Network: Computation in Neural Systems
December 2005; 16(4): 341–356

Taylor & Francis
Taylor & Francis Group

Local image structures and optic flow estimation

S. KALKAN,[1] D. CALOW,[2] F. WÖRGÖTTER,[1]
M. LAPPE,[2] & N. KRÜGER[3]

[1]*Psychology, University of Stirling, Scotland,* [2]*Psychology, University of Münster, Germany,*
[3]*Computer Sci., Aalborg University Esbjerg, Denmark*

(Received 22 November 2004; revised 3 October 2005; accepted 7 October 2005)

Abstract
Different kinds of local image structures (such as homogeneous, edge-like and junction-like patches) can be distinguished by the intrinsic dimensionality of the local signals. Intrinsic dimensionality makes use of variance from a point and a line in spectral representation of the signal in order to classify it as homogeneous, edge-like or junction-like. The concept of intrinsic dimensionality has been mostly exercised using discrete formulations; however, recent work (Felsberg & Krüger 2003; Krüger & Felsberg 2003) has introduced a continuous definition. The current study analyzes the distribution of local patches in natural images according to this continuous understanding of intrinsic dimensionality. This distribution reveals specific patterns than can be also associated to local image structures established in computer vision and which can be related to orientation and optic flow features. In particular, we link quantitative and qualitative properties of optic-flow error estimates to these patterns. In this way, we also introduce a new tool for better analysis of optic flow algorithms.

Keywords: *Optic flow, intrinsic dimension, image structures*

Introduction

Natural images are dominated by specific local image structures, such as homogeneous patches, edges, corners, or textures. Sub-domains of Computer Vision have extracted and analyzed such structures in edge detection (see, e.g., Canny 1986), junction classification (see, e.g., Rohr 1992) and texture interpretation (see, e.g., Sochen et al. 1998; Ribeiro & Hancock 2001). There exists evidence that also in human vision such structures are processed (see, e.g., Hubel & Wiesel 1969; Shevelev et al. 1995) and that they play distinguishable roles in the information processing (see, e.g., Koenderink & Dorn 1982).

The intrinsic dimension (see, e.g., Zetzsche & Barth 1990; Felsberg 2002) has proven to be a suitable descriptor that distinguishes such structures. Homogeneous image patches have an intrinsic dimension of zero (i0D); edge-like structures are intrinsically one-dimensional (i1D) while junctions and most textures have an intrinsic dimension of two (i2D).[1]

The association of intrinsic dimension to a local image structure has been done mostly by a discrete classification (Zetzsche & Barth 1990; Jähne 1997; Felsberg 2002). A *continuous* definition of intrinsic dimensionality has been recently given in Felsberg & Krüger (2003);

Correspondence: S. Kalkan, Psychology, University of Stirling, Scotland. Tel: xxx. Fax: xxx. E-mail: sinan@cn.stir.ac.uk

ISSN: 0954-898X print / ISSN 1361-6536 online
DOI: 10.1080/09548980500445005

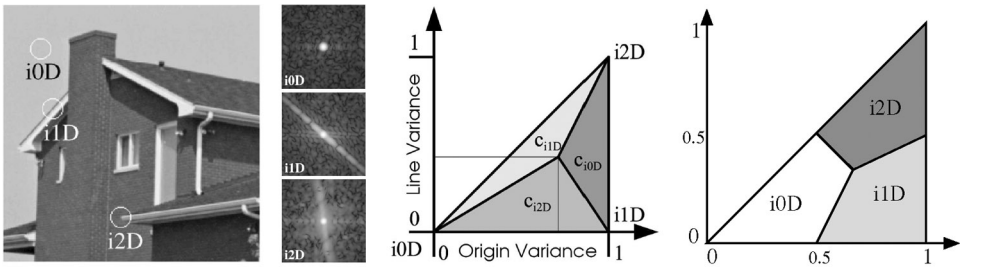

Figure 1. Illustration of intrinsic dimensionality ((a,b,c) taken from Felsberg & Krüger (2003). (a) Three image patches for three different intrinsic dimensions, (b) The 2D spatial frequency spectrum of the local patches in (a), from top to bottom: i0D, i1D, i2D. (c) The topology of intrinsic dimensionality. Origin variance is variance from a point, i.e., the origin. Line variance is variance from a line; this measures the junction-ness of the signal. c_{jND} for N = 0,1,2 stands for confidence for being i0D, i1D and i2D, respectively, (d) The decision areas for local image structures.

Krüger and Felsberg (2003). In these studies, it has also been shown that the topological structure of intrinsic dimension essentially has the form *of a triangle* which is spanned by two axes corresponding to amplitude and variance in orientation (see Figure 1). In this paper, we will use this continuous definition to investigate the structure of the distribution of signals in natural images according to their intrinsic dimensionality. More specifically, we will show that:

D0. i0D signals split into two clusters; one peak corresponding to over-illuminated (white) or under-illuminated (black) patches and a Gaussian-shaped cluster corresponding to image noise at homogeneous but not under- or over-illuminated image patches (see Figure 2a).

D1. For i1D signals, there exists a concentration of signals in a stripe-shaped cluster corresponding to high origin variance (high amplitude) *and* low line variance (see Figure 2a). This also reflects the importance of an orientation criterion that is based on local amplitude and orientation information (see, e.g., Princen et al. 1990).

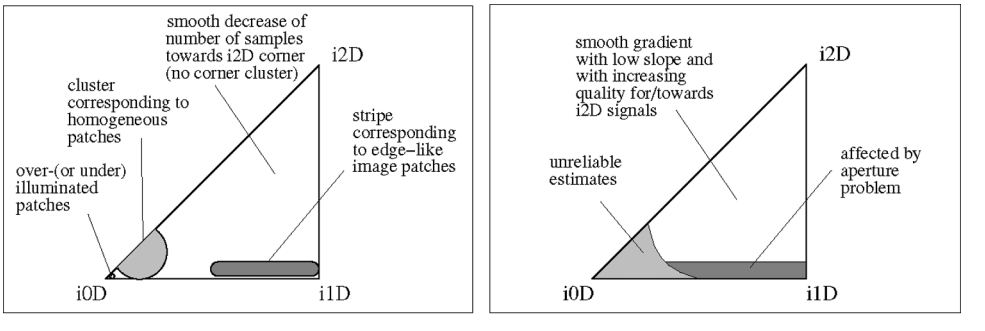

Figure 2. a) Schematic representation of the distribution of local image patches in natural images according to their intrinsic dimension, b) Schematic representation of the quality of optic flow estimation according to the intrinsic dimension of the underlying signal.

D2. In contrast to the i0D and i1D cases, there exists no cluster for i2D signals but there is a smoothly decreasing surface towards the i2D corner. This continuity in the distribution for the i2D case indicates that it is rather difficult to formulate a purely local criterion to detect corners in natural images.

