

Tensor-based Image Sequence Processing Techniques for the Study of Dynamical Processes

Horst HAUSSECKER, Hagen SPIES and Bernd JÄHNE

Interdisciplinary Center for Scientific Computing, IWR
University of Heidelberg
Im Neuenheimer Feld 368, D-69120 Heidelberg
GERMANY

E-mail: {horst.haussecker,hagen.spies,bernd.jaehne}@iwr.uni-heidelberg.de

Commission V, Working Group IC V/III

KEY WORDS: Image Sequence Analysis, Optical Flow, Structure Tensor, Filter Optimization

ABSTRACT

Quantitative analysis of dynamical processes requires a precise estimation of the optical flow field from image sequences. Most articles evaluating the performance of optical flow techniques focus on the initial formulation of the minimization problem to solve the ill posed brightness change constraint equation. Performance differences are attributed to slight differences in the formulation of the minimization and the numerical solution is taken for granted. It can be shown, however, that most differential techniques can be formulated in a generalized way and be solved by two major numerical estimation techniques: least squares and total least squares. We will conclude that the least squares technique does only vary two of three parameters of the spatiotemporal optical flow vector while the latter varies all three parameters which leads to a more precise solution. Total least squares estimation of optical flow is equivalent to a tensor representation of the spatiotemporal image structure. In addition to the optical flow, measures of certainty and type of motion, quantifying the presence of an aperture problem are directly obtained by analyzing the eigenvalues of the so called *structure tensor*. These measures give a unified perspective of common quality measures proposed by various techniques. Another crucial factor influencing the accuracy of any differential technique is the choice of appropriate differential kernels. With optimized differential filters, errors can be reduced by more than one order of magnitude.

ZUSAMMENFASSUNG

Die quantitative Analyse von dynamischen Prozessen erfordert eine präzise Schätzung des optischen Flusses in Bildsequenzen. Die meisten vergleichenden Übersichtsartikel zu verschiedenen Techniken beschränken sich auf die Formulierung des Minimierungsproblems zur Lösung der unterbestimmten Kontinuitätsgleichung des optischen Flusses. Genauigkeitsunterschiede der Verfahren werden den leicht unterschiedlichen Formulierungen des Minimierungsproblems zugeschrieben und die numerische Lösung des resultierenden Problems als gegeben hingenommen. Es zeigt sich jedoch, daß sich eine verallgemeinerte Formulierung der meisten differentiellen Verfahren finden läßt, die durch zwei grundlegende numerische Techniken gelöst werden kann: 'least squares' und 'total least squares'. Während das Standard 'least squares' Verfahren nur zwei Parameter des raum-zeitlichen optischen Flusses variiert, werden beim 'total least squares' Verfahren alle drei Komponenten variiert, was zu einem präziseren Ergebnis führt. Eine 'total least squares' Schätzung des optischen Flusses ist äquivalent zu einer Tensor-Repräsentation der raum-zeitlichen Grauwertstruktur. Zusätzlich zum optischen Fluß liefert eine Analyse der Eigenwerte des sogenannten Strukturtenors Maße der Zuverlässigkeit und der Ausprägtheit des Blendenproblems. Diese Maße vereinheitlichen gebräuchliche Qualitätsmaße anderer Techniken. Ein weiterer kritischer Faktor, im Hinblick auf die Genauigkeit des Verfahrens, ist die Wahl geeigneter Ableitungsfiler. Durch optimierte Ableitungsfiler läßt sich eine Reduktion der Fehler in der Bewegungsschätzung um mehr als eine Größenordnung erreichen.

1. INTRODUCTION

Motion analysis remains one of the fundamental problems in image sequence processing. The only accessible motion parameter from image sequences is the optical flow, an approximation of the two-dimensional motion field on the image sensor. The optical flow field can be used as input for a variety of subsequent processing steps including motion detection, motion compensation, three-dimensional surface reconstruction, autonomous navigation and the analysis of dynamical processes in scientific applications. As only the apparent motion in the sequence can be extracted, further a priori assumptions on the constancy of image brightness and the relation between relative three dimensional scene motion and the projection onto the two-dimensional image sensor are necessary for quantitative scene analysis.

In contrast to the more qualitative requirements of standard computer vision applications, such as motion detection or collision avoidance, quantitative measurement tasks require precise and dense optical flow fields in order to reduce the propagation of errors in subsequent processing steps. In addition to the optical flow field, measures of confidence have to be provided to discard erroneous data points and quantify measurement precision.

