Characterization of keypoints in images

Markus Michaelis † and Gerald Sommer ‡

† GSF-Medis-Institut, Neuherberg, Germany Email: michaeli@gsf.de
‡ Institut für Informatik, Christian-Albrechts-Universität, Kiel, Germany

ABSTRACT

In this contribution we investigate the performance of steerable functions to characterize keypoints. Steerable functions were introduced recently by Perona[9] as an efficient method to calculate the response of a filter in a continuum of orientations, scales, and other parameters. For the analysis of points with events at multiple orientations, functions with a high orientational resolution are needed. We discuss criteria to judge the quality of a function to serve for orientation analysis. To handle line as well as edge junctions, we use a complex function with the real and imaginary part approximately in quadrature. An associated one-sided function allows to distinguish between terminating and nonterminating edges and lines. To analyse thick lines and blurred edges the function is also steered in scale.

1. INTRODUCTION

Points where several edges or lines meet are called junctions or keypoints. These points are important for object recognition and scene analysis (occlusions), because of their invariance properties and their high information content. Another important field of application is the matching of points in several images of the same scene, a problem that occurs in stereo vision, optical flow analysis, and multimodal medical imaging. Unlike isolated edges or lines which have 1D models these points have a local 2D structure. Yet standard edge detectors which are tuned to 1D edge models cannot detect or characterize 2D structures[3,8,9] . According to the many degrees of freedom of junctions, more information about the neighborhood of the keypoint must be obtained.

Corner detectors were developed, among others, by Deriche and Giraudon[3], Kitchen and Rosenfeld[6], Nagel[7], and Nobel[8] . In principle all these approaches use some differential geometric measures derived from a local second order Taylor approximation. Except Noble, who shows applications to T and Y junctions, the underlying models are restricted to corners (L junctions). In none of these methods the detected keypoints are classified. The classification of more complicated keypoints like T junctions or a mixture of lines and edges needs more information than is contained in a second order Taylor approximation. Rosenthaler[11] et.al. proposed a scheme based on oriented energy maps to detect keypoints of every kind, but again without a classification.

Work on the classification of keypoints has been done by Guiducci[5] , Rohr[10] , and Brunnström[2] et.al. . Guiducci estimates the parameters of a L junction (amplitude, aperture, smoothness of edges) from second order Taylor approximations. More complicated junctions are not dealt with. Rohr identifies junctions where several edges meet by fitting wedge models to the image. Brunnström et.al. evaluate the local histogram in the neighborhood of the keypoint to derive a hypothesis about the junction. The hypothesis is verified using the edge image, which is provided by a Canny-Deriche edge detector. A weakness of this edge detector is its poor orientation selectivity and that it gives no reliable response in the vicinity of keypoints[8,9] . Neither the method of Rohr nor the one of Brunnström et.al. is applicable to junctions of lines instead of wedges.

In this contribution we investigate the performance of steerable functions to characterize keypoints. Steerable functions were introduced recently by Perona[9] and are explained in more detail in the next section. We assume that the keypoints have already been detected by one of the methods mentioned above. As every other scheme, our approach has an underlying range of models it can handle. It will be confused by image structures which do not fit into this range. In section 4 the keypoints are junctions of step edges and thin lines with a rectangular profile. In

Substance Identification Analytics
SPIE Proceedings Series Vol. 2093
J.L. - Flanagan et al. (Eds.) 450-461, 1994

section 5 we add some blur and allow thick lines. All basis functions to which the image is projected are centered at the keypoint, so no convolution, even in a small neighborhood of the keypoint, is carried out.

2. STEERABLE FUNCTIONS

Functions which are used for convolution or projection in image processing usually have free parameters, e.g. their orientation and scale. It is often desirable not to fix these parameters, but to know the response of the function in a continuum of orientations, scales or other degrees of freedom. To avoid unreasonable computational costs this problem should be solved by interpolation. If F is the function and θ the free parameter a finite number N of basis functions $G_k, k = 1 \ldots N$ and interpolation functions b_k are searched, so that:

$$F_\theta(\vec{x}) = \sum_{k=1}^{N} b_k(\theta) G_k(\vec{x}) \tag{1}$$

Convolution and projections commute with this interpolation, because they are linear operators. If $I(\vec{x})$ is the image function then (the asterisk denotes complex conjugation): $\int F_\theta^*(\vec{x}) I(\vec{x}) d^2\vec{x} = \int (\sum_k b_k^*(\theta) G_k^*(\vec{x}))\ I(\vec{x}) d^2\vec{x} = \sum_k b_k^*(\theta) \int G_k^*(\vec{x}) I(\vec{x}) d^2\vec{x}$. Of course not every function can be approximated by a small or finite number of basis functions but one can ask for the best approximation with a given number of basis functions. A solution to this problem has recently been proposed by Perona[9] who uses the singular value decomposition (SVD) to obtain the functions G_k and b_k.

In the following, we give a short derivation of the formulas to steer a function in its orientation, thereby introducing all concepts and notations which are necessary to understand the following sections. The derivation differs slightly from the one given by Perona because we avoid arguing with Hilbert-Schmidt operators and the SVD. In the case of rotations, F_θ is a periodic function in θ and the interpolation problem can be solved by ordinary Fourier theory. This derivation is not applicable to the case of steering other parameters than the orientation. We will deal with this in section 5.