Thus, the continuous formulation of intrinsic dimension allows for a more precise characterization of established local image structures in terms of their statistical manifestation in natural images. As a consequence, properties and inherent problems of classical computer vision algorithms can be reflected.

The property of optic flow estimation at homogeneous image patches, edges and corners has been discussed extensively (see, e.g., Zetzsche et al. 1991; Barron et al. 1994; Mota & Barth 2000). It has been argued that many different motion detectors specialised to particular image structures exist in human vision (for a discussion, see Cavanagh and Mather 1989; Johnston and Clifford 1995). In general, it is acknowledged that:

A0. Optic flow estimates at homogeneous image patches tend to be unreliable as the lack of structure makes it impossible to find correspondences in consecutive frames.
A1. Optic flow at edge-like structures faces the aperture problem; i.e., using local information, only the normal flow for these structures can be computed.
A2. Only for i2D structures, optic flow estimation can lead to true optic flow estimates by local methods. However, many i2D structures result from depth discontinuities where optic flow estimation algorithms fail to estimate the true motion. In order to get the true motion field, flow algorithms need to deal with at least two different motions in the local area (Bayerl & Neumann 2005).

In this paper, we investigate these claims more closely for several optic flow algorithms (Lucas & Kanade 1981; Nagel & Enkelmann 1986; Gautama & Van Hulle 2002). We will show that the continuous formulation of intrinsic dimensionality allows for a better quantitative investigation and characterization of the quality of optic flow estimation (and hence the optic flow properties as stated in A0–A2) depending on the local signal structure. In this paper, we will show that:

- The algorithms we have tested all had problems with local image structures that were very close to the i0D corner of the iD-triangle (see Figure 2b).
- The performance for image structures in the stripe shaped cluster corresponding to edge-like structures was effected by the aperture problem (see Figure 2b). However, the results depend both quantitatively and qualitatively on the different algorithms and even on different parameters when the same algorithm was used.
- The improvement of performance for signals in the i2D area was visible but small. Average performance increases smoothly and slightly towards the i2D corner (see Figure 2b).

These results support the above-mentioned statements (A0–A2) about optic flow estimation. However, by making use of a continuous understanding of intrinsic dimensionality, these statements can be made quantitatively more specific in terms of (1) characterization of subareas for which they hold and (2) their strength. Our analysis suggests a relationship between the distribution of the signals in the continuous intrinsic dimensionality space and properties of optic flow estimation. In this way, we also introduce a new tool for better analysis of optic flow algorithms.

There has been other works addressing the analysis of errors in optic flow estimates (Simoncelli et al. 1991; Nagel & Haag 1998; Fermueller et al. 2001) In Simoncelli et al.

(1991), using a probabilistic framework for estimating optical flow, it is proven that uncertainty is involved in this estimation process due to several causes such as image noise and inherent limits of motion estimation. In Nagel and Haag (1998), it is shown that gradient-based motion estimation methods underestimate the true flow. In Fermueller et al. (2001), too, it is analytically shown that certain kinds of bias in different classes of optic flow algorithms caused by noise in the image data usually lead not only to underestimate of the magnitude of optical flow and but also to consistent bias in the estimation of the direction. In contrast to the investigations in these studies, in this paper, we are specifically interested in the quality of flow estimates depending on the local image structure. This is done not by analytic means but by statistical comparisons using ground truth data.

The outline of the paper is as follows: in the next section, we introduce the concept of intrinsic dimensionality in its continuous definition as first formulated in Felsberg and Krüger (2003) and Krüger and Felsberg (2003). After this, we investigate the distribution of local image patches according to their intrinsic dimensionality and discuss its consequences for image processing. Then, we look at the statistics of orientation in natural images. Following this, we give an overview of the optic flow algorithms we used in this paper while in the penaltimated section, we look at the statistics of optic flow and the error of optic flow estimation.

Intrinsic dimensionality

When looking at the spectral representation of a local image patch (see Figure 1a and b), we see that the energy of an intrinsically zero-dimensional signal is concentrated in the origin (Figure 1b-top), the energy of an intrinsically one-dimensional signal is concentrated along a line (Figure 1b-middle) while the energy of an intrinsically two-dimensional signal varies in more than one dimension (Figure 1c-bottom).

In image processing, the intrinsic dimensionality was introduced by Zetzsche and Barth (1990) and was used to formalise a *discrete distinction* between edge-like and junction-like structures. This corresponds to a classical interpretation of local image structures in computer vision. A large variety of edge and corner extraction algorithms have been developed over the last twenty years, and their role in artificial as well as biological systems has been discussed extensively (Koenderink & Dorn 1982; Parida et al. 1998). The three kinds of signals have quite specific characteristics and problems that have been addressed in different contexts (Granlund & Knutsson 1995; Jähne 1997).

Recently, it has been shown (Felsberg & Krüger 2003; Krüger & Felsberg 2003) that the topological structure of the intrinsic dimensionality can be understood as a triangle that is spanned by two measures: origin variance and line variance. The origin variance describes the deviation of the energy from a concentration at the origin while the line variance describes the deviation from a line structure (see Figure 1b and 1c); in other words, origin variance measures non-homogeneity of the signal whereas the line variance measures the junctionness. The corners of the triangle then correspond to the 'ideal' cases of intrinsic dimensionality. The surface of the triangle corresponds to signals that carry aspects of the three 'ideal' cases, and the distance from the specific corners indicates the similarity (or dissimilarity) to ideal i0D, i1D and i2D signals.

The triangular topological structure of the intrinsic dimension is counter-intuitive, in the first place, since it realizes a two-dimensional topology in contrast to a linear one-dimensional structure that is expressed in the discrete counting 0, 1 and 2. As shown in Felsberg and Krüger (2003) and Krüger and Felsberg (2003), this triangular interpretation allows for a *continuous formulation* of intrinsic dimensionality in terms of three confidences assigned to

each discrete case. This is achieved by first computing two measurements of origin and line variance. There are different ways to compute these (see Felsberg & Krüger 2003; Krüger & Felsberg 2003). Our approach is based on local magnitude and orientation measurements (see Krüger & Felsberg 2003). Using these measurements, we can associate the local magnitude to the origin variance and associate the variance of the computed orientation in a local neighbourhood to the line variance.[2] These two measurements define a point in the triangle (see Figure 1c). The bary-centric coordinates (see, e.g., Coxeter 1969) of this point of the triangle directly lead to a definition of three confidences that add up to one. These three confidences reflect the volume of the three areas constituted by the points (see Figure 1c). For example, for a point P in the triangle, the area of the sub-triangle i0D-P-i1D denotes the confidence for i2D as shown in Figure 1c. That leads to the decision areas for i0D, i1D and i2D as seen in Figure 1d. We would like to note that the bary-centric coordinates are only one possibly way to distinguish the different structures (see, e.g., Wegmann & Zetzsche 1990).