Quantitative image sequence analysis requires the entirety of quantitative visualization, geometric and radiometric calibration and a quantitative error analysis of the entire chain of image processing algorithms. The final results are only as precise as the least precise part of the system. Quantitative visualization of object properties is up to the special requirements of applications and cannot be discussed in general. Without doubt, camera calibration is an important step towards quantitative image analysis and has been extensively investigated by the photogrammetric society. This article will focus on the algorithmic aspects of low-level motion estimation in terms of performance and error sensitivity of individual parts, given a calibrated image, eventually corrupted by sensor noise. It will be shown how a combination of radiometric uniformity correction, filter optimization and careful choice of numerical estimation techniques can significantly improve the overall precision of low-level motion estimation. Starting with the brightness change constraint equation (Section 2), we will show how a local estimate on optical flow can be obtained by using a weighted standard least squares estimation proposed by Lucas and Kanade (1981) (Section 3). This technique can be improved by using total least squares estimation instead of standard least squares. This directly leads to a tensor representation of the spatiotemporal brightness distribution, such as the structure tensor technique (Haussecker and Jähne, 1997, Haussecker, 1998) (Section 4). In this section we will further detail how a fast and efficient implementation can be achieved by using standard image processing

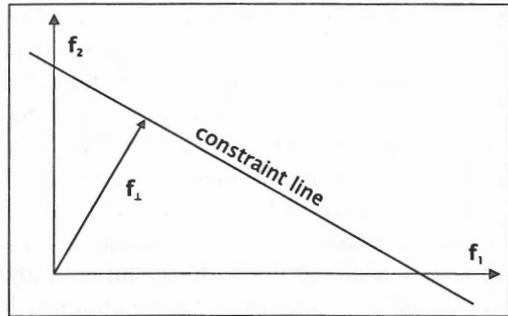


Figure 1: Illustration of the constraint line defined by (1). The normal optical flow vector, \mathbf{f}_\perp , is pointing perpendicular to the line and parallel to the local gradient $\nabla g(\mathbf{x}, t)$.

operators, which is an important requirement for dynamic analysis. Coherency and type measures are obtained from the solution of the structure tensor technique in a straightforward way. They allow to quantify the confidence of the optical flow estimation as well as the presence of an aperture problem. It will be shown how they compare to other measures, previously proposed by Barron et al. (1994) and Simoncelli (1993). In Sections 5 and 6 we will show how optimization of derivative filters and uniformity correction significantly improve the performance of any differential technique. We will conclude with results from both test patterns and application examples in Section 7 and a final discussion in Section 8.

2. OPTICAL FLOW CONSTRAINT

A common assumption on optical flow is that the image brightness $g(\mathbf{x}, t)$ at a point $\mathbf{x} = [x, y]^T$ at time t should be conserved. Thus, the total temporal derivative, dg/dt , needs to equal zero, which directly yields the well known *brightness change constraint equation*, BCCE (Horn and Schunk, 1981):

$$\frac{dg}{dt} = (\nabla_{\mathbf{x}}g)^T \mathbf{f} + g_t = 0, \quad (1)$$

where $\mathbf{f} = [f_1, f_2]^T$ is the optical flow, $\nabla_{\mathbf{x}}g$ defines the spatial gradient, and g_t denotes the partial time derivative $\partial g/\partial t$.

This relation poses a single local constraint on the optical flow at a certain point in the image. It is, however, ill posed as (1) constitutes only one equation of two unknowns. This problem is commonly referred to as the *aperture problem* of motion estimation, illustrated in Figure 1. All vectors along the constraint line defined by (1) are likely to be the real optical flow \mathbf{f} . Without further assumptions only the flow \mathbf{f}_\perp perpendicular to the constraint line can be estimated. In order to solve this problem a variety of approaches have been proposed that try to minimize an objective function pooling constraints over a small finite area. An excellent overview of optical flow techniques is given by Barron et al. (1994). They

conclude that differential techniques, such as the local weighted least squares method proposed by Lucas and Kanade (1981) perform best in terms of efficiency and accuracy. Phase-based methods (Fleet and Jepson, 1990) show slightly better accuracy but are less efficient in implementation and lack a single useful confidence measure. Bainbridge-Smith and Lane (1997) come to the same conclusion comparing the performance of differential methods. Performing analytical studies of various motion estimation techniques, Jähne (1993, 1997) showed that the three-dimensional structure tensor technique yields best results with respect to systematic errors and noise sensitivity. This could be verified by Jähne et al. (1998), analyzing a calibrated image sequence with ground truth data provided by Otte and Nagel (1994).

3. LOCAL WEIGHTED LEAST SQUARES

Lucas and Kanade (1981) propose a local weighted least squares estimate of the constraint (1) on individual pixels within a local spatial neighborhood U by minimizing:

$$\int_{-\infty}^{\infty} h(\mathbf{x} - \mathbf{x}') ((\nabla_{\mathbf{x}}g)^T \mathbf{f} + g_t)^2 d\mathbf{x}' \quad (2)$$

with a weighting function $h(\mathbf{x})$. In practical implementations the weighting is realized by a Gaussian smoothing kernel. Minimizing (2) with respect to the two components f_1 and f_2 of the optical flow \mathbf{f} yields the standard least squares solution

$$\underbrace{\begin{bmatrix} \langle g_x g_x \rangle & \langle g_x g_y \rangle \\ \langle g_x g_y \rangle & \langle g_y g_y \rangle \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}}_{\mathbf{f}} = - \underbrace{\begin{bmatrix} \langle g_x g_t \rangle \\ \langle g_y g_t \rangle \end{bmatrix}}_{\mathbf{b}}, \quad (3)$$

with the abbreviation

$$\langle a \rangle = \int_{-\infty}^{\infty} h(\mathbf{x} - \mathbf{x}') a d\mathbf{x}'. \quad (4)$$