The functions b_k are now complex exponentials and (1) becomes the Fourier decomposition of F_θ with respect to the parameter θ.

$$F_\theta(\vec{x}) = \sum_{k=-\infty}^{\infty} A_k(\vec{x})\, e^{jk\theta} \tag{2}$$

$$A_k(\vec{x}) = \tfrac{1}{2\pi} \int_0^{2\pi} F_\theta(\vec{x})\, e^{-jk\theta} d\theta \tag{3}$$

Now the crucial point is that the infinite sum in (2) must be truncated at some point. We have to know the basis functions A_k which contribute the most to F in the L^2 sense. Calculating the norm of A_k gives (the hat denotes the Fourier transform):

$$\begin{aligned} \int A_k^* A_k d^2\vec{x} \quad &= \\ \frac{1}{4\pi^2} \int\int F_\theta^*(\vec{x})\, e^{jk\theta} d\theta \int F_{\theta'}(\vec{x})\, e^{-jk\theta'} d\theta' d^2\vec{x} \quad &\overset{(a)}{=} \quad \int\int\int F_\varphi^* F_0 d^2\vec{x}\, e^{jk\varphi} d\varphi d\theta' \\ \overset{(b)}{=} \quad \frac{1}{2\pi} \int h(\varphi)\, e^{jk\varphi} d\varphi \quad &= \quad \hat{h}(k). \end{aligned} \tag{4}$$

At (a) we changed the order of integration, introduced the new variable $\varphi := \theta - \theta'$ and exploited the fact that $\int F_\theta^* F_{\theta'} d^2\vec{x}$ only depends on $\theta - \theta'$. At (b) the integration with respect to θ' gives merely a factor of 2π because the

integrand is independent of θ'. As the result, we see that the contribution of A_k to F is given by the k'th Fourier coefficient $\hat{h}(k)$ of the function:

$$h(\varphi) = \int F_\varphi^* F_0 d^2\vec{x}. \tag{5}$$

This is not surprising because $h(\varphi)$ is the autocorrelation function of F with respect to θ which shows the angular width or influence of F. In wavelet theory it is called the reproducing kernel which governs the sampling scheme for complete wavelet bases (Antoine[1]). To get a better understanding of the basis functions A_k we rewrite (3).

$$\begin{aligned} A_k(\vec{x}) &= \frac{1}{2\pi}\int_0^{2\pi} F_\theta(\vec{x})\, e^{-jk\theta} d\theta \\ \overset{(a)}{=} \frac{1}{2\pi}\int_0^{2\pi} F_\theta(\rho,\phi)\, e^{-jk\theta} d\theta &\overset{(b)}{=} \frac{1}{2\pi}\int_0^{2\pi} F_0(\rho,\varphi)\, e^{-jk(\phi-\varphi)} d\varphi \\ = \frac{1}{2\pi} e^{-jk\phi}\int_0^{2\pi} F_0(\rho,\varphi)\, e^{jk\varphi} d\varphi &=: \frac{1}{2\pi} e^{-jk\phi} C_k(\rho) \end{aligned} \tag{6}$$

In (a) polar coordinates $\vec{x} \rightarrow \rho, \phi$ (thus ρ and ϕ are functions of $\vec{x}$) and in (b) a new variable $\varphi := \phi - \theta$ are introduced. Hence the A_k's are polar separable with the k'th harmonic as their angular part and a radial part which depends on F. Examples of such a decomposition are shown in Fig.5. Perona proves that this is the best steerable L^2 approximation to F with a given number of basis functions. This is a remarkable result, even though in our application the best L^1 approximations would be better. In the continuum case the approximated function is steerable without any unisotropy. In section 3.3, we show experimental results of the unisotropy in the discrete case.

It should be emphasized that the basis functions A_k are not rotated copies of the function F as they are in the work about steerable functions of Freeman and Adelson[4]. This fact has advantages as well as disadvantages. One disadvantage is that not all L^2 errors impair the performance to the same extent. In most cases, small deviations from the original function where it has the bulk of its energy do not disturb. Instead in the approximation (2) the missing A_k's of high angular frequency (large k) of the truncated sum put ripples to the approximation where the original function has no energy (see Fig.4d for an example). This produces spurious peaks in the angular response. Another disadvantage is that functions which need high frequency components for a good approximation are problematic because the high frequency basis functions cannot be rendered in the discrete case without interpolating the grid (see Fig.5d).

Nevertheless, these basis functions have one main advantage which makes them favorable for us. The symmetries of the function and the corresponding symmetries of the basis functions can be exploited (1) to save half of the basis functions (the negative frequencies) and (2) to steer, by the same basis functions, an one-sided kernel F^{1S} associated to F. If $F_0 \approx \sum_{k=-N}^{N} A_k$, then $F_0^{1S} \approx \sum_{k=0}^{N} A_k$. These functions are important in the analysis of keypoints in section 4 because of their 2π orientational periodicity. In reality the sum will not run from zero to N because we have to take the N A_k's with the largest L^2 norm. Thus $F \approx \sum_{k=0}^{N} A_{\nu_k}$, where $\|A_{\nu_0}\| > \|A_{\nu_1}\| > \ldots > \|A_{\nu_N}\|$. Examples of such one-sided functions are shown in Fig.2c,d and Fig.4c,d. In the following k is the index for the basis functions, ordered by their L^2 norm, and ν the angular frequency.

3. ORIENTATION SELECTIVE FUNCTIONS

3.1. Quality criteria for orientation selective functions

We use the following criteria to judge the quality of a function serving for orientation analysis.

- The main criterion is the reproducing kernel h, which is defined in (5) and its Fourier transform $\hat{h}$. As discussed in section 2, h reveals the orientation selectivity of the function. If $\Delta\theta$ is the half amplitude width

of h, approximately $2\pi/\Delta\theta$ rotated copies of h would be necessary to have a complete basis of orientation selective functions (Antoine[1]) if the method of Perona were not used. In the optimal case, h has no side lobes. Its Fourier transform $\hat{h}$ determines the number of basis functions that are necessary to steer the approximated function. Thus h has to optimize an uncertainty relation. $\hat{h}_s$ is a version of $\hat{h}$ that is sorted by the magnitude of the coefficients $\hat{h}(\nu)$ (Fig.4f).

- The response to an ideal step edge $g(\theta) := \int_{\mathbb{R}} \int_{x>0} F_\theta(x,y)dxdy$ (F is centered at $(0,0)$). The energy as well as the real and imaginary part are important. If the symmetry of an event is known, the real (even) or imaginary (odd) part is used to derive the exact orientation. The function should have a zero DC response, that is $g(\theta)$ has no offset from the baseline.
- The quality of the associated one-sided function F_0^{1S}: the energy of the one-sided function should be located mostly on one side of the center (steering point), giving it an angular periodicity of 2π. The functions h and g can be calculated for this function too.
- The extension of the function in position space should be small to have low computational costs. Of course there are limitations because of the size of the events and the discretization. A function of two pixels length cannot determine exact orientations.
- The limitations in the discrete case as discussed in section 3.3 have to be considered. It makes no sense to use a theoretically optimal function that loses its qualities in the discrete case.
- The quality of the approximation with a predefined number of basis functions has to be considered. In the L^2 sense it is given by $\hat{h}$, but not every L^2 error counts the same. Especially strong ripples in the approximation can be disadvantageous (see Figs.4d,10).

