Integrated Edge and Junction Detection with the Boundary Tensor

Ullrich Köthe

University of Hamburg, Cognitive Systems Lab
koethe@informatik.uni-hamburg.de

Abstract

The boundaries of image regions necessarily consist of edges (in particular, step and roof edges), corners, and junctions. Currently, different algorithms are used to detect each boundary type separately, but the integration of the results into a single boundary representation is difficult. Therefore, a method for the simultaneous detection of all boundary types is needed. We propose to combine responses of suitable polar separable filters into what we will call the boundary tensor. *The trace of this tensor is a measure of boundary strength, while the small eigenvalue and its difference to the large one represent corner/junction and edge strengths respectively. We prove that the edge strength measure behaves like a rotationally invariant quadrature filter. A number of examples demonstrate the properties of the new method and illustrate its application to image segmentation.*

1. Introduction

Boundary detection is an important task in image analysis since boundaries provide crucial cues for object recognition and scene understanding. In this paper, we will be concerned with low-level, bottom-up boundary detection. From topology we know that a boundary in the plane necessarily consists of edges *and* junctions [12]. Edges in turn come in two kinds: step edges and roof edges (the latter ones also termed lines, bars, ridges), whereas junctions can be distinguished by their degree – junctions of degree two are called corners or L-junctions, of degree three T- and Y-junctions (according to their geometrical layout), and of degree four X- and Ψ-junctions. Psychological studies [1, 10] reinforce that *all* these boundary features are important.

Currently, it is common to use different algorithms for detecting each type of boundary feature separately. For example, one can use Canny's algorithm [3] for step edges, eigenvalues of the Hessian matrix for roof edges, and the corner response function [9] for corners. These detectors yield stable, rotationally invariant results for their respective feature types. But for a complete boundary description the different features types have to be *integrated*. This integration is not straightforward. Simple ideas like linear combinations of the responses of several detectors don't work:

First, it's impossible to find weights for the combination that work equally well accross the entire image or a set of images. Second, each detector generates spurious responses where it encounters a kind of border feature it wasn't designed for. For example, a first derivative detector (designed for step edges) yields a bimodal response when applied to a roof edge. In the vicinity of junctions it often leaves gaps or hallucinates non-existing edges. Similarly, most corner detectors find L-shaped corners, but create multiple or no responses for more complex junctions. The spurious responses of each detector for features of the "wrong" kind show up as artifacts in the combined signal. This makes a reliable interpretation of the combination very difficult.

Therefore, the responses of each operator are usually transformed into a symbolic form (e.g. edgels and corner points) and then combined by some grouping procedure. In the most common case, only edgels are considered. By means of proximity and continuity criteria, edgels are linked into a boundary graph. Grouping heuristics are needed to close gaps at junctions, merge nearby double edges resulting from lines etc. [2,15]. Unfortunately, these heuristics often make mistakes that lead to unstable or erroneous boundary graphs. Combining edge and corner detectors doesn't help either, because the corners signalled by standard corner detectors are severely displaced from their true locations (typically 2-3 pixels [14]) so that again unreliable heuristics are needed to tell which corner belongs to which edge [4].

In this paper, we propose the *boundary tensor* as a new approach to integrated boundary detection. It shall have the following properties: (i) It provides rotationally invariant edge and corner/junction strength measurements whose sum is an integrated boundary strength. (ii) It allows the estimation of the orientation and sub-pixel location of edges. (iii) It reacts uniformly to both step edges and roof edges.

2. Prior Work

To solve the integration problem, it would be most natural to combine features at the signal level. One possibility for signal level integration is the *quadrature filter approach*, e.g. [8]. Here, the responses of a pair of even and odd symmetric filters are combined into an edge energy which has the same magnitude irrespective of whether there was a step

or a roof edge. In order for the integration to work, the filters must be related by the *Hilbert transform*:

$$\mathcal{F}[k_{\text{odd}}] = -j\,\text{sign}(u)\mathcal{F}[k_{\text{even}}] = -j\frac{u}{|u|}\mathcal{F}[k_{\text{even}}] \qquad (1)$$

where $\mathcal{F}$ denotes the Fourier transform, u is the frequency coordinate, and j the imaginary unit. However, this formula only applies to 1-dimensional signals. There are several ways to generalize it to 2D, but none of them is completely satisfying. The most common one is to use an oriented quadrature filter pair that is locally tuned to the orientation of the edge under investigation. Orientation tuning can be realized very efficiently by means of steerable filters [7]. Quadrature filters elegantly solve the integration problem for step end roof edge detection, but leave open the corner/junction problem. At a corner/junction location, there often is no clearly defined local orientation, so that orientation tuning becomes unstable.

A possibility to detect edge and corner information simultaneously is the use of 2nd *order tensors*. In image analysis, the structure tensor is the most popular one:

$$\text{StructureTensor} = \begin{pmatrix} \overline{g_x^2} & \overline{g_x g_y} \\ \overline{g_x g_y} & \overline{g_y^2} \end{pmatrix} \qquad (2)$$

where g_x and g_y are the components of the gradient of the image (calculated by a suitable gradient filter) and the bar denotes spatial averaging (usually by means of a Gaussian filter). The trace of the tensor encodes boundary strength, and its orientation is perpendicular to the local edge orientation. While the gradient itself encodes only step edge information, the averaging distributes this information over a neighborhood, and points that receive contributions from edges with multiple orientations are considered junctions. This can be decided by looking at the tensor's eigenvalues: if they are about equal, the local structure is "intrinsically 2-dimensional", i.e. a corner or junction, otherwise it is "intrinsically 1-dimensional", i.e. an edge.