Statistics of intrinsic dimensionality

We use a set of seven natural sequences with ten images each (see Figure 6). The images have a resolution of 1276×1016. For our statistics, we first compute a measure for the origin and line variance (for details see Krüger and Felsberg 2003). This corresponds to one point in the triangle of Figure 1c. The distribution of the frequency of these points in the triangular structure is shown in Figure 3a. Since there exist large differences in the histogram, the logarithm is shown.

In the rest of the paper, a signal is inD for $n = 0, 1, 2$ if the associated confidence is the highest. That is, a signal will be called i0D, i1D or i2D if the confidence for i0D, i1D or i2D, respectively, is the highest. That means that 'non-ideal' cases of i0D, i1D and i2D structures also become labelled as being inD.

The distribution shows two main clusters. The peak close to the origin corresponds to low origin variance. It is visible that most of the signals that have low origin variance have high line variance. These correspond to nearly homogeneous image patches. Since the orientation is almost random for such homogeneous image patches, it causes high line variance. There

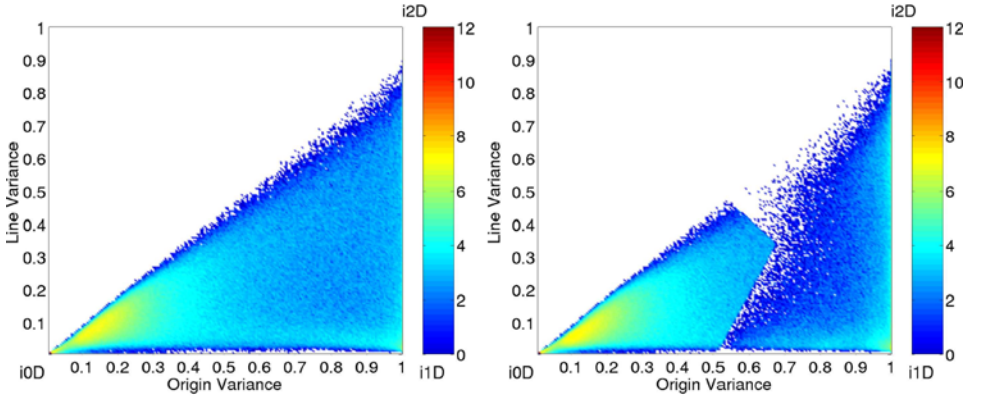

Figure 3. Logarithmic plot of the distribution of intrinsic dimensionality, (a) The distribution for regularly sampled points, (b) The distribution when the positions are modified according to iD (see the text for details of this modification).

is also a small peak at position (0, 0) that corresponds to saturated/black image patches. The other cluster is for high origin variance signals with low line variance, corresponding to edge-like structures. The form of this cluster is a small horizontal stripe rather than a peak. Finally, there is a smooth decrease while approaching to the i2D area of the triangle. That means that there does not exist a cluster for corner-like structures like the ones for homogeneous image patches or edges. Along the origin variance axis, a small continuous gap is observed. This gap suggests that there are no signals with zero line variance. This is due to the fact that at positions with positive origin-variance (i.e., positive magnitude) there is always noise included which causes some line variance.

We also see from the figure that there are far more i0D signals than i1D or i2D signals. That means we have much more homogeneous structures than edge-like or junction-like structures. Besides, it is clear that there are more i1D structures than i2D structures in natural images. The percentages of i0D, i1D and i2D structures are 86%, 11% and 3%, respectively.

Edges or corners are structures that are bound to a specific position. For example, the position of an edge is supposed to be placed directly on the image discontinuity; or, for a corner, in which a certain number of lines intersect, we want the corner to be placed directly on the intersection. We can achieve this positioning by making use of the local amplitude information in the image depending on the intrinsic dimensionality which is described in detail in Krüger et al. (2004). Note also that features such as orientation and optic flow depend on this positioning. When we determine the position for edges and corner-like structures accordingly, we get the distribution shown in Figure 3b. It is qualitatively similar to the distribution achieved with regular sampling. However, since the position is determined depending on the local amplitude (and in this way by maximizing origin variance; see Krüger and Felsberg 2003), there is a shift towards positions with higher amplitude that constitute the gaps at the border between i0D, i1D and i2D signals and the stripe along the i1D–i2D border of the triangle. In the later stages of our analysis, we adopted this positioning. Positioning is illustrated in Figure 4.

Wegmann and Zetzsche (1990) also investigated the distribution of local image structures. They analyzed the multi-dimensional hyperspace which was constructed from all possible combinations of orientation filter outputs. The hyperspace consisted of m axes corresponding to m different orientations such that the origin denoted the homogeneous signals; the axes and the planes between the neighboring axes denoted the i1D structures; and, the planes between the non-neighboring axes denoted i2D signals. Zetzsche and his colleagues could drive proportions of the different local structures (which basically reflect the percentages we have provided above) and visualize clusters of the structures for a few orientation pairs. Due to the complexity of the hyperspace, however, the visualisation becomes more complex than our triangular reresentation.

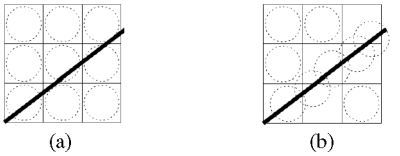

Figure 4. Illustration of positioning for an edge, (a) Without positioning, (b) With positioning as explained in the text.

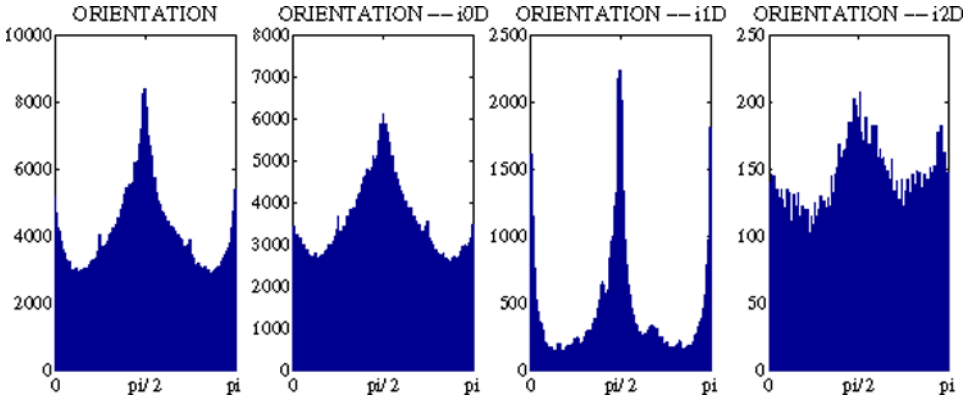

Figure 5. Orientation distribution depending on iD. The first image shows the total distribution. The sequences we have used are introduced in the 'Statistics of intrinsic dimensionality' section.