As an example we show the functions $h, \hat{h}$, and g for the first (G_1) and second (G_2) derivative of an isotropic Gaussian in Fig.1. G_1 is the kernel of the Canny edge detector. The half height width of $g(\theta)$ for G_1 is π, thus multiple orientations at one point cannot be resolved, even if they differ in their orientation by $\pi/2$. It should be evident that there is no single benchmark for the quality of an orientation selective function nor is there a function that is optimal for all the criteria. Different tasks will have different optimal functions.

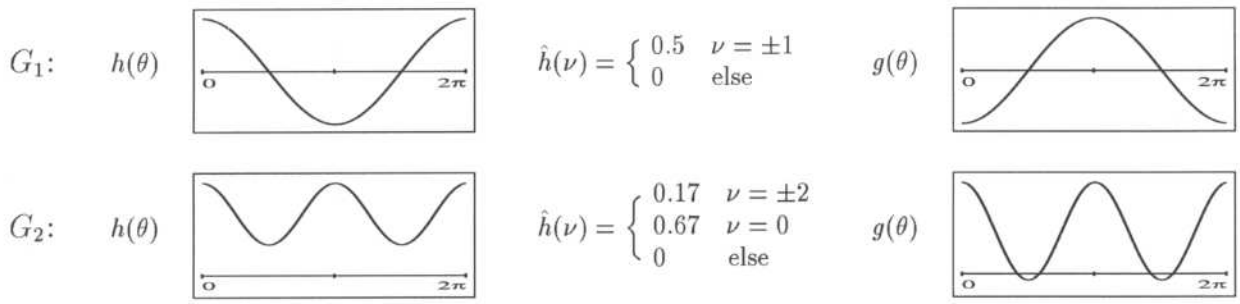


Figure 1: The reproducing kernel $h(\theta)$, its Fourier transform $\hat{h}(\nu)$, and the angular response to an edge $g(\theta)$ for the first (G_1) and second (G_2) derivative of an isotropic Gaussian. For G_1 the half amplitude width of h is $2\pi/2$ consistent with 2 nonzero Fourier components. Thus 2 basis functions are needed to steer G_1. Accordingly G_2 needs 3 basis functions. h and g show the very poor orientational resolution of both filters, making a detection of multiple orientations at one point impossible.

3.2. Double Hermite function

The example from Fig.1, where the functions are derived from an isotropic Gaussian, shows that we need elongated functions. Figure 2 gives an example, derived from an elongated Gaussian. This function has the shortcoming that the bulk of its energy is around the center. This has several drawbacks: (1) the one-sided function has its energy also around the center, (2) for the angular resolution the discretization at the keypoint is worst, (3) for many real life

keypoints the immediate neighborhood of the center point has a confusing structure. As a consequence, we choose a function that has the main part of its energy some distance away from the center. In the following we will refer to it as the 'double Hermite function'. The complete description of a function with the criteria of section 3.1 would need a wealth of data and we tried to figure out the most expressive data in Figs.3,4,5,6. There are functions that perform better in some of the mentioned criteria, e.g. a sharper $h(\theta)$ with a faster decay of $\hat{h}_s$, but with severe shortcomings in other criteria. We found the double Hermite function a good compromise.

$$\mathrm{H(x,y)} \;=\; \mathrm{N}\;\left(1-\left(\frac{\mathrm{y}}{\sigma}\right)^2+\mathrm{jy}\right)\;\mathrm{e}^{-\frac{\mathrm{y}^2}{2\sigma^2}}\;\left(\mathrm{e}^{-\frac{(\mathrm{x}-2\epsilon\sigma)^2}{2(\epsilon\sigma)^2}}+\mathrm{e}^{-\frac{(\mathrm{x}+2\epsilon\sigma)^2}{2(\epsilon\sigma)^2}}\right) \tag{7}$$

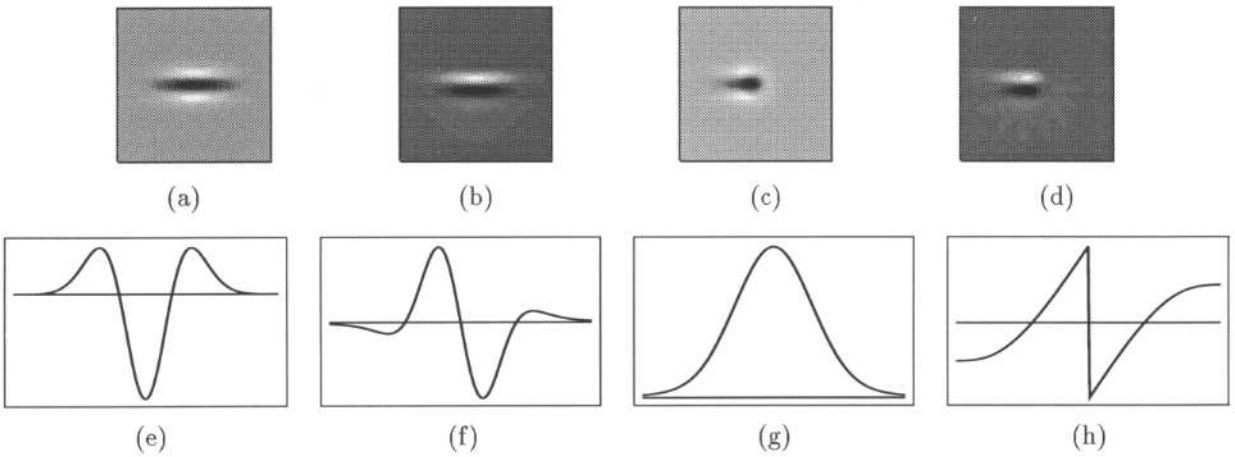


Figure 2: The function used by Perona for junction analysis. In x-direction it is a Gaussian, in y-direction it is for the real part (a) the second derivative of a Gaussian, for the imaginary part (b) the Hilbert transform of the real part. (c) and (d) show the associated one-sided function. It has the shortcoming that its energy is not at one side of the center. This is especially true for the real part where the center is inside the dark blob. (e) and (f) show a cross-section in y-direction of (a) and (b). (g) is the energy and (h) the phase of this cross-section.