The solution of (3) is given by $\mathbf{f} = \mathbf{A}^{-1}\mathbf{b}$, provided the inverse of \mathbf{A} exists. If all gradient vectors within U are pointing into the same direction, \mathbf{A} gets singular and the aperture problem remains within the local neighborhood. These cases can be identified by analyzing the eigenvalues of the symmetric matrix \mathbf{A} prior to inversion (Barron et al., 1994, Simoncelli, 1993) and only the normal flow f_{\perp} is computed from (1) by $f_{\perp} = -g_t / \|\nabla g\|$. It is, however, a critical issue to obtain information about the presence of an aperture problem from the numerical instability of the solution. Thresholds on the eigenvalues proposed by Barron et al. (1994) and Simoncelli (1993) have to be adapted to the image content which prevents a versatile implementation.

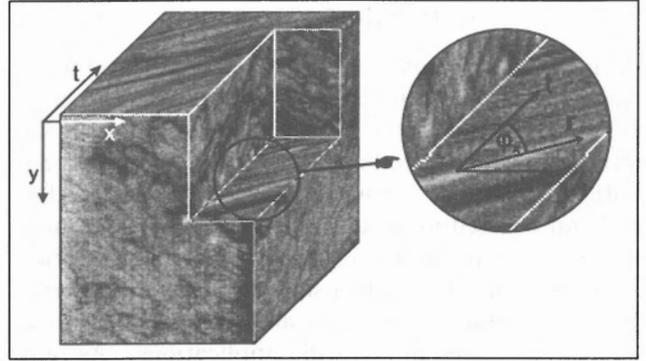


Figure 2: Illustration of the spatiotemporal brightness distribution of moving patterns. The sequence shows infrared images of the ocean surface moving mainly in positive x-direction.

Jähne (1997) shows that an extension of the integration in (2) into the temporal domain yields a better local regularization if the optical flow is modeled constant within the spatiotemporal neighborhood U . This does, however, not change the minimization procedure and results in the same linear equation system (3) with spatiotemporal integration of the individual components $\langle a \rangle$.

From a probabilistic point of view, the minimization of (2) corresponds to a maximum likelihood estimation of the optical flow, given Gaussian distributed errors at individual pixels. Black and Anandan (1996) show that the Gaussian assumption does not hold for motion discontinuities and transparent motions. By replacing the least squares estimation with robust statistics they come up with an iterative estimation of multiple motions.

4. STRUCTURE TENSOR APPROACH

The displacement of gray value structures within consecutive images of a sequence yields inclined structures with respect to the temporal axis of spatiotemporal images (Figure 2).

The orientation of iso-grey-value lines within a three-dimensional spatiotemporal neighborhood U can mathematically be formulated as the direction $\mathbf{r} = [r_1, r_2, r_3]^T$ being as much perpendicular to all grey value gradients ∇g in U as possible. The direction \mathbf{r} and the optical flow \mathbf{f} are related by $\mathbf{f} = r_3^{-1} [r_1, r_2]^T$. With the definition of \mathbf{r} , (1) can be formulated as:

$$[g_x, g_y, g_t] \begin{bmatrix} f_1 \\ f_2 \\ 1 \end{bmatrix} = r_3^{-1} (\nabla_{\mathbf{x}t}g)^T \mathbf{r} = 0, \quad (5)$$

where $\nabla_{\mathbf{x}t}g$ denotes the spatiotemporal gradient vector $\nabla_{\mathbf{x}t}g = [g_x, g_y, g_t]^T$. It is important to note that (1) and (5) are mathematically equivalent formulations and no constraint is added by extending the formulation of the brightness change constraint into three-dimensional space.

The direction \mathbf{r} can be found by minimizing

$$r_3^{-2} \int_{-\infty}^{\infty} h(\mathbf{x} - \mathbf{x}') ((\nabla_{\mathbf{x}t}g)^T \mathbf{r})^2 d\mathbf{x}', \quad (6)$$

which is mathematically equivalent to (2). Solving the quadratic terms, (6) can be written as matrix equation

$$\mathbf{r}^T \mathbf{J} \mathbf{r} \rightarrow \text{minimum}, \quad (7)$$

with the three-dimensional structure tensor

$$\mathbf{J} = \begin{bmatrix} \langle g_x g_x \rangle & \langle g_x g_y \rangle & \langle g_x g_t \rangle \\ \langle g_x g_y \rangle & \langle g_y g_y \rangle & \langle g_y g_t \rangle \\ \langle g_x g_t \rangle & \langle g_y g_t \rangle & \langle g_t g_t \rangle \end{bmatrix}. \quad (8)$$

The components of \mathbf{J} are given by

$$J_{pq} = \langle g_p g_q \rangle = \int_{-\infty}^{\infty} h(\mathbf{x} - \mathbf{x}') g_p g_q d\mathbf{x}'. \quad (9)$$

Again, the spatial integration can be extended into the time domain for local regularization without changing the results of the following minimization procedure (Jähne, 1997).