Distribution of orientation depending on intrinsic dimensionality

It is known (see, e.g., Coppola et al. 1998; Krüger 1998) that the distribution of orientations of i1D signals shows strong peaks at horizontal and vertical structures (i.e., for the values 0, $\pi/2$ and π). However, neither for a completely homogeneous image patch nor for a corner the concept of orientation (although computable) makes sense and the computed orientation is random. Therefore, it is expected that the distribution of orientations of i0D and i2D signals should be homogeneous.

The distribution of the orientation of signals and the quantitative differences depending on the intrinsic dimensionality of the patches is displayed in Figure 5. The figure shows that there are significant peaks for the i0D and i2D signals although they are smaller than the peaks in the distribution of i1D signals. This suggests that orientation is a meaningful

Figure 6. Some of the image sequences used in our analysis. The first 3 images are from one of the sequences (the starting image, the middle image and the last image). Remaining figures are the images from other sequences.

concept for some non-i1D signals, too. This also stresses the advantages of a continuous understanding of intrinsic dimensionality.

Optic flow algorithms

In this section, we give general information about the optic flow algorithms we have used in our analysis. We applied examples of algorithms corresponding to two important classes of optic flow algorithms: differential techniques (Lucas–Kanade and Nagel–Enkelmann) and phase-based techniques.

The Lucas–Kanade algorithm

The Lucas–Kanade algorithm works by minimizing the following functional (Lucas & Kanade 1981) over a spatial neighborhood Ω:

$$\int\int_{\Omega} W^2(x,y)(\nabla I(x,y,t)\cdot \mathrm{v} + I_t(x,y,t))^2 dxdy \tag{1}$$

where $W(x,y)$ is the window function over Ω that gives more influence to constraints at the center of the neighborhood; $\nabla I(x,y,t)$ denotes the intensity gradient at time t at spatial location (x,y); v is the velocity field to be found; and, I_t denotes the derivative of I with respect to t. Basically, the Lucas–Kanade algorithm makes use of the well-known gradient constraint equation $\nabla T^I \cdot \mathrm{v} + I_t = 0$ where weighting is performed over a local neighbourhood.

The Lucas–Kanade is an optic flow algorithm which uses first order derivatives. Due to its smaller complexity when compared with others, it is known as a fast algorithm.

The Nagel–Enkelmann algorithm

The Nagel–Enkelmann algorithm (Nagel & Enkelmann 1986) also makes use of the gradient constraint equation but applies a second order derivative constraint in addition. The following functional is minimized:

$$\int\int (\nabla I^T \mathrm{v} + I_t)^2 + \frac{\alpha^2}{\|\nabla I\|_2^2 + 2\delta}[(u_x I_y - u_y I_z)^2 + (v_x I_y - v_y I_x)^2 + \delta\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right)dxdy \tag{2}$$

where a and δ are constants; $u$and v are respectively the horizontal and the vertical components of the velocity vector v; and, for a function F, F_z denotes the partial derivative of F with respect to variable z. Main terms of the formula are $(u_x I_y - u_y I_x)^2 + (v_x I_y - v_y I_x)^2$ and $(u_x^2 + u_y^2 + v_x^2 + v_y^2)$. The first term smoothes velocity anisotropically, i.e., orthogonal to the intensity gradient. The second isotropic term states that velocity should be constant over position.[3]

Since the Nagel–Enkelmann algorithm can be interpreted as a diffusion process (see Alvarez et al. 2000) with fixed number of iterations, an increase in the number of iterations means an increase in the region of influence used in the computation, and hence, using more global information. The Nagel–Enkelmann algorithm encourages slow variations in the gradient of the vector field by the smoothing term in Equation 2. This leads with increasing number of iterations (i.e., increasing diffusion) naturally to a more regular distribution of directions (as visible in the first two rows of Figure 7). In this study, we will show how using more global information affects the accuracy of the flow estimation.

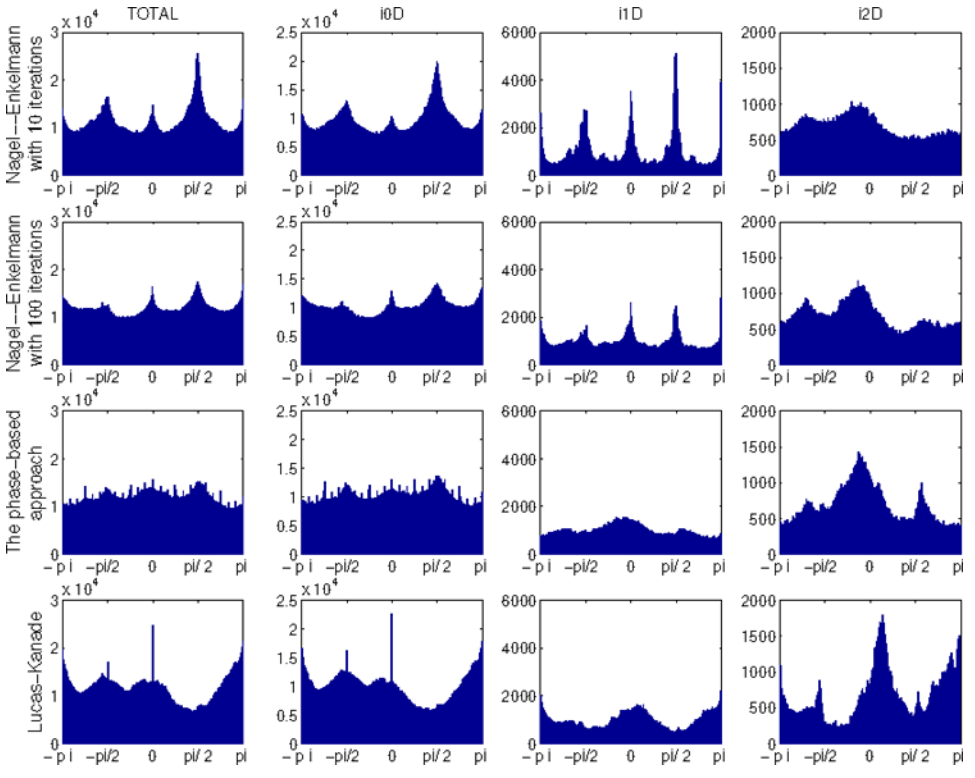

Figure 7. Distribution of direction of optic flow vectors depending on the intrinsic dimension. The histograms show the summed up distributions over the sequences which are introduced earlier. From top to bottom: the Nagel–Enkelmann algorithm with 10 iterations; the Nagel–Enkelmann algorithm with 100 iterations; the phase-based approach; the Lucas–Kanade algorithm. From left to right: the total distribution; the distribution for i0D signals; the distribution for i1D signals; the distribution for i2D signals.

