Estimation of Multiple Orientations at Corners and Junctions

Cicero Mota[1], Ingo Stuke[2], Til Aach[2], and Erhardt Barth[1]

[1] Institute for Neuro- and Bioinformatics, University of Luebeck, Germany
{mota, barth}@inb.uni-luebeck.de
[2] Institute for Signal Processing, University of Luebeck, Germany
{aach, stuke}@isip.uni-luebeck.de

Abstract. Features like junctions and corners are a rich source of information for image understanding. We present a novel theoretical framework for the analysis of such 2D features in scalar and multispectral images. We model the features as occluding superpositions of two different orientations and derive a new constraint equation based on the tensor product of two directional derivatives. The eigensystem analysis of a 3×3-tensor then provides the so-called mixed-orientation parameters (MOP) vector that encodes the two orientations uniquely, but only implicitly. We then show how to separate the MOP vector into the two orientations by finding the roots of a second-order polynomial. Based on the orientations, the occluding boundary and the center of the junction are easily determined. The results confirm the validity, robustness, and accuracy of the approach.

1 Introduction

It is well known that corners and junctions are a rich source of information for image understanding: T-junctions are associated to object occlusions; L- and Y-junctions to object corners; X-junctions to the occurrence of transparencies; and Ψ-junctions to the presence of bending surfaces in the object [2, 12]. Accordingly different approaches for junction localization have been reported [9, 11, 13, 14, 17, 19, 22].

In addition to the above semantic importance of junctions and corners, their significance is determined by basic properties of the image function itself. Flat regions in images are the most frequent but also redundant, their intrinsic dimension [23] is zero. Intrinsically one-dimensional features like straight edges are still redundant since intrinsically two-dimensional regions have been shown to fully specify an image [8, 15]. Corners and junctions belong to the last category of most significant image features.

In this paper we model an image junction as a superposition of oriented structures and show how to estimate the multiple orientations occurring at such positions. Our approach differs from previous attempts [10, 18, 19] in that we provide a closed-form solution.

Our results are an extension of earlier results that have dealt with the problems of estimating transparent motions [16, 21], occluded motions [5, 4], and multiple orientations in images based on an additive model [20, 1].

2 Theoretical Results

Let $\mathbf{f}(\mathbf{x})$ be an image that is ideally oriented in a region Ω, i.e., there is a direction (subspace) E of the plane such that $\mathbf{f}(\mathbf{x}+\mathbf{v}) = \mathbf{f}(\mathbf{x}) \quad$ for all $\mathbf{x}, \mathbf{v}$ such that $\mathbf{x}, \mathbf{x}+\mathbf{v} \in \Omega$, $\mathbf{v} \in E$. This is equivalent to

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{v}} = \mathbf{0} \quad \text{for all } \mathbf{x} \in \Omega \text{ and } \mathbf{v} \in E \tag{1}$$

which is a system of q equations for $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^q$. For intensity images $q = 1$ and for RGB images $q = 3$. The direction E can be estimated as the set of vectors that minimize the energy functional

$$\mathcal{E}(\mathbf{v}) = \int_\Omega \left| \frac{\partial \mathbf{f}}{\partial \mathbf{v}} \right|^2 \mathrm{d}\Omega = \mathbf{v}^T \mathbf{J} \mathbf{v} \ , \tag{2}$$

where $\mathbf{J}$ is given by

$$\mathbf{J} = \int_\Omega \begin{bmatrix} |\mathbf{f}_x|^2 & \mathbf{f}_x \cdot \mathbf{f}_y \\ \mathbf{f}_x \cdot \mathbf{f}_y & |\mathbf{f}_y|^2 \end{bmatrix} \mathrm{d}\Omega \ . \tag{3}$$

In the above equation, $\mathbf{f}_x$, $\mathbf{f}_y$ are short notations for $\partial \mathbf{f}/\partial x$, $\partial \mathbf{f}/\partial y$.