While this approach is very useful for corner detection [9,14], it also has several disadvantages: First, it cannot handle roof edges. Second, it tends to give multiple responses for junctions with degree higher than 2. Third, the isotropic smoothing employed by definition (2) does not only produce the desirable corner responses, but also leads to an undesirable blurring of the edges. The last problem could potentially be solved by means of *unisotropic tensor diffusion* [16], but we are not aware of any systematic investigation of this possibility in the structure tensor context. As to the first problem, Granlund and Knutsson [8] proposed a method to derive tensors from the responses of a family of oriented quadrature filters. These tensors encode step and roof edge information simultaneously. Their behavior on edges has been intensively studied, but the suitability for corner/junction detection has apparently never been investigated.

3. Mathematical Preliminaries

Since we want to derive rotationally invariant boundary strength measures and estimate edge orientation, we need a well-defined way to describe the behavior of our image measurements under rotations. This makes 2-dimensional Cartesian tensors a natural representation choice. A 2^p tuple of real numbers $[T_{i_1 \dots i_p}, \quad i_k = 1, 2]$ is called a Cartesian tensor of order p in 2-dimensional space, if a Euclidean coordinate transform (rotation and translation) induces the following transform of the tensor elements (see e.g. [11]):

$$\tilde{T}_{i_1 \dots i_p} = \sum_{j_1=1}^{2} \cdots \sum_{j_p=1}^{2} s_{i_1 j_1} \dots s_{i_p j_p} T_{j_1 \dots j_p} \qquad (3)$$

where s_{ij} are the elements of the 2D rotation matrix:

$$(s_{ij}) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix} \qquad (4)$$

In particular, tensors of order 0 (also called *scalars*) are rotationally invariant. Linear combinations of tensors of equal order are again tensors. New tensors can also be formed by means of the *tensor product* and *tensor contraction*:

$$T_{i_1 \dots i_{p+q}} = U_{i_1 \dots i_p} V_{i_{p+1} \dots i_{p+q}}$$
$$T_{i_1 \dots i_{m-1} i_{m+1} \dots i_{n-1} i_{n+1} \dots i_p} =$$
$$\sum_{j=1}^{2} T_{i_1 \dots i_{m-1} j\, i_{m+1} \dots i_{n-1} j\, i_{n+1} \dots i_p}$$

The first operation transforms two tensors of order p and q into a tensor of order $p+q$, while the second creates a tensor of order $p-2$ from a tensor of order p. Many common operations, such as the squared norm of a vector or matrix, the matrix product, and the trace of a matrix, can be expressed as sequences of tensor products and contractions.

Tensors of order 2 are particularly suited for the simultaneous analysis of intrinsically 1D and 2D image structure. A positive semi-definite symmetric tensor of order 2

$$T = \begin{pmatrix} t_{11} & t_{12} \\ t_{12} & t_{22} \end{pmatrix}$$

has two non-negative eigenvalues

$$\lambda_{1,2} = \frac{1}{2} \left(t_{11} + t_{22} \pm \sqrt{(t_{11} - t_{22})^2 + 4t_{12}} \right) \qquad (5)$$

which encode the magnitudes of the quantity of interest in the directions given by the corresponding orthogonal eigenvectors $\vec{e}_1$ and $\vec{e}_2$. The orientation of $\vec{e}_1$ is

$$\psi = \frac{1}{2} \arctan \left(\frac{2t_{12}}{t_{11} - t_{22}} \right) \qquad (6)$$

T can be decomposed into intrisically 1D (edge) and 2D (junction) parts as follows:

$$T = T_{\text{edge}} + T_{\text{junction}} = (\lambda_1 - \lambda_2)\vec{e}_1\vec{e}_1^T + \lambda_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad (7)$$

$\lambda_1 - \lambda_2$ can be interpreted as an edge strength, and $2\lambda_2$ as a junction strength. Formally, these quantities are obtained by contraction of T_{edge} and T_{junction} along the main diagonal, i.e. by the matrix trace. They are thus rotationally invariant.

In order to construct tensors we have to combine raw filter responses in such a way that the transformation property (3) is fulfilled. Thus, the filters must possess suitable *angular* behavior. In contrast, requirement (iii) from the introduction – uniform reaction to step and roof edges – can be considered a 2D generalization of the quadrature property (1). This is esentially a constraint on the filters' frequency behavior, e.g. their *radial* shape. Since the angular and radial shapes should be optimized independently, it is natural to use *polar separable filters*, e.g. products of angular and radial functions. To facilitate optimization of the frequency behavior, we define these filters in the Fourier domain:

$$K(\rho, \varphi) = K_\varphi(\varphi) K_\rho(\rho) \tag{8}$$

where ρ, ϕ denote polar coordinates in the Fourier domain. Filters whose angular part is proportional to $\sin(n\varphi)$ or $\cos(n\varphi)$ are called *polar harmonic filters*. They are separable in both the spatial and Fourier domains, with the same angular function. Their inverse Fourier transforms are:

$$\mathcal{F}^{-1}\left[\cos(n\varphi)K_\rho(\rho)\right](r, \vartheta) = \frac{j^n}{4\pi^2}\cos(n\vartheta)k_r^{(n)}(r)$$