The phase-based approach

Phase-based optic flow algorithms make use of the phase gradient for finding the flow. It has been shown that temporal evolution of contours of constant phase provides a better approximation to local flow (see e.g., Fleet & Jepson 1990). The basic assumption is that phase contours should be constant over time (Fleet & Jepson 1990; Gautama & Van Hulle 2002). This assumption can be formulated as $\phi(x, y, t) = c$, where $\phi(x, y, t)$ denotes the phase component at spatial location (x, y) at time t. Taking differentiation with respect to time, the following constraint is found:

$$\nabla\phi(x, y, t) \cdot (\nabla(x, y), 1) = \nabla\phi(x, y, t) \cdot (\mathrm{v}, 1) = 0. \tag{3}$$

Among phase-based approaches, we chose a recent implementation (Gautama & Van Hulle 2002) in which the constraint (3) is solved for a number of Gabor filters and the flow orthogonal to the orientation of each filter is found. Combining the solutions reached by the filters yields the true displacement. We will show that in this way, even for a large number of i0D signals good optic flow can be estimated.

Optic flow and intrinsic dimensionality

In this section, we investigate the statistics of optic flow direction and the error of optic flow estimation and its relation to the iD triangle.

Optic flow direction

The distribution of the flow direction of the optic flow vectors (using the Nagel–Enkelmann algorithm with 10 and 100 iterations, and the phase-based approach) is shown in Figure 7.[4]

The distribution of the direction varies significantly with the intrinsic dimensionality. The statistics of the true flow can be expected to show some homogeneity since a translational forward motion is dominant in the sequences that leads to a regular flow field (see, e.g., Lappe et al. 1999). A detailed discussion of first order statistics of optic flow in natural scenes can be found in (Calow et al. 2004). They showed that the main factor for irregularity is the large amount of structure near in the lower visual field as compared to the lack of structure in the upper visual field causes larger flow in the lower visual field. This, however, does not effect the magnitude but only the orientation. However, for the Nagel–Enkelmann algorithm with 10 iterations (Figure 7, top row), the distribution of the direction of optic flow vectors of i1D signals directly reflects the distribution of orientation of i1D signals. Since only the normal flow can be computed for ideal i1D signals (using local information only), the dominance of vertical and horizontal orientations leads to peaks at horizontal and vertical flows. The fact that basically there exits a direct quantitative equivalence of the distribution of i1D orientations and the distribution of optic flow directions reflects the seriousness of the aperture problem. In contrast, the distribution of direction of optic flow vectors of i0D and i2D signals is much more homogeneous. When the number of iterations is increased (and hence, more global information is used in the computation of the flow as explained in the previous section), the peaks that correspond to horizontal and vertical lines become smaller (Figure 7, middle row). For the phase-based approach and Lucas–Kanade, a different picture occurs (Figure 7, last two rows): the peaks are less apparent.

As a summary, in Figure 7 we observe that there is a relation between the direction of estimated optic flows and the orientation distribution of signals. However, the strength of this relation depends on the particular algorithm and its parameters. For example, when the used information is very local, the Nagel–Enkelmann algorithm computes basically the normal flow which results in a strong relation between the distribution of optic flow direction and distribution of orientations in the images. However, when the number of iterations is increased, this relation becomes weaker because of the decrease of the aperture effect due to using more global information.

Analysis of quality of optic flow estimation using the intrinsic dimensionality triangle

We now analyze the qualities of the optic flow estimation depending on the intrinsic dimension. For this, we need to compare the computed flow with a ground truth. For our investigations, we used the Brown Range Image Database (BRID), a database of 197 range images collected by Ann Lee, Jinggang Huang and David Mumford at Brown University (see also Huang et al. 2000). The range images are recorded with a laser range-finder.[5] The data of each point consist of four values: the distance, the horizontal angle and the vertical angle in spherical coordinates and a value for the reflected intensity of the laser beam (see Figure 8).

The knowledge about the 3D data structure allows for a simulation of a moving camera in a scene and is used to estimate the correct flow for nearly all pixel positions of a frame of an

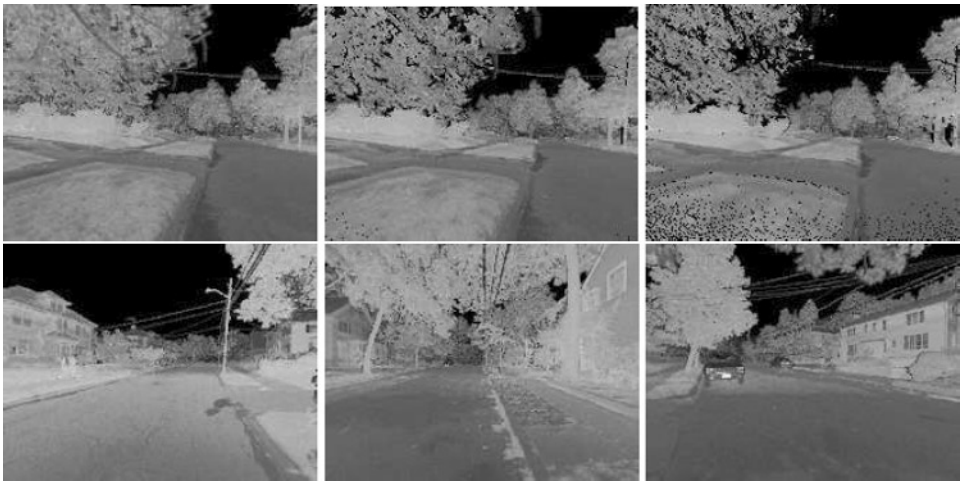

Figure 8. Sequences with ground truth optic flow. Reflected intensity of the laser beam is shown. The first line is from the same sequence (the starting image, the middle image and the last image). Other images are taken from the other sequences.

image sequence. We would like to note that this approach cannot produce correct flow for occlusions.

The simulated motion is forward translation. Different motion types (such as rotation, and rotation and translation) may produce different global motion types. Therefore, the results in this paper are valid only for translational motions, and other types of motions should be expected to yield quantitatively if not qualitatively different results.

Different flow estimation algorithms yield flow fields with different densities; i.e., they can make an estimation of the motion for a certain proportion of the image data. By adjusting the parameters of the flow algorithms we have used, the flow fields were made as dense as possible for our analysis, which happened to be on the average 100%, 90% and 86% respectively for the Nagel–Enkelmann, the Lucas–Kanade algorithms and the phase-based approach.

The quality of optic flow estimation is displayed in a histogram over the iD triangle (see Figures 9, 10). The error is calculated using the well known measure $e(\mathrm{u}, \mathrm{v}) = acos(\frac{\mathrm{u}\cdot\mathrm{v}+1}{\mathrm{u}\cdot\mathrm{u}+1)(\mathrm{v}\cdot\mathrm{v}+1)})$, wnere u and v the flow vectors (see also Barron et al. 1994).[6] We call this error 'combined error'.