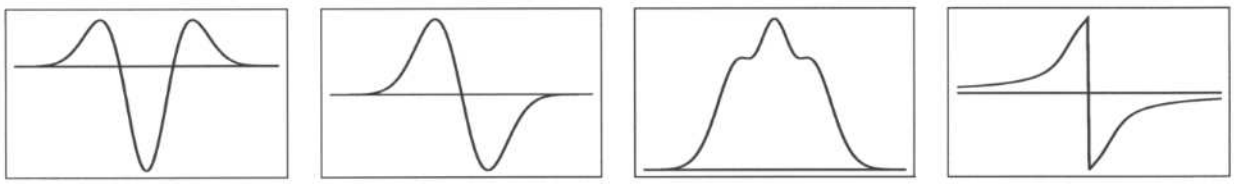


Figure 3: A cross-section in y-direction of the function from equation (7). The real and imaginary parts are not exact Hilbert partners, thus the phase is not linear and the energy is at the borderline of not being monomodal. In return the central peak is sharper than the peak for the function from Fig.2.

3.3 Limitations of steerability in the discrete case

In the continuous case the approximation to a function, given by a truncation of the sum in (2) is perfectly (isotropic) steerable. The quality of the approximation in the L^2 sense is governed completely by $\hat{h}$. In the discrete case both statements are not exactly true. Fig.6 shows the L^1 norm of a double Hermite function steered in orientation and scale (see section 5 for steering scale). As can be expected the error increases with a decreasing size of the function. The source of this error is the nonsymmetric sampling of the different orientations in a rectangular grid. For a basis function where the negative and positive parts are not symmetric with respect to the sampling grid (real part of A_8 in Fig.5b) this leads, besides other effects, to a nonvanishing DC-component. The DC-component

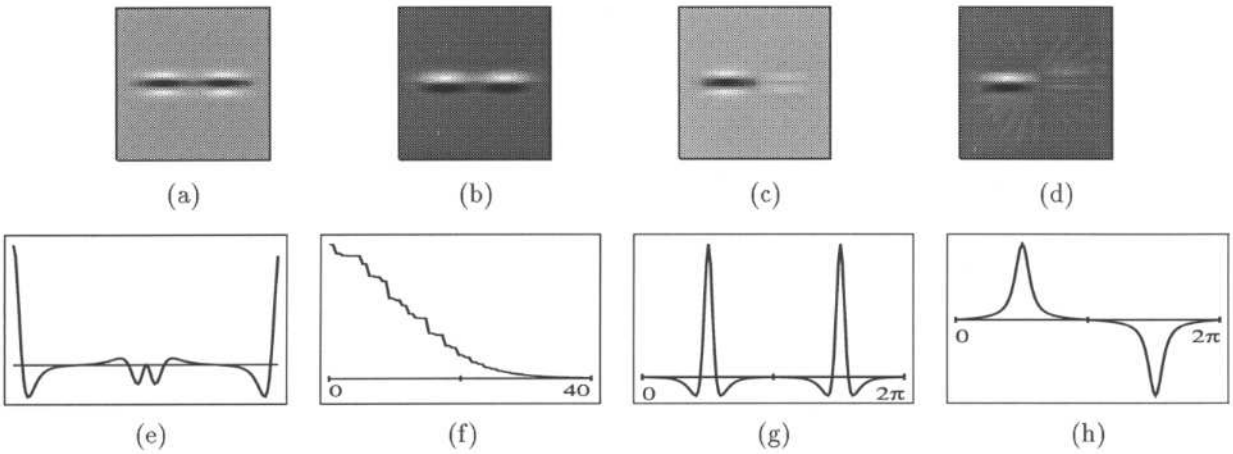


Figure 4: The function from equation (7). (a) real part, (b) imaginary part. (c) and (d) are the real and imaginary part of the associated one-sided function. The ring of ripples in (d) is the typical approximation error. (e) is the reproducing kernel $h(\theta)$, (f) is the sorted Fourier transform $h_s(k)$, that gives the L^2 norm of the basis functions A_k. The faster it decays, the less basis functions are needed for a good approximation. The angular response to a step edge, $g(\theta)$, is depicted in (g) (real) and (h) (imaginary).

of the approximated function can be zeroed by zeroing the DC-components of all basis functions, because the basis functions are fix, and the approximation is linear.

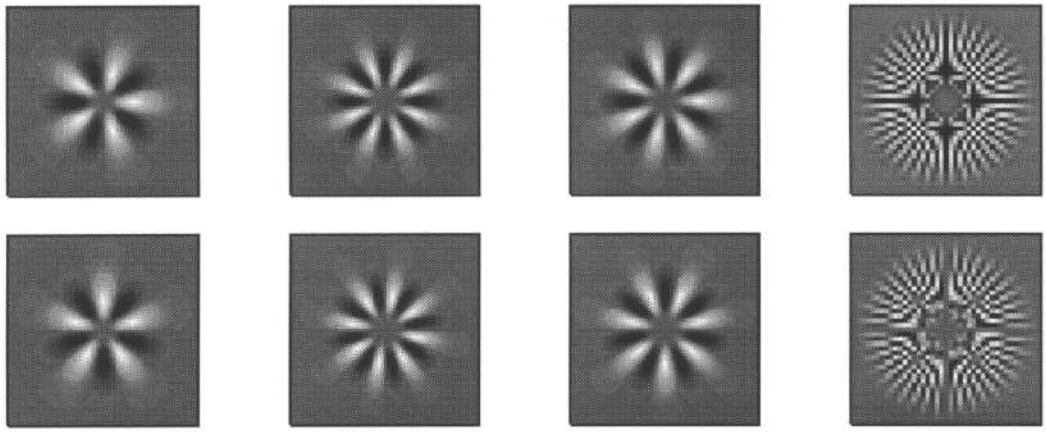
Another limitation arises when basis functions with large angular frequencies (large ν) are necessary for a good approximation. Such basis functions cannot be rendered properly near the origin because of their polar separability. Two consequences of this are (1) moiré effects as for the basis function in Fig.5d and (2) $\hat{h}_s$ does not converge to zero if it is calculated numerically. For the double Hermite function with $\varepsilon = 3$ that we used for the keypoint characterization in the following chapter, this is not a limiting factor. The largest angular frequency that is used to approximate it with 23 basis functions is $\nu = 24$ for $k = 22$. Nevertheless, functions with a higher orientational resolution can have this problem.