$$\mathcal{F}^{-1}\left[\sin(n\varphi)K_\rho(\rho)\right](r, \vartheta) = \frac{j^n}{4\pi^2}\sin(n\vartheta)k_r^{(n)}(r)$$

where r, ϑ are polar coordinates in the spatial domain. The radial functions $k_r^{(n)}(r)$ do not only depend on $K_\rho(\rho)$, but also on the order of the angular harmonic. They are obtained by the nth-order Hankel transform $\mathcal{H}_n$:

$$k_r^{(n)}(r) = \mathcal{H}_n[K_\rho(\rho)] = 2\pi \int_0^\infty K_\rho(\rho) J_n(2\pi r\rho)\, \rho \, d\rho$$

where $J_n(t)$ is the nth order Bessel function of the first kind:

$$J_n(t) = \frac{1}{2\pi}\int_{-\pi}^\pi e^{j(t\sin\phi - n\phi)}d\phi$$

A set of polar harmonic separable filters of the form $\{(-j)^n\cos(n\varphi)K_\rho(\rho), (-j)^n\sin(n\varphi)K_\rho(\rho); n = 0, 1, ...\}$ will be called a *polar filter family*. The factor $(-j)^n$ is introduced to make all filters real in the spatial domain, provided K_ρ is real. The filters in a family are angular modulations of the *same* base filter, i.e. have the same frequency behavior. The two filters with the same n form a pair. The result of convolving an image f with a polar harmonic filter pair will be called *polar channel* $C^{(n)}$:

$$C^{(n)} = \begin{pmatrix} C_1^{(n)} \\ C_2^{(n)} \end{pmatrix} = \begin{pmatrix} f \star \left(\frac{1}{4\pi^2}\cos(n\vartheta)k_r^{(n)}(r)\right) \\ f \star \left(\frac{1}{4\pi^2}\sin(n\vartheta)k_r^{(n)}(r)\right) \end{pmatrix} \tag{9}$$

($\star$ is the convolution symbol). Channel 0 is simply defined as $C^{(0)} = f \star \frac{1}{4\pi^2}k_r^{(0)}(r)$. The set of polar channel responses up to order n will be called the *polar n-jet* of the image f. This term stresses the similarity to the standard n-jet introduced by Koenderink et al. [6] which, however, is based on derivatives rather than angular modulations of a base filter.

If the coordinate system is rotated, the polar channels transform according to the *steering equations* (cf. [7]):

$$\begin{aligned} C_{1,\vartheta_0}^{(n)} &= \quad\cos(n\vartheta_0)C_1^{(n)} + \sin(n\vartheta_0)C_2^{(n)} \\ C_{2,\vartheta_0}^{(n)} &= -\sin(n\vartheta_0)C_1^{(n)} + \cos(n\vartheta_0)C_2^{(n)} \end{aligned} \tag{10}$$

where $C_{1,\vartheta_0}^{(n)}$ and $C_{2,\vartheta_0}^{(n)}$ denote the components of the rotated channel n. It should be noted that steerability is independent of how the radial part of the filters is defined. It is only necessary that the two filters of the same order (i.e. in the same pair) have identical radial functions.

4. Definition of the Boundary Tensor

4.1. Tensors from Polar Filter Responses

Before we define the boundary tensor, we want to introduce a number of simpler tensors that are created directly from the polar filter responses. The simplest such tensor is the channel 0 response itself: Since $\cos(0)k_r^{(0)}$ is a rotationally invariant filter, $C^{(0)}$ is a 0th order tensor. If the radial function K_ρ in the Fourier domain is a bandpass, the channel 0 filter can be interpreted as a generalized *Laplacian operator*. We get the standard Laplacian of Gaussian by further restricting the bandpass shape to $K_\rho = -\rho^2\exp(\rho^2\sigma^2/2)$. A generalized Laplacian yields a high magnitude at points of high mean curvature: roof edges and local extrema.

According to the steering equations (10), the components of channel 1 fulfill the requirements of an order 1 tensor. Since $\frac{\partial}{\partial x} = \cos(\vartheta)\frac{\partial}{\partial r}$ and $\frac{\partial}{\partial y} = \sin(\vartheta)\frac{\partial}{\partial r}$, the first derivatives of some radially symmetric kernel k_r are special cases of channel 1 kernels with $k_r^{(1)} = \frac{\partial}{\partial r}k_r^{(0)}$. In other words, channel 1 generalizes the *gradient vector*, and the special choice $K_\rho = -j\rho\exp(\rho^2\sigma^2/2)$ yields the Gaussian gradient. Therefore, channel 1 mainly responds at points of high contrast, i.e. step edges.