For the Nagel–Enkelmann algorithm with ten iterations, the combined error computed using the original ground truth (see Figure 9a) is high for signals close to the i0D corner of the triangle as well as on the horizontal stripe from the i0D to the i1D corner. In the other parts, there is a smooth surface which shows that the error decreases towards the i2D corner (note that the peaks in the middle of the triangle are due to few samples and can be ignored). This is in accordance with the notion that optic flow estimation at corner-like structures is more reliable than for edges and homogeneous image patches (A2). However, in Figure 9a, it becomes obvious that the area where more reliable flow vectors can be computed is very broad and covers also i0D and i1D signals. Furthermore, the decrease of error is rather slight which points to the fact that the quality of flow computation is limited in these areas, as well. When the number of iterations is increased (Figure 9c), the estimation of flow becomes better. In fact, we observe that the area where optic flow is estimated with low error covers almost the whole triangle except some parts in the i1D area of the triangle.

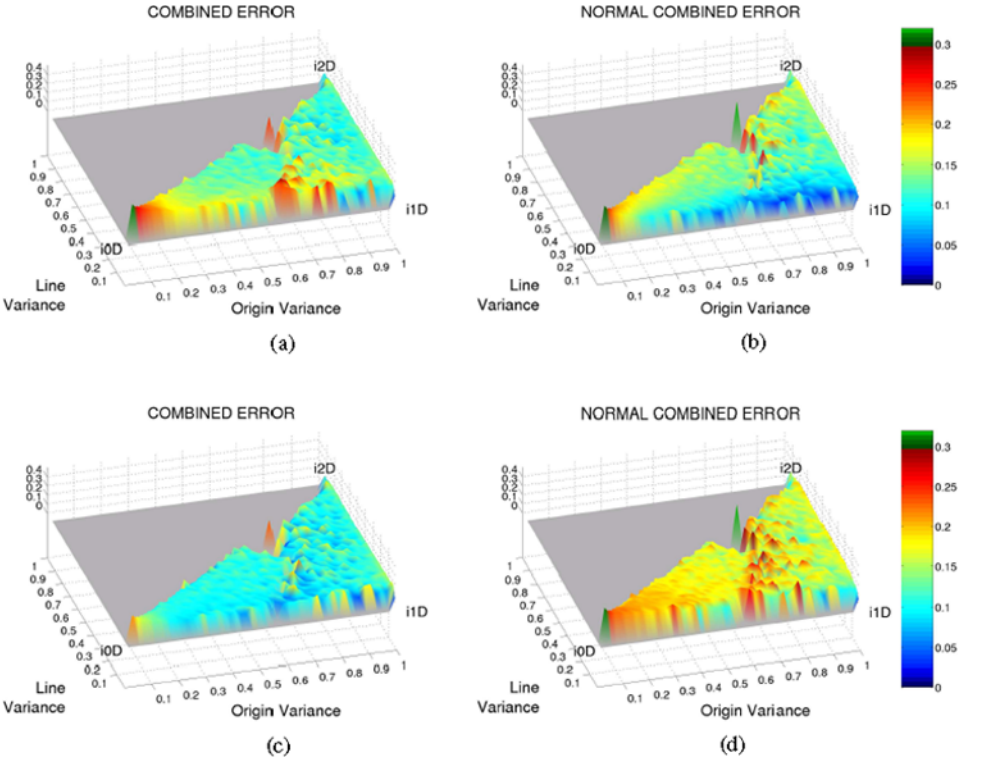

Figure 9. Qualities of the optic flow algorithms depending on iD. Left column shows the errors between computed flow and the ground truth. Right column shows the errors between computed flow and the ground truth projected orthogonal to the orientation of signals. Color bars show the error values for corresponding colors of corresponding graphs. (a,b) Nagel–Enkelmann with 10 iterations. (c,d) Nagel–Enkelmann with 100 iterations.

Among all, the phase-based approach produces the best results for quite a large area (Figure 10c). However, for many i1D signals the estimate is more unreliable than for most i0D and i2D cases. For the Lucas–Kanade algorithm, we again observe that the area where error is low smoothly extends to some i0D and i1D signals (Figure 10a). The figures 9a, 9c, 10a and 10c suggest that this only slight decrease is a general property of the investigated optic flow algorithms.

For better analysis of the aperture problem, we have also computed errors with the projection of the ground truth over the normal vectors (i.e., the true normal flow). For this, we first compute the normal vector of the image patch using the local orientation; then, the ground truth is projected over this vector, and the error is computed between the optic flow vector and this projected ground truth. In Figures 10 and 9, this is called as normal combined error.

For the error computed using the normal ground truth (see Figures 9b, 9d), a different picture occurs. For the Nagel–Enkelmann algorithm with 10 iterations (Figure 9b), the error is very low for a horizontal stripe from the middle point between the i0D and i1D corners to the i1D corner. When compared to Figure 9a, this figure reflects the effect of the aperture problem when only local information is used. When the number of iterations is increased, we see that using more global information decreases the effect of the aperture problem (Figures 9c, 9d). However, it is visible that the quality of the estimated error area in the i1D area of the triangle is always significantly lower than the quality of the estimated normal flow with a small number of iterations (i.e., using only very local information).

