necessary. The faster the decay of $\hat{h}(\nu)$ the less basis functions are necessary for a good approximation. According to Fourier theory we have an uncertainty relation: better orientation selectivity implies more basis functions.



(a) one-s. Perona23 (b) one-s. Perona15 (c) one-s. Hermite23 (d) one-s. Hermite15

Fig. 8. Depicted is the energy of the responses of the one-sided Hermite function and one-sided Perona's function for an horizontal endstopping. To demonstrate the effect of approximations with only a few basis functions the responses for 15 and for 23 basis functions are shown. The main effect of taking less basis functions is not a worse orientation selectivity but strong ripples. The same is true for the two-sided original functions.

5 Conclusions

We presented a method to classify junctions by detecting the edges and lines that constitute them. For this we use a function that we designed for this particular task and evaluate its response in a continuum of orientations and scales. Our method is applicable to junctions of lines as well as edges and it is robust under real world conditions: noise, shaded neighborhoods, not ideal edges and lines. Depending on the required orientation selectivity a large number of basis functions is necessary to steer the analysing function. Hence, it is not intended to apply this method by convolution to the whole image. Preceeding processing steps are needed which guide the attention to a selection of interesting candidate points. This method is assumed to be a part of complex vision systems that use attentional mechanisms to allow more involved algorithms for the foveated parts of the image. The presented evaluation of the responses of the analysing function is not the optimum that is possible. Depending on the situation a flexible use of the energy, the real, and the imaginary response of the two-sided as well as the one-sided function for the detection and orientation estimation of the events would be better. In addition the shape of the peaks in the response can be analysed as in [1]. However, one has to keep in mind that more sophisticated evaluations are more dependent on the models and hence are more likely to give false results in real world images. Another open question is an efficient handling when the keypoint is not detected properly. We already mentioned that this results in scale dependent orientation estimates. At larger scales the orientations are exact. But this works only if the detection error is not too big. Otherwise a method that brings the keypoint into the focus would perform better.

References

1. M.T. Andersson, *Controllable multidimensional filters and models in low level computer vision*, PhD Thesis, Linköping University, S-58183 Linköping, Sweden, 1992, Diss. No. 282, ISBN 91-7870-981-4.
2. K. Brunnström, J.-O. Eklundh and T. Lindeberg, *On scale and resolution in active analysis of local image structure*, Image and Vision Computing **8** (1990) 289-296.
3. W.T. Freeman and E.H. Adelson, *The design and use of steerable filters for image analysis, enhancement, and wavelet representation*, IEEE PAMI **13** (1991) 891-906.
4. A. Guiducci, *Corner characterization by differential geometry techniques*, Pattern Recognition Letters **8** (1988) 311-318.
5. M. Michaelis and G. Sommer, *Keypoint characterization in images*, Proceedings of the SPIE **2093**, in press, (1993).
6. J.A. Noble, *Finding corners*, Image and Vision Computing, **6** (1988) 121-128.
7. P. Perona and J. Malik, *Detecting and localizing edges composed of steps, peaks and roofs*, UC Berkeley, Tech. Rep. UCB/CSD 90/590 (1990).
8. P. Perona, *Steerable-scalable kernels for edge detection and junction analysis*, ECCV 92 (1992) 3-18.
9. K. Rohr, *Recognizing corners by fitting parametric models*, Int.J. of Computer Vision **9** (1992) 213-230.