The components of the channel 2 response can be used to define a tensor of order 2:

$$T = \begin{pmatrix} C_1^{(2)} & C_2^{(2)} \\ C_2^{(2)} & -C_1^{(2)} \end{pmatrix} \tag{11}$$

The proof of the tensor property is straightforward when the rotated channel responses are expanded by means of the steering equations. We can again relate this to a well-known differential quantity, namely the *Hessian matrix*, by adding

$C^{(0)}$ to the diagonal elements:

$$T^{(2)} = \begin{pmatrix} C^{(0)} + C_1^{(2)} & C_2^{(2)} \\ C_2^{(2)} & C^{(0)} - C_1^{(2)} \end{pmatrix} \qquad (12)$$

This is a tensor since $T + C^{(0)} I$ is a linear combination of tensors (I is the unit tensor). The relationship to the Hessian can be seen by defining second derivative filters in the Fourier domain. According to the derivative theorem of Fourier theory, the horizontal second derivative of a rotationally symmetric spatial filter w_r is:

$$\begin{aligned} \mathcal{F}\left[\frac{\partial^2}{\partial x^2} w_r\right] &= -\cos(\varphi)^2 \rho^2 W_\rho(\rho) \\ &= -\frac{1}{2}(1 + \cos(2\varphi))\rho^2 W_\rho(\rho) \end{aligned} \qquad (13)$$

where $W_\rho(\rho)$ denotes the Fourier transform of w_r. Inverse Fourier transform yields

$$\frac{\partial^2}{\partial x^2} w_r = \frac{1}{8\pi^2}\left(\hat{w}_r^{(0)} + \cos(2\varphi)\hat{w}_r^{(2)}\right) \qquad (14)$$

where $\hat{w}_r^{(0)}$ and $\hat{w}_r^{(2)}$ are the 0th and 2nd order Hankel transforms of $-\rho^2 W_\rho(\rho)$ respectively. The result of convolution with this filter is obviously a special case of the expression $C^{(0)} + C_1^{(2)}$ with the appropriate choices of radial functions. Analogous arguments apply to the other components of $T^{(2)}$. The generalized Hessian is sensitive to points of high Gaussian curvature, i.e. roof edges, local extrema and saddle points. In a similar way, one can define higher order tensors, but these will not be needed in this paper.

4.2. Rotationally Invariant Quadrature Filters

The *energy* in polar channel n is defined as the sum of the squares of the two filter responses in the channel. If the coordinate system is rotated, the transformation of the squared filter responses is given by the steering equations (10):

$$\begin{aligned} \left(C_{1,\vartheta_0}^{(n)}\right)^2 =\ &\cos(n\vartheta_0)^2 \left(C_1^{(n)}\right)^2 + \sin(n\vartheta_0)^2 \left(C_2^{(n)}\right)^2 \\ &+ 2\cos(n\vartheta_0)\sin(n\vartheta_0)C_1^{(n)}C_2^{(n)} \\ \left(C_{2,\vartheta_0}^{(n)}\right)^2 =\ &\sin(n\vartheta_0)^2 \left(C_1^{(n)}\right)^2 + \cos(n\vartheta_0)^2 \left(C_2^{(n)}\right)^2 \\ &- 2\cos(n\vartheta_0)\sin(n\vartheta_0)C_1^{(n)}C_2^{(n)} \end{aligned}$$

By adding both equations, we see that the energy E_n in every channel is rotationally invariant, irrespective of the radial function $k_r^{(n)}$:

$$E_n = \left(C_{1,\vartheta_0}^{(n)}\right)^2 + \left(C_{2,\vartheta_0}^{(n)}\right)^2 = \left(C_1^{(n)}\right)^2 + \left(C_2^{(n)}\right)^2$$

The meaning of these energies can be derived from the tensor properties discussed in the previous section. $E_0 = (C^{(0)})^2$ is the squared response of a generalized Laplacian operator and indicates local extrema and roof edges. $E_1 = (C_1^{(1)})^2 + (C_2^{(1)})^2$ is a generalized squared gradient magnitude and thus indicates step edges. In case of E_2, a simple calculation reveals that $(C_1^{(2)})^2 + (C_2^{(2)})^2 = (\lambda_1 - \lambda_2)^2$, where $\lambda_{1,2}$ are the eigenvalues of the tensor $T^{(2)}$. E_2 is large when the squared difference of these eigenvalues is large, i.e. when one eigenvalue is large and the other is small (roof edge), or both are large and have opposite signs (saddle point). In other words, the energies up to order 2 encode many of the most important structural image features.

We show now that certain linear combinations of polar energies behave like quadrature filters. Quadrature filters are characterized by the fact that their energy has a single peak at an edge, irrespective of whether the edge exhibits even or odd symmetry (i.e. is a roof or step edge). To generalize this to 2D, we define "intrinsically 1D images" by the requirement that all sections in a certain direction $\vec{n}$ are constant, whereas all sections perpendicular to this direction have the same characteristic profile, i.e. $f(\vec{x}) = g(\vec{x}^T \vec{n}_\perp)$.

Theorem 1 (Rotationally Invariant Quadrature Energy) *Let $\{E_n\}$ denote the set of polar energies which are obtained by applying a polar filter family with radial component K_ρ to an intrinsically 1D image f. Then all sums of an even and an odd energy*

$$E_{\text{quadrature}} = E_l + E_m, \quad l \in \{0, 2, \ldots\}, m \in \{1, 3, \ldots\}$$

are equal, independent of l and m. $E_{\text{quadrature}}$ is an intrinsically 1D image whose characteristic profile is identical to the energy response obtained by applying a 1D quadrature filter with frequency behavior K_ρ to the original image's characteristic profile g.