The tensor $\mathbf{J}$ is the natural generalization of the structure tensor [6, 7, 9, 11] for multi-spectral images. Since $\mathbf{J}$ is symmetric and non-negative, Eq. (1) is equivalent to $\mathbf{J}\mathbf{v} = \lambda \mathbf{v}, \mathbf{v} \neq \mathbf{0}$, where ideally $\lambda = 0$. This implies that E is the null-eigenspace of $\mathbf{J}$ and in practice estimated as the eigenspace associated to the smallest eigenvalues of $\mathbf{J}$. Confidence for the estimation can thus be derived from the eigenvalues (or, equivalently, scalar invariants) of $\mathbf{J}$ (see [16]): two small eigenvalues correspond to flat regions of the image; only one small eigenvalue to the presence of oriented structures; and two significant eigenvalues to the presence of junctions or other 2D structures. Below we show how to estimate the orientations at junctions where two oriented structures predominate.

2.1 Multiple Orientations

Let Ω be a region of high confidence for a junction. We model junctions by the following constraint on $\mathbf{f}(\mathbf{x})$ that is the occluded superposition

$$\mathbf{f}(\mathbf{x}) = \chi(\mathbf{x}) \mathbf{g}_1(\mathbf{x}) + (1 - \chi(\mathbf{x})) \mathbf{g}_2(\mathbf{x}) \ , \tag{4}$$

where $\mathbf{g}_1(\mathbf{x}), \mathbf{g}_2(\mathbf{x})$ are ideally oriented with directions $\mathbf{u} = (u_x, u_y)^T$ and $\mathbf{v} = (v_x, v_y)^T$ respectively; and $\chi(\mathbf{x})$ is the characteristic function of some half-plane P through the 'center' (to be defined later) of the junction. This model is appropriate for the local description of junction types T, L and Ψ. X-junctions better fit a transparent model and have been treated in [1, 20].

The Constraint Equation. To estimate two orientations in Ω, we observe that Eq. (4) is equivalent to

$$\mathbf{f}(\mathbf{x}) = \begin{cases} \mathbf{g}_1(\mathbf{x}) \text{ if } \mathbf{x} \in P \\ \mathbf{g}_2(\mathbf{x}) \text{ otherwise.} \end{cases} \tag{5}$$

Therefore, $\partial \mathbf{f}(\mathbf{x})/\partial \mathbf{u} = \mathbf{0}$ if $\mathbf{x}$ is inside of P and $\partial \mathbf{f}(\mathbf{x})/\partial \mathbf{v} = \mathbf{0}$ if $\mathbf{x}$ is outside of P. From the above we can draw the important and, as we shall see, very useful conclusion that the expression

$$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{u}} \otimes \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{v}} = \mathbf{0} \tag{6}$$

is valid everywhere except for the border of P where it may differ from zero. The symbol $\otimes$ denotes the tensor product of two vectors. Eq. (6) may not hold at the border of P because there the derivatives of the characteristic function $\chi(\mathbf{x})$ are not defined. This is not the case if $\mathbf{u}$ and the border of P have the same direction, e.g. in case of a T-junction. Given Eq. (6), the tensor product should be symmetric in $\mathbf{u}$ and $\mathbf{v}$. Since in practice symmetry might be violated, we expand the symmetric part of the above tensor product to obtain

$$c_{xx}\mathbf{f}_x \otimes \mathbf{f}_x + \frac{c_{xy}}{2}(\mathbf{f}_x \otimes \mathbf{f}_y + \mathbf{f}_y \otimes \mathbf{f}_x) + c_{yy}\mathbf{f}_y \otimes \mathbf{f}_y = \mathbf{0} \tag{7}$$

where

$$c_{xx} = u_x v_x, \quad c_{yy} = u_y v_y, \quad c_{xy} = u_x v_y + u_y v_x \ . \tag{8}$$

Note that for an image with q spectral components, the system in Eq. (7) has $q(q+1)/2$ equations, which makes the system over-constrained if $q > 2$. The vector $\mathbf{c} = (c_{xx}, c_{xy}, c_{yy})^T$ is the so-called *mixed orientation parameters vector* and is an implicit representation of the two orientations.

Estimation of the Mixed Orientation Parameters. An estimator of the mixed orientation parameters is obtained by a least-squares procedure that finds the minimal points of the energy functional

$$E(\mathbf{c}) = \int_\Omega \left\| c_{xx}\mathbf{f}_x \otimes \mathbf{f}_x + \frac{c_{xy}}{2}(\mathbf{f}_x \otimes \mathbf{f}_y + \mathbf{f}_y \otimes \mathbf{f}_x) + c_{yy}\mathbf{f}_y \otimes \mathbf{f}_y \right\|^2 \mathrm{d}\Omega, \tag{9}$$