4. KEYPOINT CHARACTERIZATION

In this section, we apply the steerable functions to the characterization of keypoints. The keypoints are junctions of step edges and thin lines with a rectangular profile. We assume that the keypoint has already been detected by one of the methods from the references. It is important to locate the center of the function as close as possible at the real keypoint. The lines and edges are detected at the orientations of the maxima in the energy of the response. The evaluation of merely the maxima limits of course the resolution. The two main limiting effects are described in Fig.7.

The problems discussed in the caption to Fig.7 arise because of the linear superposition of the responses of all events that are 'seen' by the function. Looking to the response at just one orientation the information about the number of events which caused the response and their relative strength is lost. The evaluation of just the maxima can distinguish the events if for each event it is assured that there is an orientation where only this event is 'seen'. Of course, in the whole angular response there is more information. E.g. the shape and symmetry of the peaks could be analysed to improve the resolution. The drawback is that more sophisticated methods distinguish between different events in the range of the models by exploiting more knowledge about the models. Hence, because there are more assumptions about the analysed events, they are less robust when they are applied to events that are not in the range of models.

The phase of the response at the orientation of the maximum in the energy shows the symmetry type of the event. Lines (even events) have a phase near 0 or $\pm\pi$, edges (odd events) have a phase near $\pm\pi/2$. If the event is



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

Figure 5: (a),(b),(c) show the first three basis functions for the double Hermite function (real part top, imaginary part bottom). The angular frequencies are (a) $\nu = 5$, (b) $\nu = 8$, (c) $\nu = 7$. (d) shows an example of the moire effects for large angular frequencies ($\nu = 40$).

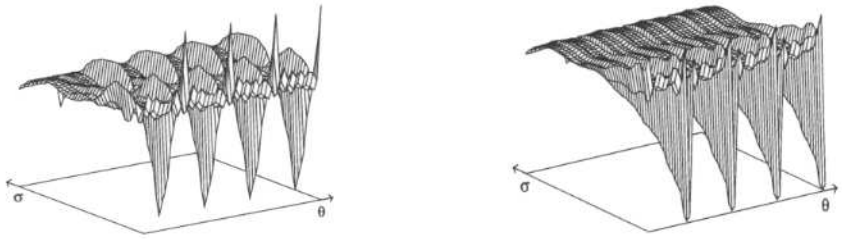


Figure 6: Limitations of perfect steerability in the discrete case: The pseudo 3D plots show the L^1 norm of the double Hermite function steered in orientation and scale (real part left, imaginary part right). Perfect steerability would result in a constant norm and the graph should be flat. The plateau is at $L^1 = 1$ and the largest deviations are -0.89 and +1.19 for the real part and -0.88 and +1.09 for the imaginary part. The maximal deviations are at multiples of 90°, because at this orientations the symmetry axses of the functions and the grid coincide. Parameters for the double Hermite function: $\sigma_{min} = 0.75, \sigma_{max} = 2, \varepsilon = 2$. For the reconstruction 23 basis functions in orientation and 5 in scale are used.

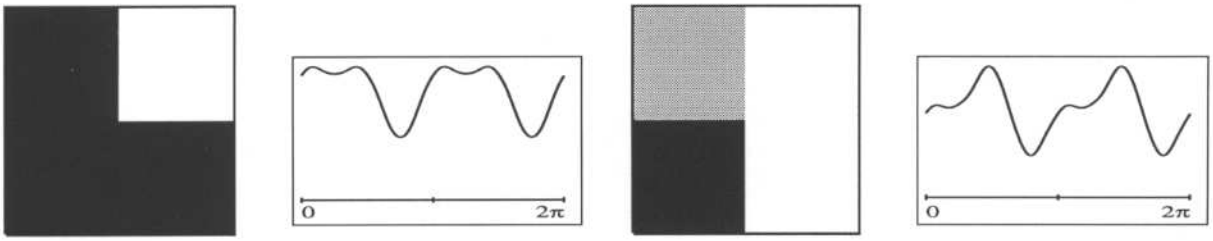
of one of these pure types it is preferable to determine the exact localization in the real or imaginary part of the response Fig.10.

For a function that is symmetric or anti-symmetric with respect to the center, the energy of the angular response to an edge at a certain orientation θ is the same as for $\theta + \pi$. Thus one cannot distinguish between lines or edges that terminate at the junction and lines or edges that do not terminate. This information is contained in the response of the one-sided function, because of its 2π periodicity. Remember that the calculation of this response needs no extra computational costs, because it is derived from the same basis functions.

The angular responses of Figs.8,9,10,11 were all derived with 23 basis functions. The orientation $\theta = 0$ is from the center to the right. Positive rotations are clockwise. To understand the one-sided responses note that the one-sided function has its energy to the left of the center (orientation π) (Fig.4). The effect of using less basis functions is shown in Fig.12.