In order to avoid the trivial solution of (7) the constraint $\|\mathbf{r}\|_2 = 1$ has to be imposed on \mathbf{r} . Solving (7) by the method of Lagrangian multipliers gives the solution that the minimum of (7) is reached, if the vector \mathbf{r} is given by the eigenvector of the tensor \mathbf{J} to the minimum eigenvalue. This method is known as orthogonal L^2 approximation and can be shown to be mathematically equivalent to total least squares estimation.

It is important to note that the difference between the least squares method of Lucas and Kanade (1981) and the structure tensor formulation is neither imposed by the formulation of the minimization nor by the extension into the temporal domain but rather by the minimization procedure. While least squares estimation only varies the objective function with respect to the two components of \mathbf{f} , the total least squares technique varies all three components of the spatiotemporal vector \mathbf{r} under the constraint $\|\mathbf{r}\|_2 = 1$.

4.1 Computing the structure tensor

The implementation of the tensor components can be carried out very efficiently by standard image processing operators. Identifying the convolution in (9) with a three-dimensional spatiotemporal smoothing of the product of partial derivatives, each component of the structure tensor can be computed as

$$J_{pq} = \mathcal{B}(\mathcal{D}_p \cdot \mathcal{D}_q), \quad (10)$$

with the 3D spatio-temporal smoothing operator \mathcal{B} and the differential operator \mathcal{D}_p in the direction of the

coordinate x_p . Using a binomial operator the smoothing can be performed very. A more critical point is the choice of an appropriate differential operator. It can be shown that derivative filters optimized for a minimum deviation from the correct direction of the gradient vector reduce the error by more than an order of magnitude as compared to standard differential operators (Section 5).

4.2 Eigenvalue analysis

A standard procedure in numerical eigenvalue analysis is the Jacobi transformation (Press et al., 1992). The Jacobi method is absolutely foolproof for all real symmetric matrices. This is very advantageous, because it does not depend on the image content. In order to speed up the whole eigenvalue analysis, a major decrease in computation time can be achieved by pre-selecting interesting image regions by thresholding the trace of the matrix, $\text{trace}(\mathbf{J}) = J_{xx} + J_{yy} + J_{tt}$, for each point before starting the Jacobi transformation.

4.3 Computing displacements

Different classes of 3D spatio-temporal structures can be identified without explicitly solving the eigenvalue problem. The structure tensor contains the entire information on the first-order structure of the grey value function in a local neighborhood. By analyzing the rank of the matrix four different cases of spatio-temporal structures can be distinguished (Table 1). The two extreme cases of $\text{rank}(\mathbf{J}) = 0$ and $\text{rank}(\mathbf{J}) = 3$ represent no apparent linear motion.

In the case of $\text{rank}(\mathbf{J}) = 1$ an already oriented image structure moves with a constant velocity. This is the well known *aperture problem* in optical flow computation. Only one of the three eigenvectors has an eigenvalue larger than zero. This eigenvector $\vec{e}_i = (e_{i,x}, e_{i,y}, e_{i,t})$ points normal to the plane of constant grey value in 3D space and can be used to compute the normal optical flow f_{\perp} :

$$f_{\perp} = -\frac{e_{i,t}}{\sqrt{e_{i,x}^2 + e_{i,y}^2}}. \quad (11)$$

For $\text{rank}(\mathbf{J}) = 2$ an isotropic grey value structure moves with a constant velocity. No aperture problem is present in the spatio-temporal neighborhood. The orientation of the 3D iso-grey-value line yields the two components f_1 and f_2 of the optical flow. The eigenvector $\vec{e}_s = (e_{s,x}, e_{s,y}, e_{s,t})$ to the smallest eigenvalue pointing into the direction of the line corresponds to the initially defined vector \mathbf{r} , and \mathbf{f} can be computed as:

$$\mathbf{f} = \begin{pmatrix} e_{s,x} & e_{s,y} \\ e_{s,t} & e_{s,t} \end{pmatrix}. \quad (12)$$

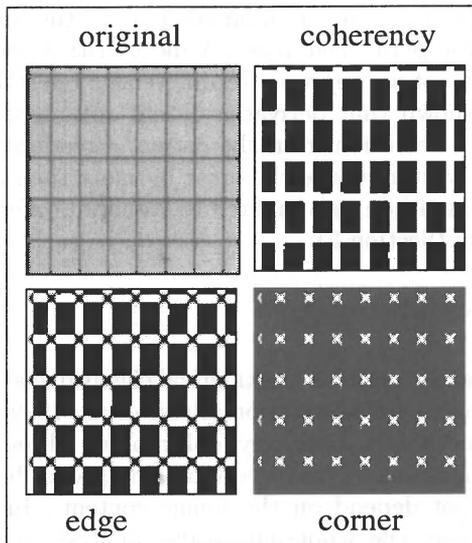


Figure 3: Illustration of the coherency and type measures for a moving grid on a linear positioning table.