i.e., $\mathbf{c}$ is estimated as an eigenvector $\hat{\mathbf{c}}$ associated to the smallest eigenvalue of

$$\mathbf{J}_2 = \int_\Omega \begin{bmatrix} |\mathbf{f}_x|^4 & |\mathbf{f}_x|^2 \mathbf{f}_x \cdot \mathbf{f}_y & |\mathbf{f}_x \cdot \mathbf{f}_y|^2 \\ |\mathbf{f}_x|^2 \mathbf{f}_x \cdot \mathbf{f}_y & \frac{1}{2}\left(|\mathbf{f}_x|^2 |\mathbf{f}_y|^2 + |\mathbf{f}_x \cdot \mathbf{f}_y|^2\right) & |\mathbf{f}_y|^2 \mathbf{f}_x \cdot \mathbf{f}_y \\ |\mathbf{f}_x \cdot \mathbf{f}_y|^2 & |\mathbf{f}_y|^2 \mathbf{f}_x \cdot \mathbf{f}_y & |\mathbf{f}_y|^4 \end{bmatrix} \mathrm{d}\Omega \ . \tag{10}$$

The actual region of integration can be kept smaller for multi-spectral images since the system in Eq. (7) is over-constrained in this case. Note, however, that $\hat{\mathbf{c}}$ represents two orientations only if it is consistent with Eq. (8). This is the case if and only if $c_{xy}^2 - 4c_{xx}c_{yy} > 0$.

Separation of the Orientations. To separate the orientations it suffices to know the matrix

$$\boldsymbol{C} = \begin{bmatrix} u_x v_x & u_x v_y \\ u_y v_x & u_y v_y \end{bmatrix} \tag{11}$$

because its rows represent one orientation and its columns the other, cf. [1] for the case of transparency. Since we already know that $c_{xx} = u_x v_x$ and $c_{yy} = u_y v_y$, we need to only obtain $z_1 = u_x v_y$ and $z_2 = u_y v_x$. For this, observe that $z_1 + z_2 = c_{xy}$ and $z_1 z_2 = c_{xx} c_{yy}$. Therefore, z_1, z_2 are the roots of

$$Q_2(z) = z^2 - c_{xy} z + c_{xx} c_{yy} \; . \tag{12}$$

Pruning of the Orientation Fields. After the separation, each point of a junction neighborhood has two directions assigned to it, see Fig. 1 (b,e) and Fig. 2 (b,e). Since only one of these is correct, we need to prune the other. For this, we observe that at each position only one of the equations

$$\partial \mathbf{f}(\mathbf{x})/\partial \mathbf{u} = \mathbf{0}, \quad \partial \mathbf{f}(\mathbf{x})/\partial \mathbf{v} = \mathbf{0} \tag{13}$$

is valid. To prune the wrong vector at a given position $\mathbf{p}$, we first compute the local histogram of the orientations in a small neighborhood (3×3 pixels) of $\mathbf{p}$ and separate the two orientations by the median. We then assign to $\mathbf{p}$ the correct direction depending on which equation in (13) is better satisfied in the sense that the sum of squares is lowest. This is equivalent to a procedure that would choose the direction of smallest variation of the image $\mathbf{f}(\mathbf{x})$.

Junction Localization. Since measures of confidence only give us a region where multiple orientations can occur, it is useful to have a method for deciding which point in this region is actually the center of the junction. We follow the approach in [9] for the localization of the junction. Let Ω represent a region of high confidence for a junction. For an ideal junction located at $\mathbf{p}$ we have