Proof: Since the polar energies are rotationally invariant, we can choose the orientation of the image profile arbitrarily. Let the image vary in the horizontal direction (i.e. it is constant along vertical sections). Then its Fourier transform is only non-zero on the horizontal u-axis. Thus, only the filter values on this axis, i.e. at angles 0 and π, have any influence on the filter result. It holds that $\sin(n0) = \sin(n\pi) = 0$, so the sine components $C_2^{(n)}$ vanish for all n. For the cosine components we have $\cos(n0) = 1$, $\cos(l\pi) = 1$ (as l is even) and $\cos(m\pi) = -1$ (as m is odd). Therefore, for any even l and odd m it holds along the horizontal u-axis that

$$K^{(l)}(u) = \pm K_\rho(|u|)$$

$$K^{(m)}(u) = \pm(-j)\,\text{sign}(u)K_\rho(|u|)$$

$$\Rightarrow \quad K^{(l)}(u) = \pm(-j)\,\text{sign}(u)K^{(m)}(u)$$

Thus, except for the sign all even filters yield identical results, likewise all odd filters. Hence, the quadrature energy is independent of l and m. Furthermore, the last expression is identical to the defining equation (1) of a quadrature filter pair in 1D (again except for the sign), so that a 1D qudrature filter would yield the same energy profile. ∎

Thus, we can say that pairs of even/odd channels generalize the 1D notion of a quadrature filter pair to 2D. Instead of two filters we now need four (unless $l = 0$, where three are sufficient). In contrast to the standard approach to 2D quadrature filters which requires orientation tuning, our quadrature filters provide a *rotationally invariant* energy response. By combining even/odd channels of different orders, we can construct infinitely many different such filters even when the radial function is kept constant. This allows us to select channels in such a way that we also obtain desirable responses for corners and junctions.

Since linear combinations of rotationally invariant quantities are still rotationally invariant, we can generalize the above result even further:

$$E_{\text{quadrature}} = \sum_{l \text{ even}} a_l E_l + \sum_{m \text{ odd}} b_m E_m, \quad \sum a_l = \sum b_m$$

Any linear combination of polar energies is a rotationally invariant quadrature filter as long as the coefficients of the odd and even energies balance. It is interesting to note that the *monogenic signal* which was introduced as the first rotationally quadrature filter by [5], can be interpreted as the simplest special case of our approach with $l = 0$ and $m = 1$.
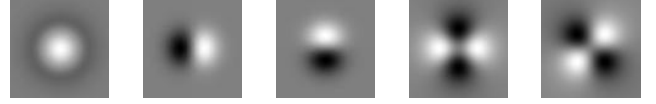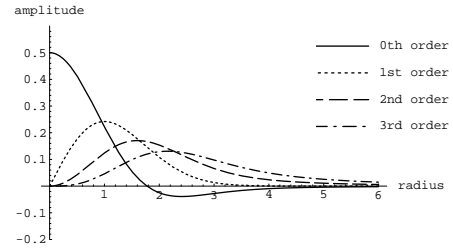
4.3. The Boundary Energy

We have shown in section 4.1 that the polar 2-jet responds to step and roof edges as well as local extrema (junctions of roof edges) and saddle points (checker-board junctions of step edges). Therefore, it is natural to base our integrated boundary filter on the polar 2-jet. Although we do not explicitly arrange for the detection of all junction types (especially not T-junctions) it turns out that the boundary filter is also able to detect them. On the other hand, dropping the second order would be of no use because it would deprive us of the possibility to detect saddle junctions and to determine the orientation of roof edges.

The *boundary energy* is now defined as the sum of the channel energies up to order 2, where even channels receive the weight $\frac{1}{2}$:

$$E_{\text{Boundary}} = \frac{1}{2} E_0 + E_1 + \frac{1}{2} E_2 \tag{15}$$

Since the boundary energy is a linear combination of rotationally invariant energies, it is itself rotationally invariant. It also has the quadrature property for intrinsically 1-dimensional features because the coefficients of odd and

Figure 1: Top: Hankel transforms of order 0 to 3 for the band-pass $K_\rho(\rho, \sigma) = \rho \exp(-\rho^2/2)$. The filter maxima move towards higher radius as the order increases, and the tails are longer than those of Gaussian derivatives. Bottom: resulting polar harmonic filters up to order 2.

Figure 2: Left to right: original image (circular step and roof edges); gradient magnitude; squared Laplacian; oriented quadrature filter; boundary energy. Only the boundary energy yields a rotational invariant unimodal response for both step and roof edges. Filtering is done in the Fourier domain throughout the paper.

even order terms balance. It is an integrated detector for step edges, roof edges, and junctions thereof. A model for these features is implicitly defined by the choice of the radial filter function. Experimentally, we have found that good results are obtained if the radial function is chosen so that the first order filters are the first derivatives of a Gaussian. That is, we set $K_\rho(\rho, \sigma) = \rho \exp(-\rho^2 \sigma^2/2)$ in the Fourier domain, where σ is the filter scale. Figure 1 depicts this function's Hankel transforms (i.e. the radial filter components in the spatial domain) and the resulting spatial domain filters. It will be interesting to investigate whether the feature model implied by this or other choices can be made explicit, or if an optimal radial function can be derived by starting from a particular model. Figure 2 compares several common boundary detectors and shows that only the boundary energy yields a rotationally invariant unimodal response.

4.4. The Boundary Tensor

The edge energy defined above is a rotationally invariant measure for the boundary strength at each image location. However, this is not sufficient for a detailed characterization of the local boundary type. Therefore, we combine the

components of the polar 2-jet into the *boundary tensor*. This second order tensor should have the following properties: (i) The trace of the tensor equals the boundary energy. (ii) The eigenvalues of the tensor encode the oriented energy, measured along the two major tensor axes. Since the eigenvalues encode energy they must not be negative. (iii) The eigenvector corresponding to the large eigenvalue indicates the local orientation, which is the direction of maximal signal variation, i.e. the direction perpendicular to an edge. (iv) The local strength of the intrinsically 1D information is given by the difference of the eigenvalues, while the intrinsically 2D energy equals twice the small eigenvalue.