4.4 Coherency and type measures

Although the rank of the structure tensor proves to contain all information necessary to distinguish different types of motion it can not be used for practical implementations because it does not constitute a normalized measure of certainty. Additionally it is only defined for integer values $0, \dots, 3$. In real sequences usually mixtures between the types of motion occur. In this section we will introduce two normalized measures to quantify the confidence and type of motion which are suited for practical implementations. In contrast to the rank they yield multivalued numbers between 0.0 and 1.0.

Coherency (*coh*): In order to quantify the overall certainty of displacement estimation we define the normalized coherency measure

$$coh = \left(\frac{\lambda_l - \lambda_s}{\lambda_l + \lambda_s} \right)^2, \quad (13)$$

with λ_l and λ_s denoting the largest and smallest eigenvalue of the structure tensor, respectively. Table 1 shows the values of *coh* for different cases of motion. For both types of apparent motion, i. e. aperture problem and no aperture problem, *coh* is equal one. If no displacement can be computed it is identical zero. With increasing noise level *coh* approaches zero for all different types of motion.

Edge measure (*edge*): While the coherency *coh* gives a normalized estimate for the certainty of the computation it does not allow to identify areas of apparent aperture problem. In order to quantify the presence of an aperture problem we define the edge measure

$$edge = \left(\frac{\lambda_l - \lambda_m}{\lambda_l + \lambda_m} \right)^2, \quad (14)$$

where λ_m denotes the medium eigenvalue of the structure tensor. Table 1 shows the values of *edge* for different cases of motion. Only if an aperture problem is present *edge* reaches its maximum value of *coh*. For all other types of motion it is equal zero. Note that the edge measure is normalized between 0.0 and *coh* since it is not possible to detect the presence of an aperture problem more reliably than the overall certainty.

motion type	coh	edge	rank
homogeneous brightness	0	0	0
aperture problem	1	1	1
no aperture problem	1	0	2
no coherent motion	0	0	3

Table 1: Rank, coherency and edge measure for different types of motion.

Corner measure (*corner*): From the two independent measures *coh*, and *edge*, a corner measure can be computed by $corner = coh - edge$. It constitutes the counterpart of *edge*, quantifying the absence of an aperture problem, i. e. selecting points where both components f_1 and f_2 of the optical flow \mathbf{f} can be reliably computed.

Figure 3 illustrates the coherency and type measures for a moving calibration target on a linear positioning table. The coherency measure shows the entire grid without regions of homogeneous brightness. These areas split up into the edges and crossing points of the grid for the edge and corner measure, respectively.

The importance of coherency and type measures for quantitative analysis are illustrated in Figure 4. The optical flow field of a moving ring pattern shows random flow vectors in homogeneous regions with additive noise. With the coherency measure these regions can be identified. Further knowledge about the presence of an aperture problem allows to reconstruct the real flow field using the local regularization technique proposed in (Haussecker and Jähne, 1997).

Confidence measures for the local least squares approach have also been defined by Barron et al. (1994) and Simoncelli (1993). Both try to quantify the singularity of the matrix \mathbf{A} in (3) by analyzing the eigenvalues of \mathbf{A} . While Simoncelli (1993) proposes to threshold the sum of eigenvalues, Barron et al. (1994) argue that the smallest eigenvalue proved to be more reliable in practical implementations. The matrix \mathbf{A} constitutes the structure tensor of a two-dimensional subspace of the spatiotemporal neighborhood and represents local orientation in two-dimensional images (Bigün and Granlund, 1987, Knutsson, 1998). The two-dimensional equivalent to the three-dimensional coherency measure is the difference of the two eigenvalues, normalized to the sum of the eigenvalues (Jähne, 1993, 1997). Rather than thresholding parts of the information, this measure quantifies the con-

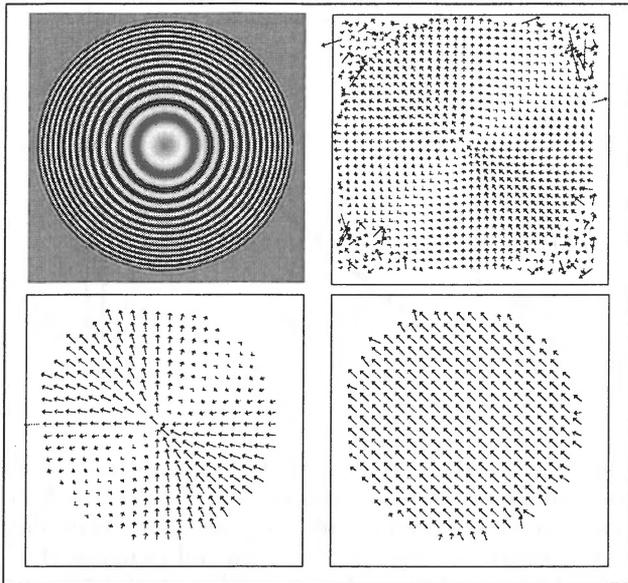


Figure 4: Illustration of the importance of confidence and type measures for a ring pattern moving with (1,1) pixels/frame towards the upper left corner (with additive noise $\sigma_n = 1$). Upper row: One frame of the moving ring pattern (left), optical flow computed for any image point without confidence and type measures (right). Lower row: Optical flow masked by the confidence measure (left), local regularization incorporating confidence and knowledge about the presence of an aperture problem (right).

confidence of local orientation and therefore the presence of an aperture problem. It also approaches zero for homogeneous brightness. The three-dimensional coherency constitutes a generalization of the two-dimensional case. It also includes information on the coherency of motion and identifies isotropic noise patterns.