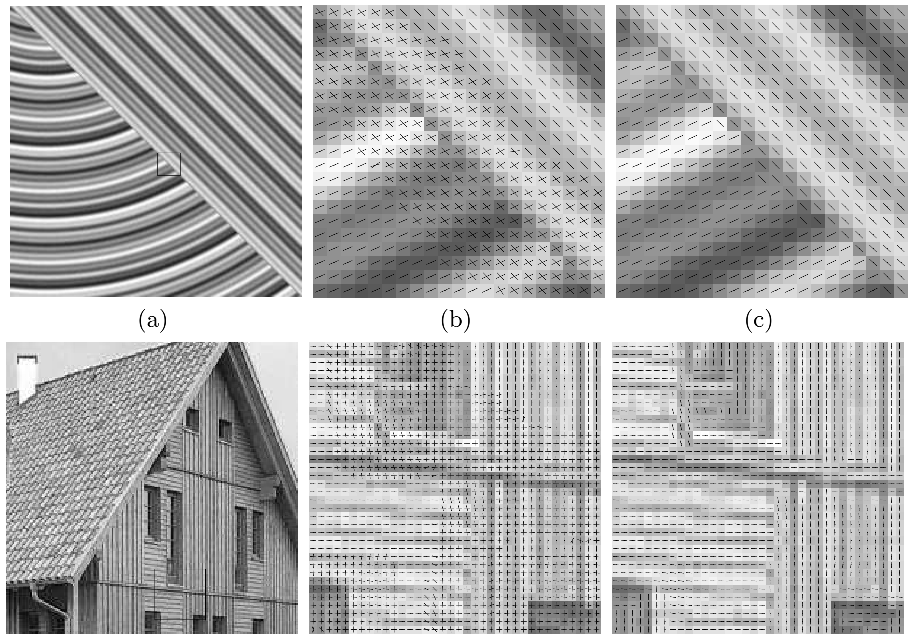
$$\mathrm{d}\mathbf{f}_{\mathbf{x}}(\mathbf{x} - \mathbf{p}) = \mathbf{0} \tag{14}$$

where $\mathrm{d}\mathbf{f}_{\mathbf{x}}$ is the Jacobian matrix of $\mathbf{f}(\mathbf{x})$. The center of the junction is therefore defined and estimated as the minimal point of

$$\int_{\Omega} |\, \mathrm{d}\mathbf{f}_{\mathbf{x}}(\mathbf{x} - \mathbf{p})|^2 \, \mathrm{d}\Omega \tag{15}$$

which gives

$$\hat{\mathbf{p}} = \mathbf{J}^{-1}\mathbf{b}, \text{ where } \mathbf{b} = \int_{\Omega} \mathrm{d}\mathbf{f}_{\mathbf{x}}^T \, \mathrm{d}\mathbf{f}_{\mathbf{x}} \mathbf{x} \, \mathrm{d}\Omega \; . \tag{16}$$



(a) (b) (c)

(d) (e) (f)

Fig. 1. Synthetic example: panel (a) depicts a sinusoidal pattern (b) the estimated orientations for the marked region and (c) the orientations after pruning. Real example: panel (d) shows a picture of a house; (e) and (f) are analogous to (b) and (c) above.

3 Results

Fig. 1 depicts the results of the estimation of single and double orientations in a synthetic and a natural image. In panel (a) two oriented sinusoidal patterns were combined to form T-junctions along the main diagonal of the image to which Gaussian white noise was added (SNR of 25 dB). The estimated orientations for the selected region in (a) are shown in panel (b). Note that in a region around the T-junction two orientations are estimated at each pixel. Panel (c) shows the result obtained after the pruning process. Note that the occluding boundary and the orientations on both sides of the boundary are well estimated. Panel (d) depicts an natural image with many oriented regions and junctions. The estimated orientations for the selected region in (d) are shown in panel (e). The orientations after the pruning process are depicted in panel (f).

Fig. 2 presents results for L-junctions of different angles. Panel (a) depicts the letter ‘A’ (image with additive Gaussian noise, SNR of 25 dB). In panel (d) a segmentation of the ‘A’ in terms of the number of estimated orientations is shown: white for no orientation, black for one, and gray for two orientations. Note that, around all corners of the letter, two orientations are found. The estimated orientations for the upper-left corner of the ‘A’ are shown in panels (b) (before pruning) and (c) (after pruning). Panel (e) depicts the estimated orientations