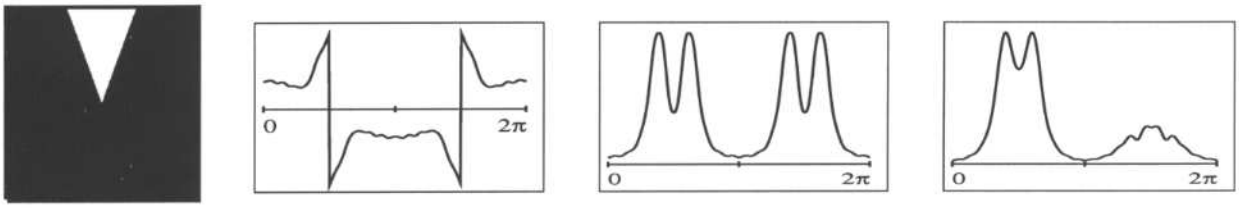
5. STEERING SCALE

In section 2 functions were steered only in their orientation. Yet there are other parameters, especially the scale, which are important in image processing. In the case of the scale the derivation from section 2 via the Fourier



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

Figure 7: Limitations of the angular resolution: (a) and (c) are the analysed junctions. (b) is the energy of the angular response to (a) and (d) the one to (c). If the two edges in (a) were a little closer in their orientation the two peaks would melt to one. The problem in (b) is that the small peak of the low contrast edge is swallowed by the large peak. In both cases some angular distance before the two peaks melt, the maxima are shifted. In this case the lines and edges can be detected, but the orientation estimation will be affected. The distinction of two weakly separated peaks as caused by two different events is impeded if the function has side lobes in its edge response function g or if its approximation has ripples (see Fig.4d, Fig.12).



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

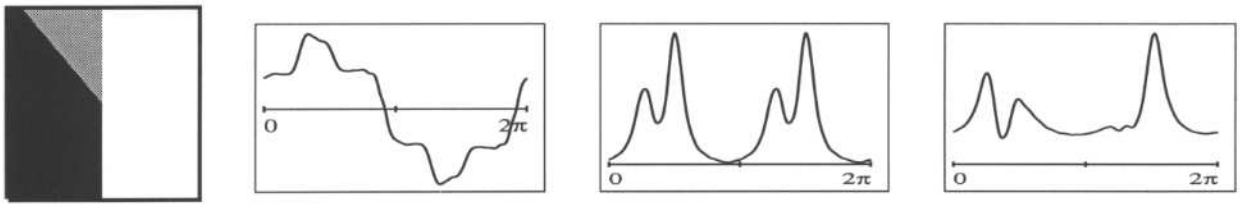
Figure 8: (a) keypoint, (b) phase, (c) energy, (d) energy for the one-sided function. The energy at the peaks shows two events at 70° and 110° (modulo 180°). The phase at the orientation of the maxima in the energy (1.78 and -1.78) reveals that they originate from two edges. The missing peaks at 250° and 290° in (d) show that the edges are terminating at the junction. Hence, the junction is a wedge.

transform is not applicable, but the general method of Perona[9], using the SVD, can handle this case. We will give here only a short sketch of the method to fix the notation and refer for the details to the reference.

To scale a function F we scale all its basis functions A_k. Because the A_k's are polar separable (see (6)) it suffices to scale the radial component $C_{ks}(\rho) := C_k(s\rho)$ (s is the scale parameter). Steering scale here again means interpolation as in (1). To arrive there we discretize ρ and s, and put all C_{ks}'s of different scales in one matrix as in the l.h.s. of (8) ($s_1 < s_2 < \ldots < s_m$).

$$\begin{pmatrix} C_{ks_1}(\rho_1) & \ldots & C_{ks_1}(\rho_n) \\ \vdots & \vdots & \vdots \\ C_{ks_m}(\rho_1) & \ldots & C_{ks_m}(\rho_n) \end{pmatrix} \overset{\text{(SVD)}}{=} (U_k)\,(W_k)\,(V_k^T) = \left(\sum_{i=1}^{n} w_i^k u_i^k(s_p) v_i^k(\rho_q) \right)_{p=1..m, q=1..n} \tag{8}$$

The rows of the l.h.s. matrix are the radial functions C at the different scales, so this matrix carries all information we need to scale F. The middle of (8) is the SVD of the matrix. (U) is a $m \times n$ matrix with orthonormal rows u_i, (V) is a $n \times n$ matrix with orthonormal rows v_i, and (W) is a diagonal matrix with the singular values w_i. The r.h.s. of (8) is just the middle written in another way and has the interpolation form of (1) where the v_i are the basis functions and the u_i the interpolation functions. Because of the L^2 normalization of the u_i and v_i the singular values w_i determine their L^2 contribution to the l.h.s. matrix of (8). To have a good approximation of the matrix, we need only a few components in the r.h.s. of (8). We used 128×128 matrices and the linpack library to perform the SVD.



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

Figure 9: (a) keypoint, (b) phase, (c) energy, (d) energy for the one-sided function. The energy at the peaks shows two events at 50° and 90° (modulo 180°). The phase at the orientation of the maxima in the energy (1.77 and 1.79) reveals that they originate from two edges. The maximum of the edge at 230° is slightly shifted by less than one degree because of the influence of the stronger edge at 90°. The missing second peak at 230° in (d) shows that one edge terminates at the keypoint. Hence, the keypoint is a T junction.

To steer the scale parameter s, the basis functions A_{ks} can now be decomposed into the new basis functions B_{ki}.