A suitable tensor for the first order (odd) information can be defined by taking the tensor product of the generalized gradient with itself:

$$T^{(\text{odd})} = \begin{pmatrix} (C_1^{(1)})^2 & C_1^{(1)}C_2^{(1)} \\ C_1^{(1)}C_2^{(1)} & (C_2^{(1)})^2 \end{pmatrix} \qquad (16)$$

This tensor is analogous to the structure tensor (2), but without the spatial averaging. It has the required properties: its trace is the first order energy, it has a single non-zero eigenvalue whose value is identical to the energy (the other eigenvalue is 0), and the tensor orientation equals the direction of the generalized gradient modulo π.

For the even components of the boundary filter we already defined the generalized Hessian (12). Although this tensor contains the right kind of information it does not itself fulfill all the above requirements, as its trace is not an energy but an amplitude. However, this problem can be solved by taking the (matrix) square of $T^{(2)}$:

$$T^{(\text{even})} = T^{(2)}T^{(2)} = \qquad (17)$$

$$\begin{pmatrix} (C^{(0)} + C_1^{(2)})^2 + (C_2^{(2)})^2 & 2C^{(0)}C_2^{(2)} \\ 2C^{(0)}C_2^{(2)} & (C^{(0)} - C_1^{(2)})^2 + (C_2^{(2)})^2 \end{pmatrix}$$

This tensor indeed has the desired properties: its trace equals the even energy $E_0 + E_2$, its orientation $\frac{1}{2}\arctan(C_2^{(2)}/C_1^{(2)})$ is identical to the orientation of the second order filter, and the non-negative eigenvalues are

$$\lambda_{1,2}^{(\text{even})} = \left(C^{(0)} \pm \sqrt{(C_1^{(2)})^2 + (C_2^{(2)})^2} \right)^2 \qquad (18)$$

The eigenvalues also meet the requirements stated above: An intrinsically 1D structure (roof edge) is signaled with the same strength by both the zeroth and second order filters, so that one eigenvalue becomes zero, and the other is equal to the 1D (edge) energy. In case of a purely 2D structure, either the zeroth or second order energies are zero (since we have a saddle or a local extremum respectively), and both eigenvalues become identical.

The complete *boundary tensor* is now defined by the sum of the even and odd tensors, where the even tensor must get weight $\frac{1}{4}$ so that the tensor's trace equals the boundary energy:

$$T^{(\text{Boundary})} = T^{(\text{odd})} + \frac{1}{4}T^{(\text{even})} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$b_{11} = (C_1^{(1)})^2 + \frac{1}{4}\left((C^{(0)} + C_1^{(2)})^2 + (C_2^{(2)})^2 \right)$$

$$b_{12} = b_{21} = C_1^{(1)}C_2^{(1)} + \frac{1}{2}C^{(0)}C_2^{(2)} \qquad (19)$$

$$b_{22} = (C_2^{(1)})^2 + \frac{1}{4}\left((C^{(0)} - C_1^{(2)})^2 + (C_2^{(2)})^2 \right)$$

Boundary energy and local orientation can be calculated form the boundary tensor as follows:

$$E_{\text{Boundary}} = \text{tr}(T^{(\text{Boundary})}) = b_{11} + b_{22} \qquad (20)$$

$$\psi = \frac{1}{2}\arctan\frac{2C_1^{(1)}C_2^{(1)} + C^{(0)}C_2^{(2)}}{(C_1^{(1)})^2 - (C_2^{(1)})^2 + C^{(0)}C_1^{(2)}} \qquad (21)$$
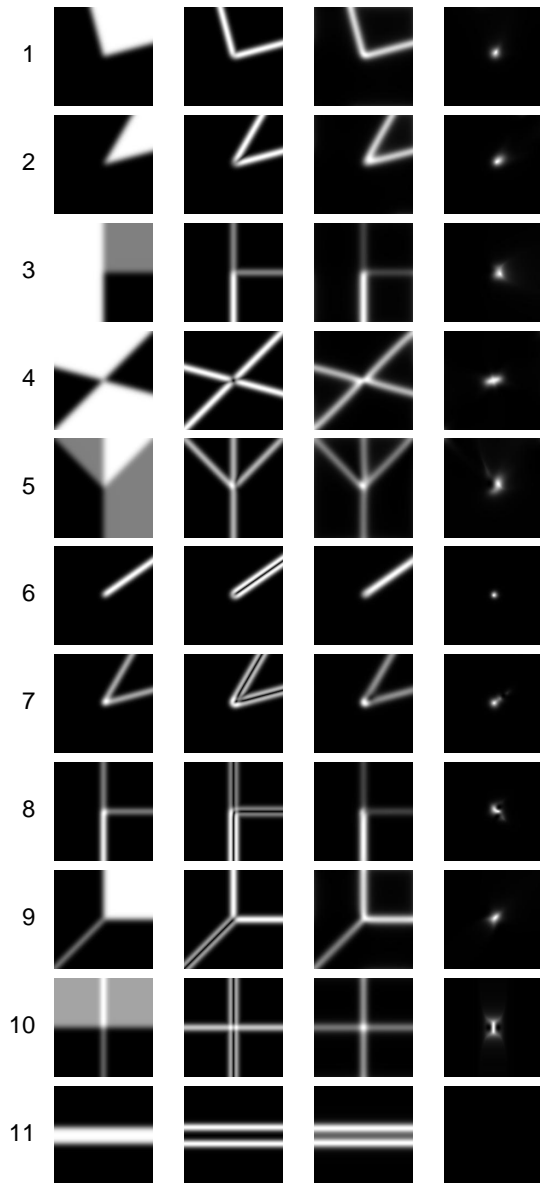
Equation (21) is a very interesting result: If the even energy is zero, ψ equals the gradient orientation, and if the odd energy is zero, ψ is identical to the orientation of the generalized Hessian. When both energies are non-zero, we get an average between the two orientations.