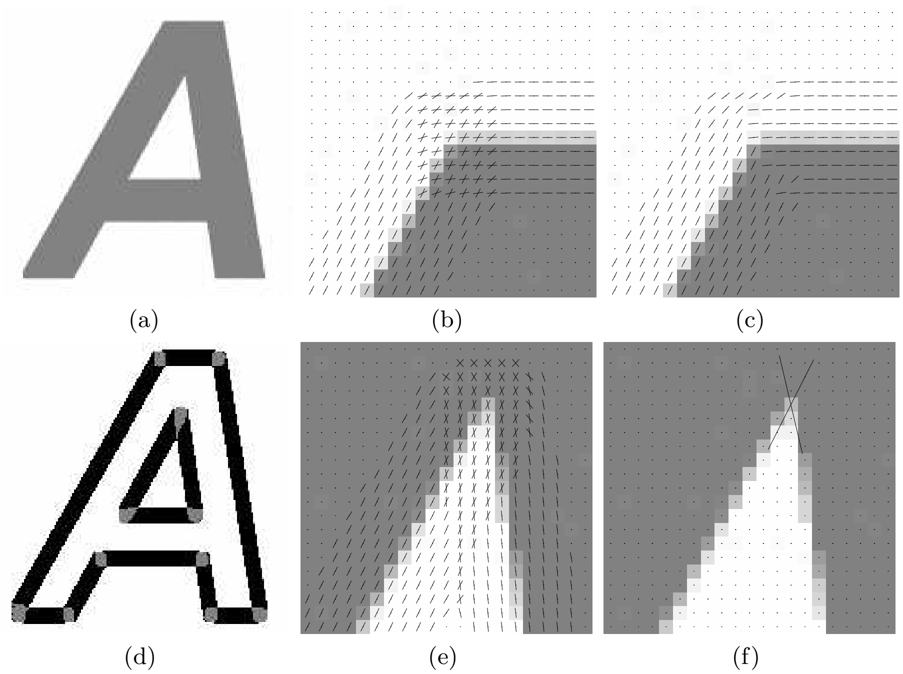


Fig. 2. Panel (a) shows the ‘A’-letter input image, estimated orientations before and after pruning are shown in (b) and (c) respectively for the upper left corner of the ‘A’. Panel (d) depicts the segmentation in terms of the number of observed orientations (see text), (e) the estimated orientations for the corner with the smallest angle, and (f) indicates the corner location estimated according to Eq. (16) and the two orientations at that location.

for the corner of the ‘A’ with the smallest angle. Pixels with two orientations are then used according Eq. (16) to locate the corner position. The result is indicated by the cross in panel (f) which also shows the corresponding orientations at the corner location.

In all examples, we first search for at least one single orientation. If confidence for at least one orientation is low, we search for double orientations. Confidence is based on the invariants of $\mathbf{J}_N$ according to [16]. Thus, for $\mathbf{J}_1$, the confidence criteria are $H > \epsilon$ and $\sqrt{K} < c_1 H$. For $\mathbf{J}_2$, the confidence criterion for two orientations is $\sqrt[3]{K} < c_2 \sqrt[2]{S}$. The numbers H, K and S are the invariants, i.e., the trace, the determinant, and the sum of the diagonal minors of $\mathbf{J}_{1,2}$. For the examples in Fig. 1, we used an integration window size of 11×11 pixels, $c_1 = 0.5$, $c_2 = 0.6.$, and $\epsilon = 0.001$. For the example in Fig. 2, we used a integration-window size of 7×7 pixels, $c_1 = 0.4$, $c_2 = 0.6$, and $\epsilon = 0.01$. The above parameter settings have been found experimentally. Derivatives were taken with a $[-1, 0, 1]^T [1, 1, 1]$ kernel in $x-$ and analogously in y direction.

4 Conclusions

We have presented a straightforward and accurate method for the estimation of two orientations at image features that satisfy an occlusion model. Typical features are corners and various kinds of junctions that occur frequently in natural images. The method is straightforward because it only involves first-order derivatives and has closed-form solutions. Iterative procedures are not involved, unless one chooses to estimate the eigenvectors of a 3×3 tensor iteratively. This can be avoided as shown in [3]. The method is accurate because junctions and corners are well described by the occlusion model that we use and because we have found the proper constraint that results from the occlusion model and an analytical solution. Our results confirm this and can be further improved.

An obvious improvement is the use of optimized derivative kernels. Derivatives could also be replaced by more general filters as a consequence of results obtained in [16]. Based on a straightforward extension of our approach, one could estimate more than two orientations.

We have formulated our results such as to include multi-spectral images in a natural but non-trivial way. If q is the number of colors, the constraint that we use consists of $q(q+1)/2$ equations. This implies that for only two colors we already have a well conditioned system and can use even smaller neighborhoods for the estimation of two orientations. Forthcoming results will show the benefit of using multi-spectral images.

The benefits of using corners and junctions for image analysis, registration, tracking etc. have often been highlighted. The estimation of the orientations that form these features may add further robustness and new kinds of invariant features. It might, for example, be easier to register a junction in terms of its orientations since the orientations will change less than the appearance and other features of the junction. The orientations seem especially useful as they can now be well estimated with little computational effort.