5. FILTER OPTIMIZATION

A crucial factor in optical flow computation is the discretization of the partial derivative operators. Figure 5 illustrates this basic fact with a simple numerical study. With the standard symmetric difference filter $1/2 [1 \ 0 \ -1]$, large deviations from the correct displacements of more than 0.1 pixels/frame occur. With an

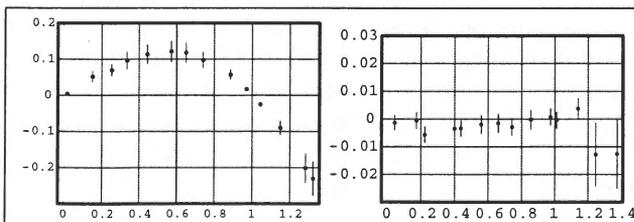


Figure 5: Systematic error in the velocity estimate as a function of the interframe displacement of a moving random pattern. Derivatives computed with the symmetric difference filter $0.5 [1 \ 0 \ -1]$ (left) and an optimized Sobel filter (right) given by (Schar et al., 1997).

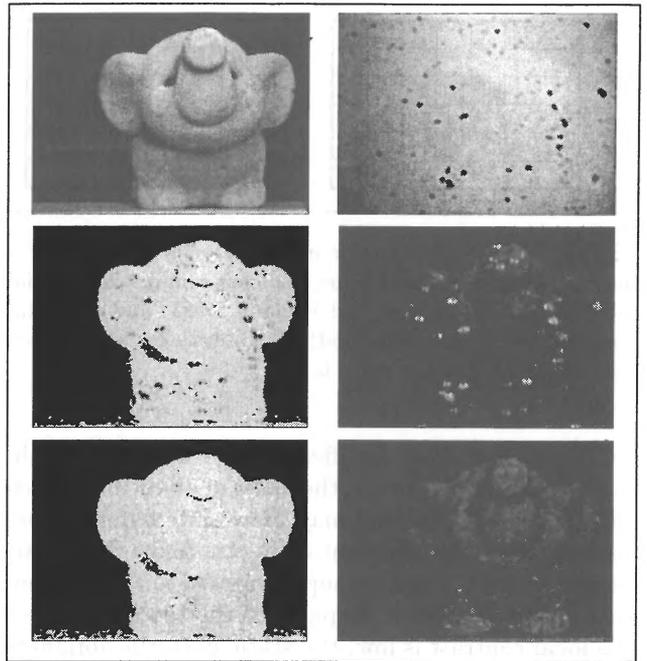


Figure 6: Demonstration of the influence of spatial sensitivity variations of the CCD sensor on motion estimation: Top row: One image of the elephant sequence (left), contrast enhanced relative responsivity (right). Middle row: Velocity component in x direction (left), smallest eigenvalue of the structure tensor (right). Bottom row: Velocity component in x direction for the corrected sequence (left), smallest eigenvalue of the structure tensor for a sequence corrected for the spatial responsivity changes (right).

optimized 3×3 Sobel-type filter (Schar et al., 1997), the error is well below 0.005 pixels/frame.

Analyzing the impact of noise on differential least-squares techniques, Bainbridge-Smith and Lane (1997) report the same results for the errors as in the left image in Figure 5. This error can be identified as discretization error of the differential operators. In order to reduce these errors they use a combination of smoothing and derivative kernels for local regularization in one dimension.

We used a new class of regularized separable first-order derivative filters with a transfer function

$$\hat{D}_p(kt) = (ik_p)\hat{B}(|kt|) \quad (15)$$

for a filter in the direction p . A similar approach has been used by (Simoncelli, 1993). While Simoncelli (1993) applied a linear error functional, a nonlinear one has been used by (Schar et al., 1997) minimizing the angle error of the gradient vector.

6. SENSOR UNIFORMITY CORRECTION

Errors of motion analysis with real sensor data are significantly higher than those obtained with computer generated sequences. The higher errors are related to imperfections of the CCD sensor/camera system. A radiometric calibration study showed that standard

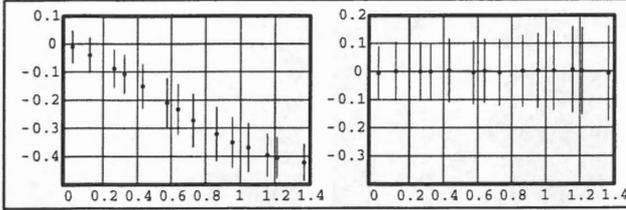


Figure 7: Systematic error in the velocity estimate of a moving random pattern using the least squares differential method (left) of (Lucas and Kanade, 1981) and the total least squares (tensor method) of (Haussecker and Jähne, 1997). Signal to noise ratio is one.