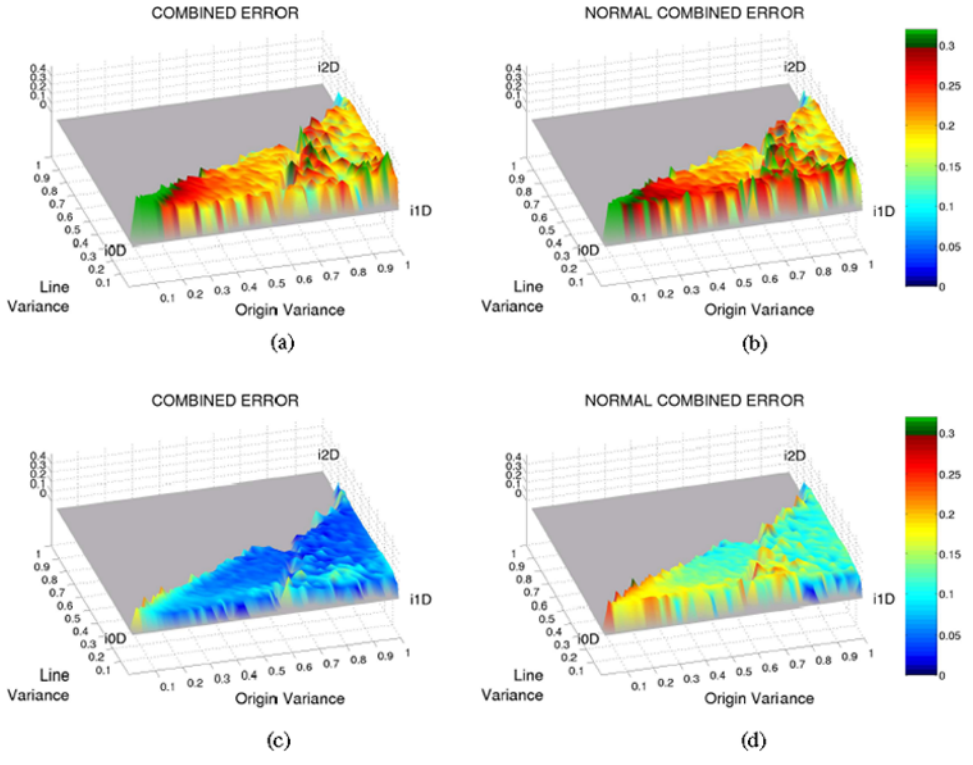

Figure 10. Qualities of the optic flow algorithms depending on iD. Left column shows the errors between computed flow and the ground truth. Right column shows the errors between computed flow and the ground truth projected orthogonal to the orientation of signals. Color bars show the error values for corresponding colors of corresponding graphs. (a,b) The Lucas–Kanade. (c,d) The phase-based approach.

Comparing the results for the Nagel–Enkelmann algorithm with 10 and 100 iterations, we observe that increasing the region of influence means an increase in the overall quality of optic flow estimation. However, it is visible that the error in estimation of flow in the i1D area of the triangle using more global information (Figure 9c) is always significantly higher than the error of the estimated normal flow using more local information (see, Figure 9b), which can be computed with high reliability. The information for such signals is of great importance since (1) there exists a large number of local image patches corresponding to such edge-like structures (see, Figure 3) and (2) constraints for global motion estimation can be defined on line-correspondences (see, e.g., Rosenhahn 2003; Krüger & Wörgötter 2004), i.e., correspondences that only require normal flow. For these tasks, a reliably estimated normal flow might be a better basis than an unreliably estimated true flow.

Discussion

We have analyzed the distribution of local signal patches according to their intrinsic dimensionality. Furthermore, we investigated how the quality of optic flow estimates depends on intrinsic dimensionality.

A continuous understanding of the concept of intrinsic dimension (Felsberg & Krüger 2003; Krüger & Felsberg 2003) allows for a more precise characterization of established structures in terms of their statistical manifestation in natural images (see Figure 2a and Figure 3). The continuous formulation of intrinsic dimensionality also allows for a more

quantitative investigation and characterization of the quality of optic flow estimation depending on local signal structures. We could justify and more precisely quantify generally acknowledged ideas about such estimates (see A0–A2 in the introduction). More specifically, we showed (see also Figure 2b, Figure 9, and 10):

Q0. In general, homogeneous structures lead to low quality optic flow estimation. However, for many i0D signals, true flow can be estimated with good accuracy. In fact, we observe that the algorithms compute the flow with quite a low error for most of the i0D structures.

Q1.1. There exist significantly more horizontal and vertical structures in natural images (see also, e.g., Coppola et al. 1998; Krüger 1998). In this paper, we show that the strength of the dominance of these structures depends crucially on the intrinsic dimension. Furthermore, we show that the distribution of orientations is directly reflected in the distribution of the estimates of optic-flow directions as an effect of the aperture problem. However, we observe that the degree of this reflection is very much dependent on the particular optic flow algorithm and the parameters used.

Q1.2. The optic flow estimates in the stripe-shaped cluster with high origin–variance and low line–variance (corresponding to edges) are in general worse than for i2D signals for all investigated algorithms. However, for the Nagel–Enkelmann algorithm, normal flow can be estimated reliably for this stripe-shaped cluster in the i1D signal domain. We think that it is important to note that this reliably computed normal flow is an important information as such. For example, line–line correspondences that can be derived from the normal flow play an important role in Rigid Body Motion estimation (see, e.g., Krüger & Wörgötter 2004; Shevlin 1998; Rosenhahn & Sommer 2002). Using the Nagel–Enkelmann algorithm, we showed how using more global information affects the quality of optic flow estimation for i1D signals.

Q2. The quality of optic flow estimation is higher for i2D signals. However, in analogy to the lack of a cluster for i2D signals, there exists a continuous signal domain (covering also sub-areas of i0D and i1D signals) for which a higher quality in the optic flow estimation can be achieved. The increase of the quality, on the other hand, is only slight which suggests that the role of i2D structures for motion estimation might not be as important as suggested by some authors (see, e.g., Mota & Earth 2000). By using different optic flow algorithms, we show that this is a general property of optic flow algorithms.

We want to stress that our aim was not to find the 'best' algorithm but to make general properties of optic flow algorithms explicit. In particular, we do not want to make a statement about a superiority of phase-based approaches to differential techniques. For example, we have started to investigate a very recently developed algorithm (Brox et al. 2004) related to the class of differential based algorithms which showed superior performance to the algorithms discussed here. The choice of the 'best' algorithm depends a lot on the context determined by time constraints or hardware constraints. Furthermore, we could show that even different parameter settings leading to qualitatively different estimates are plausible for different tasks.

Acknowledgments

We would like thank Joachim Weickert and Michael Felsberg for fruitful discussions, and Fabio Solari (and his student Javier Diaz), Temujin Gautama and Marc van Hulle for supplying us with computed flow fields for the Lucas–Kanade and the phase-based approach, respectively. This work is supported by the ECOVISION project.

End notes

1. For convenience, intrinsic dimension of zero, intrinsic dimension of one, intrinsic dimension of two and intrinsic dimensionality (and intrinsic dimension) will be called i0D, i1D, i2D and iD, respectively, in the rest of the paper.
2. Note that a continuous interpretation can be derived also from the structure tensor (Jähne 1997) by computing origin and line variance using the two eigenvalues.
3. In our simulations, we used the standard values 0.5 and 1.0 for α and δ, respectively, as suggested and usually practiced in the literature (see, e.g., Barron et al. (1994)).
4. We also analyzed the distribution of magnitude, but as expected found no dependence on the intrinsic dimensionality (for details, see Kalkan et al. (2004)).
5. Each image contains 44 $\times$ 1440 measurements with an angular separation of 0.18 degree. The field of view is 80 degree vertically and 259 degree horizontally. The distance of each point is calculated from the time of flight of the laser beam, where the operational range of the sensor is 2–200 m. The laser wavelength of the laser beam is 0.9 μm in the near infrared region.
6. We also made measurements using angular and magnitudal errors with the formulas $e_{ang}(\mathrm{u},\mathrm{v}) = acos(\frac{\mathrm{u}\cdot\mathrm{v}}{|\mathrm{u}|\cdot|\mathrm{v}|})$, $e_{mag}(\mathrm{u},\mathrm{v}) = (\frac{abs|\mathrm{u}|-|\mathrm{v}|}{|\mathrm{u}|+|\mathrm{v}|})$, respectively. However since they show similar patterns they are not shown in this paper (for details, see Kalkan et al. (2004)).