Acknowledgment. Work is supported by the *Deutsche Forschungsgemeinschaft* under Ba 1176/7-2.

References

1. T. Aach, I. Stuke, C. Mota, and E. Barth. Estimation of multiple local orientations in image signals. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Montreal, Canada, May 17–21, 2004. To appear.
2. E. H. Adelson. Lightness perception and lightness illusions. In M. Gazzaniga, editor, *The New Cognitive Neurosciences*, pages 339–51. MIT Press, 2000.
3. E. Barth. The minors of the structure tensor. In G. Sommer, editor, *Mustererkennung 2000*, pages 221–228. Springer, Berlin, 2000.
4. E. Barth, I. Stuke, T. Aach, and C. Mota. Spatio-temporal motion estimation for transparency and occlusion. In *Proc. IEEE Int. Conf. Image Processing*, volume III, pages 69–72, Barcelona, Spain, Sept. 14–17, 2003. IEEE Signal Processing Soc.

5. E. Barth, I. Stuke, and C. Mota. Analysis of motion and curvature in image sequences. In *Proc. IEEE Southwest Symp. Image Analysis and Interpretation*, pages 206–10, Santa Fe, NM, Apr. 7–9, 2002. IEEE Computer Press.
6. J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with application to texture analysis and optical flow. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(8):775–90, 1991.
7. S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33:116–25, Jan. 1986.
8. C. Z. Erhardt Barth, Terry Caelli. Image encoding, labeling, and reconstruction from differential geometry. *CVGIP: Graphical Model and Image Processing*, 55(6):428–46, Nov. 1993.
9. W. Förstner and E. Gülch. A fast operator for detection and precise locations of distinct points, corners, and centres of circular features. In *Intercommission Conference on Fast Processing of Photogrammetric Data*, pages 281–305, 1987.
10. W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.
11. C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. The Fourth Alvey Vision Conference*, pages 147–152, 1988.
12. D. A. Huffman. Impossible objects as non-sense sentences. *Machine Intelligence*, VI:295–323, 1971.
13. L. Kitchen and A. Rosenfeld. Gray level corner detection. *Pattern Recognition Letters*, pages 95–102, 1982.
14. U. Köthe. Edge and junction detection with an improved structure tensor. In B. Michaelis and G. Krell, editors, *Pattern Recognition, 25th DAGM Symposium*, volume 2781 of *LNCS*, pages 25–32, 2003.
15. C. Mota and E. Barth. On the uniqueness of curvature features. In G. Baratoff and H. Neumann, editors, *Dynamische Perzeption*, volume 9 of *Proceedings in Artificial Intelligence*, pages 175–8, Köln, 2000. Infix Verlag.
16. C. Mota, I. Stuke, and E. Barth. Analytic solutions for multiple motions. In *Proc. IEEE Int. Conf. Image Processing*, volume II, pages 917–20, Thessaloniki, Greece, Oct. 7–10, 2001. IEEE Signal Processing Soc.
17. H. H. Nagel and W. Enkelmann. Investigation of second order greyvalue variations to estimate corner point displacements. In *International Conference on Pattern Recognition*, pages 768–773, 1982.
18. L. Parida, D. Geiger, and R. Hummel. Kona: A multi-junction detector using minimum description length principle. In M. Pelillo and E. R. Hancock, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 1223 of *LNCS*, pages 51–65, 1997.
19. K. Rohr. Recognizing corners by fitting parametric models. *International Journal of Computer Vision*, 9:213–230, 1992.
20. I. Stuke, T. Aach, E. Barth, and C. Mota. Analysing superimposed oriented patterns. In *Proc. IEEE Southwest Symp. Image Analysis and Interpretation*, Lake Tahoe, NV, Mar. 28–30, 2004.
21. I. Stuke, T. Aach, C. Mota, and E. Barth. Estimation of multiple motions: regularization and performance evaluation. In B. Vasudev, T. R. Hsing, A. G. Tescher, and T. Ebrahimi, editors, *Image and Video Communications and Processing 2003*, volume 5022 of *Proceedings of SPIE*, pages 75–86, May 2003.
22. C. Tomasi and T. Kanade. Detection and tracking of point features. Technical report, Carnegie Mellon University Technical, april 1991. Report CMU-CS-91-132.
23. C. Zetzsche and E. Barth. Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, 30:1111–7, 1990.