CCD cameras show significant large-scale and small-scale spatial variations in the order of about 1%, which cannot be neglected and may show quite different spatial variations for different cameras. Since these patterns are static, they are superimposed to the real motion in the sequence. In parts of the sequence where the local contrast is low, the static patterns dominate the structure and thus a lower or even zero velocity is measured.

The influence of such static patterns can nicely be demonstrated for moving objects with low contrast such as the slightly textured elephant in 6. Dirt on the glass window of the CCD sensor causes spatial variations in the responsivity of the sensor. At the edges of the speckles, the smallest eigenvalue of the structure tensor shows high values indicating motion discontinuities. The motion field indeed shows drops at the positions of the speckles. If a simple two-point calibration is performed using the measured responsivity and an image with a dark pattern, the influence of the speckles is no longer visible both in the smallest eigenvalues and the motion field.

7. RESULTS

Figure 7 shows a comparison of the least squares approach of Lucas and Kanade (1981) and the total least squares (structure tensor) technique. While the least squares technique shows a bias for increasing velocities, the total least squares technique does not show a bias and performs better for all velocities within the range of the temporal sampling theorem.

In order to prove the theoretical error limits of the technique in real sequences, accuracy tests have been performed. Defining the *local noise portion* (LNP) as the ratio between the noise variance σ_n^2 and the local variance $var_B(g_n) = \mathcal{B}(g_n)^2 - (\mathcal{B}g_n)^2$ of the noisy image g_n

$$LNP = \frac{\sigma_n^2}{var_B(g_n)} = \frac{\sigma_n^2}{var_B(g) + \sigma_n^2}, \quad (16)$$

we get a normalized measure quantifying the presence of noise within a local neighborhood. A value of $LNP = 0.5$ means that the noise variance lies in the same

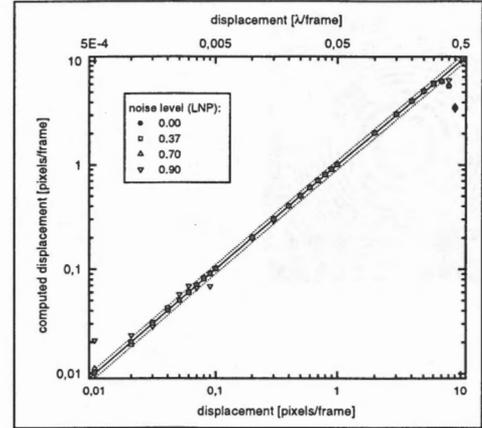


Figure 8: Subpixel accuracy of the tensor method. A sinusoidal pattern with a wavelength of 20 pixels has been shifted from 0.01 pixels/frame up to the theoretical limit of 10 pixels/frame given by the temporal sampling theorem. The contribution of noise to the signal within the spatiotemporal region of support of the filters is quantified by the local noise portion, LNP. The dashed lines indicate relative errors of $\pm 5\%$.

order as the local variance $var_B(g)$ of the noise-free pattern g .

Figure 8 shows the results of the averaged optical flow for a sinusoidal test pattern with additive noise of different LNPs. The result exhibits the high subpixel accuracy of the structure tensor technique. The structure tensor technique was applied to a variety of application examples from oceanography (IR ocean surface images), botany (growth processes) and traffic scenes. It proved to work well without any adaption to the image content. Figure 9 shows examples of such sequences. As a nice detail the pedestrian in figure 9 is about to lift his right leg which is clearly visible in the optical flow field.

8. CONCLUSIONS

The structure tensor technique allows to compute dense displacement vector fields from extended image sequences with high subpixel accuracy. Additionally, it yields a measure of certainty as well as a measure for the type of motion quantifying the presence of an aperture problem. The structure tensor technique proved to work well for a variety of applications. The measures of coherency and aperture problem allow the technique to automatically adapt to local spatiotemporal image structures without additional parameters.

ACKNOWLEDGEMENTS: We gratefully acknowledge financial support by the ‘Deutsche Forschungsgemeinschaft’, DFG, within the frame of the research unit ‘Image Sequence Analysis to Investigate Dynamic Processes’.