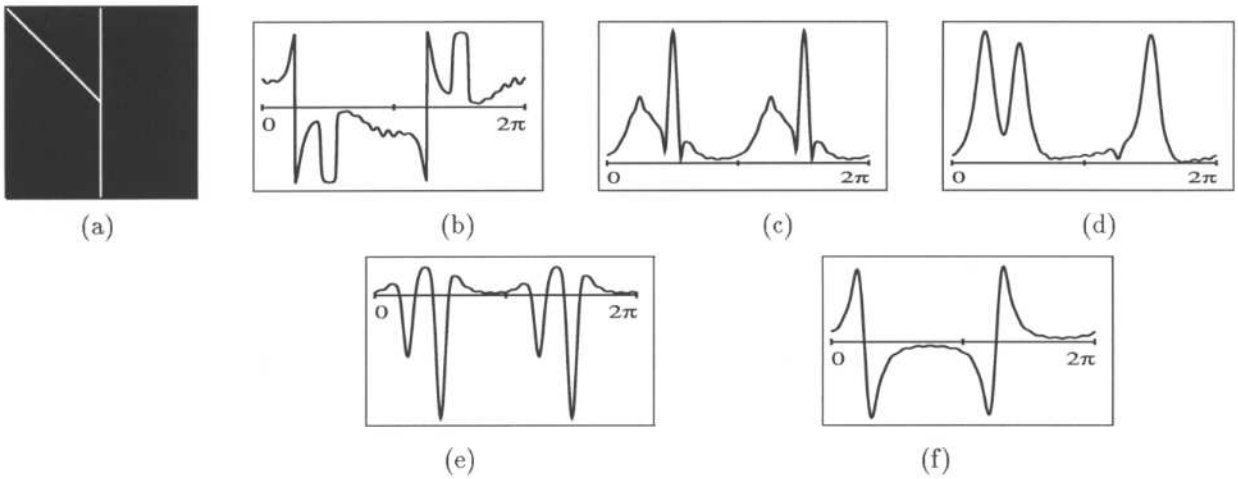
$$\begin{aligned} A_{ks}(\vec{x}) &= \frac{1}{2\pi} e^{jk\phi(\vec{x})} C_k(s\rho(\vec{x})) \quad \approx \\ \frac{1}{2\pi} e^{jk\phi(\vec{x})} \sum_{i=1}^{M} w_i^k u_i^k(s) v_i^k(\rho) &= \frac{1}{2\pi} \sum_{i=1}^{M} w_i^k u_i^k(s) B_{ki}(\vec{x}) \\ \text{where} \quad B_{ki}(\vec{x}) &= e^{jk\phi(\vec{x})} v_i^k(\rho(\vec{x})) \end{aligned} \tag{9}$$

For most functions F, which are of interest very few basis functions suffice to steer it in the scale. This is due to the large variances of their autocorrelation function (reproducing kernel) with respect to the scale $h(s) := \int F_s F_{s_0} d^2\vec{x}$ (s_0 is a reference scale). Figure 13 shows $h(s)$ for the Hermite function and the double Hermite function. The usual range of interest in the scale spans 2 to 4 octaves while the half amplitude width of $h(s)$ is typically about 2 octaves. This is a ratio of about 2, a relatively small number compared to the same ratio in the orientational case which is about 15 for the function we used. Moreover, for step edges the scale is not a sensitive parameter. Accordingly, the energy distribution against orientation and scale for the junction from Fig.14a reveals no strong dependence on the scale. In addition, we are usually not interested in an exact analysis of the response in scale but merely to find an appropriate scale for the orientation analysis. Hence, the common discretization of the scale in octave steps is not as disadvatageous as discretizing the orientation.

Nevertheless, one can take benefit from steering the scale for junctions of lines of different widths and blurred edges. The function F can be adapted in its size to improve the angular response and the width of the lines and edges can be estimated. Results for the application of steerable functions to keypoints of blurred edges and thick lines are shown in Fig.14.

6. CONCLUSIONS

We have shown that a wide range of keypoints can be characterized by steerable functions. Compared to previous work by Guiducci, Rohr, and Brunnström our approach improves in handling edge as well as line junctions. Without making the functions unreasonably large, it is not possible to distinguish two lines or edges which differ only by a small angle in their orientation. To obtain a better resolution, it would be possible to evaluate not just the extrema in the angular response signals. One could use as well the shape, width, and symmetry of the peaks for the classification, but with the drawback of false results with image structures which do not fit in the range of models. A more robust method would require more than just the evaluation of basis functions centered at the keypoint, but also in a small neighborhood. Because of the high computational costs due to the relatively large set of basis functions, this method is not intended to be applied by convolution to the whole image. Instead some preceding keypoint detection or focus of attention is required.



(a) (b) (c) (d) (e) (f)

Figure 10: (a) keypoint, (b) phase, (c) energy, (d) energy for the one-sided function, (e) real and (f) imaginary part of the response. The energy at the peaks shows two events at 45° and 90° (modulo 180°). The phase at the orientation of the maxima in the energy reveals (jump from $+\pi$ to $-\pi$, -3.08) that they originate from two lines. The missing second peak of the line at 225° in (d) shows, that it terminates at the keypoint. The leftmost peak in (b) is relatively wide and not well shaped. In this case it is the influence of the second line and not caused by a third event, but one cannot a priori know this. The phase at the maximum of the energy at 45° has a jump from $+\pi$ to $-\pi$, indicating a line. Hence, the real part of the response (e) is appropriate for the localization and has a better separation of the two peaks.

7. REFERENCES

1. J.-P. Antoine, P. Carrette, R. Murenzi, and B. Piette, "Image analysis with two-dimensional continuous wavelet transform", Signal Processing, Vol.31, 241-272, 1993.

2. K. Brunnström, J.-O. Eklundh, and T. Lindeberg, "On scale and resolution in active analysis of local image structure", Image and Vision Computing, Vol.8, No.4, 289-296. 1990.

3. R. Deriche and G. Giraudon, "Accurate corner detection: An analytical study", INRIA, Technical Report No.1420, 1991.

4. W.T. Freeman and E.H. Adelson, "The design and use of steerable filters", IEEE PAMI-13, No.9, 891-906, 1991.

5. A. Guiducci, "Corner characterization by differential geometry techniques", Pattern Recognition Letters, Vol.8, 311-318, 1988.

6. L. Kitchen and A. Rosenfeld, "Gray-level corner detection", Pattern Recognition Letters, Vol.1, 95-102, 1982.

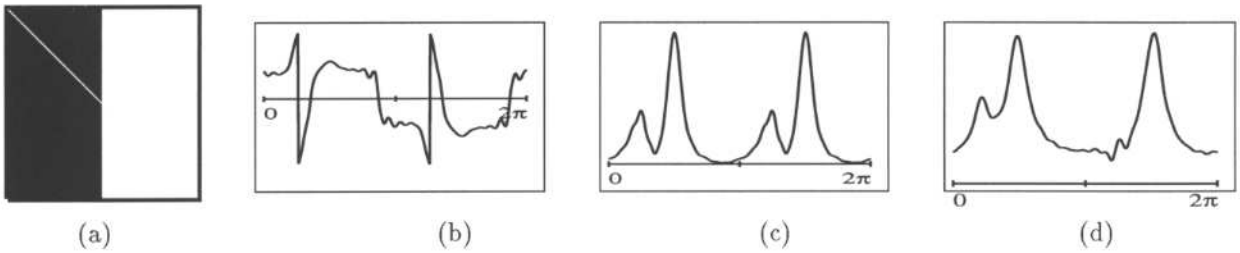
7. H.-H. Nagel, "Displacement vectors derived from second-order intensity variations in image sequences", CVGIP, Vol.21, 85-117, 1983.