When we decompose the tensor into edge and junction parts according to (7), we get the following expressions for the edge and junction energies:

$$E_{\text{Edge}} = \sqrt{ \left((C_1^{(1)})^2 - (C_2^{(1)})^2 + C^{(0)}C_1^{(2)} \right)^2 + \left(2C_1^{(1)}C_2^{(1)} + C^{(0)}C_2^{(2)} \right)^2 } \qquad (22)$$
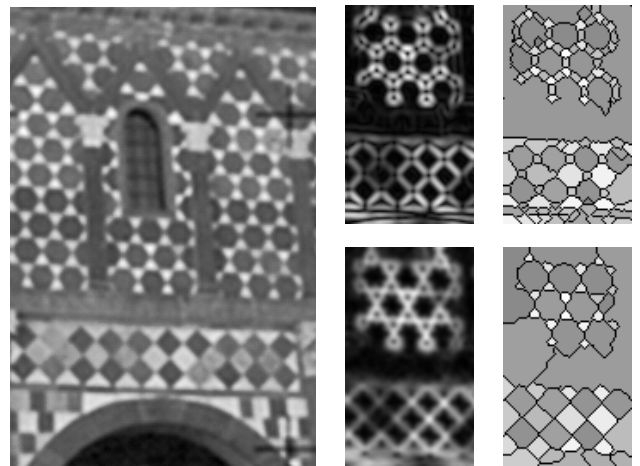
$$E_{\text{Junction}} = E_{\text{Boundary}} - E_{\text{Edge}} \qquad (23)$$

Figure 3 illustrates the boundary and junction energies for a number of test configurations and compares them with the gradient magnitude. It can be seen that the gradient is inferior to the boundary energy at junctions: in case of line junctions, the bimodal gradient responses give raise to multiple junctions, whereas for the saddle-point junction (row 4) it gives no answer at all. In contrast, single edges always yield unimodal responses of the boundary energy. Moreover, the boundary tensor properly handles line terminations and junctions with various edge angles, although a model for these feature types was not explicitly build into the tenseor. In case of a bar (Fig. 3 row 11), the boundary energy gives a bimodal edge response if the curvature at the bar's centerline is significantly below the curvature at its borders. It can also be observed that the junction energy has always a local maximum at roughly the right location, i.e. with a small offset from the true junction (we are currently working on a quantitative analysis of this statement).

Figure 3: Analysis of some test patterns. From left to right: original image, gradient magnitude, boundary energy, junction energy (filter scale: approx. $\frac{1}{4}$ line width).

Figure 4: Left: original image (tiled wall of historic building, enlarged); Right, top row: gradient magnitude and resulting segmentation (detail below the window); Right, bottom row: boundary energy and resulting segmentation.

The boundary tensor allows us to classify the local configuration into different structure types. If the edge energy exceeds the junction energy we consider the local structure as edge-like. By comparing the contributions of the even and odd energies, we can further distinguish roof and step edges. If the junction energy is larger than the edge energy, we have a corner or junction (however, accurate calculation of a junction's degree requires methods beyond the boundary tensor approach). It should be noted that a junction is also signaled when both the even and odd tensors alone indicate edges, but with perpendicular orientation (row 10 in Fig. 3). In principle, the classification can be done for any pixel, but it is, of course, most useful at boundary points, i.e. after some kind of non-maxima suppression. Away from the boundary, the boundary energy may be too small for the classification to be meaningful.
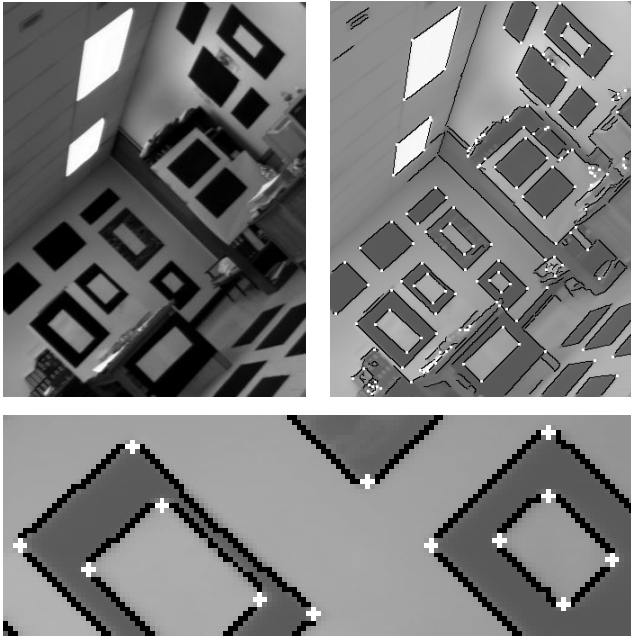
5. Examples

In this section we demonstrate that the boundary tensor indeed produces good boundary descriptions on real images. In the examples, we used filters based on the radial function $\rho \exp(-\rho^2\sigma^2/2)$ with $\sigma = 0.6$. All filters were applied in the Fourier domain. Standard non-maxima suppression and hysteresis thresholding were applied to the calculated tensor energies.