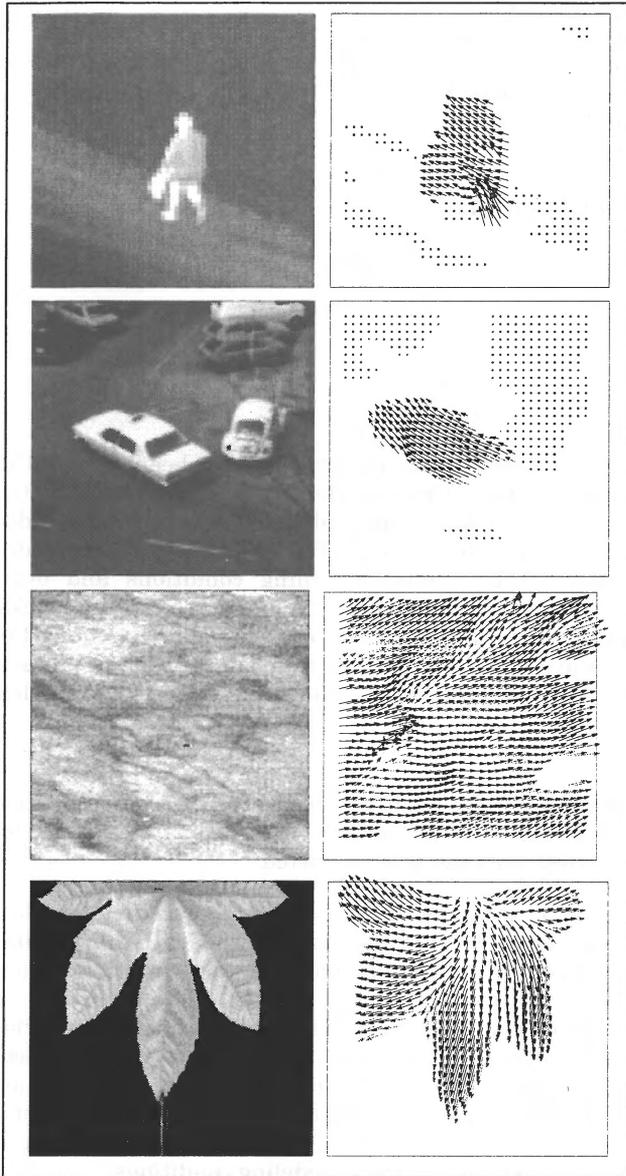


Figure 9: Application examples. From top to bottom: IR image of a pedestrian, Hamburg Taxi Scene, IR image of the ocean surface, Growing leaf of a castor-oil plant. The intensity of the arrows quantifies the coherency *coh*. A black arrow symbolizes high certainty.

REFERENCES

- Bainbridge-Smith, A., and R. G. Lane, 1997. Determining Optical Flow Using a Differential Method. *Image and Vision Computing*, (15) pp. 11-22.
- Barron, J. L., D. J. Fleet, and S. S. Beauchemin, 1994. Performance of Optical Flow Techniques. *Intern. J. Comput. Vis.* (12:1), pp. 43-77.
- Bigün, J. and G. H. Granlund, 1987. Optimal Orientation Detection of Linear Symmetry. *Proc. First Intern. Conf. on Computer Vision, ICCV, London, England, June 8-11.*
- Black, M. J., and P. Anandan, 1996. The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields. *Computer Vision and Image Understanding*, (63:1), pp. 75-104.
- Fleet, D. J. and A. D. Jepson, 1990. Computation of Component Image Velocity from Local Phase Information. *Intern. J. Comput. Vis.* (5) pp. 77-104.
- Haussecker, H., and B. Jähne, 1997. A Tensor Approach for Precise Computation of Dense Displacement Vector Fields. *Proc. Mustererkennung 1997*, F. Wahl and E. Paulus (eds.), Informatik Aktuell, Springer-Verlag, Heidelberg.
- Haussecker, H., 1998. Motion. In: *Handbook on Computer Vision and Applications Volume 2: Signal Processing and Pattern Recognition*, edited by B. Jähne, H. Haussecker and P. Geissler, Academic Press, Boston, to be published.
- Jähne, B., 1993. *Spatio-Temporal Image Processing, Theory and Applications. Lecture Notes in Computer Science*, (751), Springer, Berlin.
- Jähne, B., 1997. *Digital Image Processing: Concepts, Algorithms and Scientific Applications*, 4th ed., Springer-Verlag, Heidelberg.
- Jähne, B., H. Haussecker, H. Spies, D. Schmundt, and U. Schurr, 1998. Study of Dynamical Processes with Tensor-Based Spatiotemporal Image Processing Techniques. *Proc. Computer Vision - ECCV '98*, Springer-Verlag, in press.
- Knutsson, H., 1989. Representing Local Structure using Tensors. *Proc. 6th Scandinavian Conf. Image Analysis*, Oulu, Finland, pp. 244-251.
- Lucas, B., and T. Kanade, 1981. An Iterative Image Registration Technique with an Application to Stereo Vision. *Proc. DARPA Image Understanding Workshop*, pp. 121-130.
- Otte, M., and H.-H. Nagel, 1994. Optical Flow Estimation: Advances and Comparisons. In: J.-O. Eklundh (ed.), *Computer Vision - ECCV '94*, Springer-Verlag, pp. 51-60.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, 1992. *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York.
- Scharr, H., S. Körkel, and B. Jähne, 1997. Numerische Isotropieoptimierung von FIR-Filtern mittels Querglättung. *Proc. Mustererkennung 1997*, F. Wahl and E. Paulus (eds.), Informatik Aktuell, Springer-Verlag, Heidelberg.
- Simoncelli, E. P., 1993. *Distributed Representation and Analysis of Visual Motion. Ph.D. dissertation*, Dept. Of Electrical Engineering and Computer Science, MIT.