8. J.A. Noble, "Finding corners", Image and Vision Computing, Vol.6, No.2, 121-128, 1988.

9. P. Perona, "Steerable-scalable kernels for edge detection and junction analysis", ECCV 92, 3-18, 1992.

10. K. Rohr, "Recognizing corners by fitting parametric models", Int.J. of Computer Vision, Vol.9, No.3, 213-230, 1992.

11. L. Rosenthaler, F. Heitger, O. Kübler, and R. von der Heydt, "Detection of general edges and keypoints", ECCV 92, 78-86, 1992.



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

Figure 11: (a) keypoint, (b) phase, (c) energy, (d) energy for the one-sided function. The energy at the peaks shows two events at 44° and 90° (modulo 180°). The phase at the orientation of the maxima in the energy (3.00 and 1.77) reveals that they originate from a line at 44° and an edge at 90°. Due to the influence of the dominating edge at 90° the maximum of the line is shifted by less than one degree, while the jump in the phase ($+\pi$ to $-\pi$) is slightly shifted in the opposite direction. The missing second peak of the line at 225° in (d) shows, that it terminates at the keypoint.

Figure 12: The main effect of using less basis functions for the approximation is not an increasing width of the peaks but stronger ripples in the response. The energy responses to the keypoints of Fig.8 (left) and Fig.11 (right) are depicted. Only 15 basis functions were used for the approximation. Using less than 15 basis functions leads to many new peaks which cannot be distinguished from the peaks caused by lines or edges.

Figure 13: Reproducing kernel in scale $h(\sigma)$ for the Hermite function (left) and the double Hermite function (right). The numbers at the scale axes are in units of the reference scales σ_0 and l_0. The range of displayed scales is 4 octaves. The parameters are for the Hermite function: $\sigma_0 = 4, \sigma_{min} = 1, \sigma_{max} = 16, \varepsilon = 3$, double Hermite function: $\sigma_0 = 4, \sigma_{min} = 1, \sigma_{max} = 16, \varepsilon = 2.2$, The half amplitude widths are (in octaves) (a) 2.2, (b) 1.7. Compared to the range of interest in the scale, which is not more than 3 or 4 octaves for many tasks, the width is large.

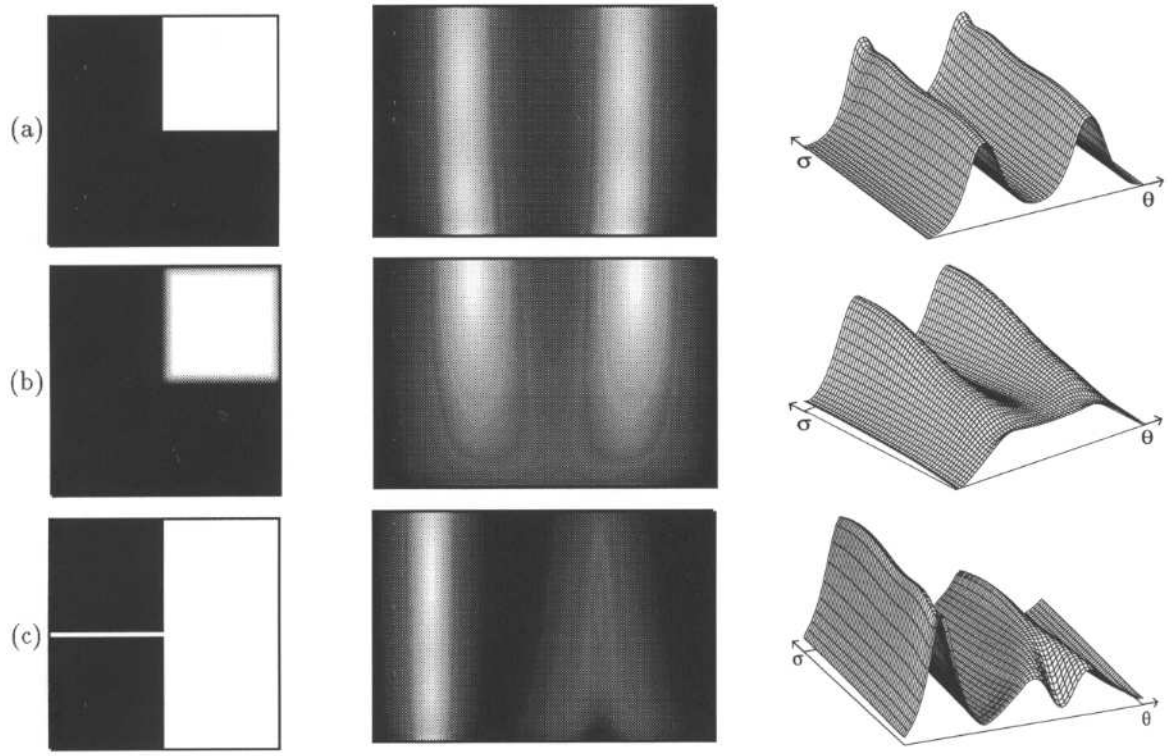


Figure 14: The figure shows the results of steering the double Hermite function in energy and scale. Depicted are from left to right the keypoint, the energy of the response as an intensity image (small scales at the top), and and the energy as an pseudo 3D plot. In (a) the keypoint is a corner of step edges and the response is approximately independent of the scale. The energy image displays the orientations from 135° to 315°. The junction in (b) is the same as in (a) but blurred by an Gaussian with $\sigma = 3$. At small scales the edges are not seen The junction in (c) is a step edge and a line of 5 pixels width. The energy is displayed in the range from 120° to 300°. At small scales the line is seen as two edges with an incorrect orientation. The continuous scale energy image allows not only to detect the line properly at coarse scales, but it gives also the information how the events at small scales are connected to those at coarse scales.