References

Alvarez L, Weickert J, Sanchez J. 2000. Reliable estimation of dense optical flow fields with large displacements. *Int. J. Computer Vision* 39:41–56.

Barron JL, Fleet DJ, Beauchemin SS. 1994. Performance of optical flow techniques. *Int J Computer Vision* 1:43–77.

Bayerl P, Neumann H. Disambiguating visual motion by form—motion interaction—a computational model. *Int. J. Computer Vision (Special Issue)* (in press).

Brox T, Bruhn A, Papenberg N, Weickert J. 2004. High accuracy optical flow estimation based on a theory for warping. *Proceeding of ECCV 2004*.

Calow D, Krüger N, Wörgötter F, Lappe M. 2004. Statistics of optic flow for self-motion through natural scenes. *Proc. Dynamic Perception Workshop.* pp. 133–138.

Canny J. 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8:679–698.

Cavanagh P, Mather G. 1989. Motion: the long and the short of it. *Spatial Vision* 4:103–129.

Coppola DM, Purves HR, McCoy AN, Purves D. 1998. The distribution of oriented contours in the real world. *PNAS* 4002–4006.

Coxeter HSM. 1969. *Introduction to Geometry (2nd ed.)*. Chichester, Wiley & Sons.

Felsberg M. 2002. *Low-Level Image Processing with the Structure Multivector.* PhD thesis, Institute of Computer Science and Applied Mathematics, Christian-Albrechts-University of Kiel.

Felsberg M, Krüger N. 2003. A probablistic definition of intrinsic dimensionality for images. *Pattern Recognition, 24th DAGM Symposium*.

Fermueller C, Shulman D, Aloimonos Y. 2001. The statistics of optical flow. *Computer Vision and Image Understanding* 82:1–32.

Fleet DJ, Jepson AD. 1990. Computation of component image velocity from local phase information. *Int J Computer Vision* 5:77–104.

Gautama T, Van Hulle MM. 2002. A phase-based approach to the estimation of the optical flow field using spatial filtering. *IEEE Trans. on Neural Networks* 13:1127–1136.

Granlund GH, Knutsson, H. 1995. *Signal Processing for Computer Vision.* Dordrecht, Kluwer Academic Publishers.

Huang J, Lee AB, Mumford D. 2000. Statistics of range images. *CVPR* 1:1324–1331.

Hubel DH, Wiesel TN. 1969. Anatomical demonstration of columns in the monkey striate cortex. *Nature* 221:747–750.

Jähne, B. 1997. *Digital Image Processing—Concepts, Algorithms, and Scientific Applications.* New York, Springer.

Johnston A, Clifford CWG. 1995. A unified account of three apparent motion illusions. *Vision Res* 35:1109–1123.

Kalkan S, Calow D, Felsberg M, Wörgötter F, Lappe M, Krüger N. 2004. Optic flow statistics and intrinsic dimensionality. *Proc. of Brain Inspired Cognitive Systems, Scotland, available at http://www.cs.stir.ac.uk/lss/BICS2004/CD/toc.html*.

Koenderink JJ, Dorn AJ. 1982. The shape of smooth objects and the way contours end. *Perception* 11:129–173.

Krüger N. 1998. Collinearity and parallelism are statistically significant second order relations of complex cell responses. *Neural Processing Letters* 8:117–129.

Krüger N, Felsberg M. 2003. A continuous formulation of intrinsic dimension. *Proceedings of the British Machine Vision Conference*.

Krüger N, Felsberg M, Wörgötter F. 2004. Processing multi-modal primitives from image sequences. *Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS*.

Krüger N, Wörgötter F. 2004. Statistical and deterministic regularities: Utilisation of motion and grouping in biological and artificial visual systems. *Advances in Imaging and Electron Physics* 131:82–147.

Lappe M, Bremmer F, van den Berg AV. 1999. Perception of self-motion from visual flow. *Trends in Cognitive Sciences* 3:329–336.

Lucas B, Kanade T. 1981. An iterative image registration technique with an application to stereo vision. *Proc. DARPA Image Understanding Workshop* pp. 121–130.

Mota C, Barth E. 2000. On the uniqueness of curvature features. *Proc. in Artificial Intelligence* 9:175–178.

Nagel H-H, Enkelmann W. 1986. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8:565–593.

Nagel H-H, Haag M. 1998. Bias-corrected optical flow estimation for road vehicle tracking. *Proc. International Conference on Computer Vision*, Bombay, India pp. 1006–1011.

Parida L, Geiger D, Hummel R. 1998. Junctions: detection, classification and reconstruction. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 20:687–698.

Princen J, Illingworth J, Kittler J. 1990. An optimizing line finder using a Hough transform algorithm. *Computer Vision, Graphics, and Image Proc* 52:57–77.

Ribeiro E, Hancock ER. 2001. Shape from periodic texture using the eigenvectors of local affine distortion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23:1459–1465.

Rohr K. 1992. Recognizing corners by fitting parametric models. *Int J Computer Vision* 9:213–230.

Rosenhahn B. 2003. *Pose Estimation Revisited PhD Thesis.* Institut für Informatik und praktische Mathematik, Chrsitian—Albrechts Universität Kiel.

Rosenhahn B, Sommer G. 2002. Adaptive pose estimation for different corresponding entities. In L. van Gool, editor, *Pattern Recognition, 24th DAGM Symposium* pp. 265–273. Springer Verlag.

Shevelev IA, Lazareva NA, Tikhomirov AS, Sharev GA. 1995. Sensitivity to cross-like figures in the cat striate neurons. *Neuroscience* 61:965–973.

Shevlin F. 1998. Analysis of orientation problems using Plücker lines. *Int Conference on Pattern Recognition, Bisbane* 1:65–689.

Simoncelli EP, Adelson EH, Heeger DJ. 1991. Probability distributions of optical flow. *Proc. IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii* pp. 310–315.

Sochen N, Kimmel R, Malladi R. 1998. A general framework for low level vision. *IEEE Transactions on Image Processing* 7:310–318.

Wegmann B, Zetzsche C. 1990. Statistical dependence between orientation filter outputs used in a human-vision-based image code. In *Proc. SPIE Vol. 1360, p. 909–923, Visual Communications and Image Processing '90: Fifth in a Series, Murat Kunt; Ed.* pp. 909–923, September 1990.

Zetzsche C, Barth E. 1990. Fundamental limits of linear filters in the visual processing of two dimensional signals. *Vision Research* 30:1111–1117.

Zetzsche C, Barth E, Berkmann J. 1991. Spatio-temporal curvature measures for flow field analysis. *Geometric Methods in Computer Vision* 1570:337–350.