Figure 4 illustrates the superior junction detection of the boundary tensor. It can be seen that the gradient-based segmentation contains spurious regions at saddle junctions that result from the gradient being zero at saddle points. The boundary tensor-based segmentation does not contain such artifacts. Figure 5 shows that the boundary tensor is indeed an integrated detector for both edges and junctions. The edges and junctions shown were derived from the same tensor measurements (using the edge and junction energies respectively) and thus exhibit much less displacement than standard edge and corner detectors.

6. Discussion

In this paper we proposed the boundary tensor as a new tool for boundary analysis. The big advantage of the tensor approach is its ability to integrate edge and junction detection.

Figure 5: Top left: original image; top right: segmentation by means of the boundary tensor (black: edges, white: corners/junctions); bottom: detail of segmentation.

By combining this approach with polar separable filters, we were able to design a tensor that responds reasonably to a wide range of different edge and junction types (including T- and X-junctions). This also allowed us to prove that the boundary energy, i.e. the trace of the boundary tensor, acts as a rotationally invariant quadrature filter. Thus the tensor is suited for both step and roof edge detection. The tensor approach also facilitates integration of information from several sources. For example, we could obtain a color boundary detector by first calculating the boundary tensors for each color band separately, and then adding the tensors at each pixel (we cannot demonstrate this due to space limitations).

Theorem 1 on rotationally invariant quadrature filters sheds new light on the discussion of how the quadrature property should be generalized to 2D. Currently, oriented quadrature filters are used most often. Felsberg and Sommer [5] proposed the first rotationally invariant generalization, the monogenic signal. Our analysis reveals that this is just the simplest case among infinitely many possible generalizations which can be defined by weighting filter orders differently. Apparently, the requirement of uniform response to even and odd 1D structures is too weak to determine how many filters are needed in 2D and how they should be related. The extra degrees of freedom can be used to define additional constraints for the detection of *intrinsically 2D structures*. In this paper, the weights were derived from our decision to use filters up to order 2. This choice

is not yet optimal. First, the junction response is not always unimodal. This means that the junction model implied by our choice of filters does not yet cover all cases. Second, at obtusely angled step edge corners above approximately 120 degrees, the junction energy does not exceed the edge energy, so the corner is missed by our structure type classification.

Junction detection could be improved by including higher order filters into the boundary tensor. But in practice the design of good quality filters for orders ≥ 3 is non-trivial because the symmetry of the filters does no longer match the symmetry of the sampling grid. Thus, it is difficult to obtain filters that exhibit precisely the desired frequency behavior and at the same time exactly fulfill the steering equations.

References

[1] F. Attneave: *Some Informational Aspects of Visual Perception*, Psychl. Rev. 61(3), pp. 183-193, 1954

[2] D. Beymer: *Finding Junctions Using the Image Gradient*, in: Proc. IEEE Computer Vision and Pattern Recognition, Lahaina, Maui, Hawaii, pp. 720-721, 1991

[3] J. Canny: *A Computational Approach to Edge Detection*, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6), pp. 679-698, 1986

[4] R. Deriche, G. Giraudon: *A Computational Approach for Corner and Vertex Detection*, Int. J. of Computer Vision, 10(2), pp. 101–124, 1993

[5] M. Felsberg, G. Sommer: *The Monogenic Signal*, IEEE Trans. on Signal Processing, 49(12), pp. 3136-3144, 2001

[6] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever: *The Gaussian Scale-Space Paradigm and the Multiscale Local Jet*, Int. J. of Computer Vision, 18(1), pp. 61-75, 1996

[7] W. T. Freeman and E. H. Adelson, *The Design and Use of Steerable Filters*, IEEE Trans. on Pattern Analysis and Machine Intelligence, 13(9), pp. 891-906, 1991

[8] G. Granlund, H. Knutsson: *Signal Processing for Computer Vision*, Kluwer, 1995

[9] C.G. Harris, M.J. Stevens: *A Combined Corner and Edge Detector*, Proc. of 4th Alvey Vision Conference, 1988

[10] G. Kanisza: *Organization in Vision*, Holt, Rinehart and Winston, 1979

[11] D.C. Kay: *Tensor Calculus*, McGraw-Hill, 1988

[12] V. Kovalevsky: *Finite Topology as Applied to Image Analysis*, Computer Vision, Graphics, and Image Processing, 46(2), pp. 141-161, 1989

[13] G. Medioni, M.-S. Lee, C.-K. Tang: *A Computational Framework for Segmentation and Grouping*, Elsevier, 2000

[14] K. Rohr: *Localization Properties of Direct Corner Detectors*, J. of Mathematical Imaging and Vision, 4, pp. 139-150, 1994

[15] C. Rothwell, J. Mundy, W. Hoffman and V.-D. Nguyen: *Driving Vision by Topology*, in: Proc. IEEE Symposium on Computer Vision, pp. 395-400, 1995

[16] J. Weickert: *Coherence-Enhancing Diffusion Filtering*, Intl. J. of Computer Vision, 31(2/3), pp. 111-127; 1999