Performance Characteristics of Low-Level Motion Estimators in Spatiotemporal Images

Bernd Jähne

Research Group Image Processing

Interdisciplinary Center for Scientific Computing

Heidelberg University, Im Neuenheimer Feld 368, 69120 Heidelberg

Bernd.Jaehne@iwr.uni-heidelberg.de

Abstract

The performance characteristics of gradient-based, tensor-based, and phase-based low-level motion estimators are studied analytically. All three techniques yield exact results under ideal conditions (constant motion field and perfect discretization). With appropriate derivative filters, accuracy well below a hundredth pixel/frame can be achieved. All techniques are quite insensitive to inhomogeneous and accelerated motion. The differential method, however, is biased by noise while the tensor method is not biased by noise. Under certain conditions, the phase method is insensitive against global illumination changes.

1 Introduction

This paper presents an analytical, numerical, and experimental study of the performance of low-level motion estimators in spatiotemporal images. Motivation for this work arose from scientific applications of image sequence processing within the frame of an interdisciplinary research unit where the study of transport, exchange, and growth processes with various imaging techniques requires highly accurate velocity estimates [Jähne *et al.*, 1996]. These high accuracy demands triggered a revisit of the basic theory of motion estimation in spatiotemporal images.

In this paper only low-level motion estimators are discussed. This is, of course, only a part of the picture. However, systematic errors in low-level estimators will propagate and thus cause also systematic errors in subsequent processing. Only a few systematic studies of the performance characteristics of low-level motion estimators are available in the literature. *Kearney et al.* [1987] performed an error analysis of optical flow estimation with gradient-based methods, while *Barron et al.* [1994] used a set of computer-generated and real image sequence to compare various approaches to compute optical flow. *Haglund and Fleet* [1994] used a natural image and warped it to study the performance of phase-based and energy-based techniques.

None of the above cited papers includes an analytical study of the performance of different techniques. This is the main new point in this paper. The paper starts with a brief outline of differential, tensorial, and phase-based techniques formulated

in a unified framework as nonlinear filter operations in continuous spatiotemporal images in Sect. 2. This section forms the basis for the discussion of their performance characteristics. Section 3 discusses the exactness of the techniques under ideal conditions with some remarks on systematic errors introduced by the discretization of the filter operations. Sect. 4 deals with the bias introduced by noise. In the remaining sections a systematic study of the influence of deviations from a constant motion field on the accuracy of the motion estimate is performed: Unsteady motion (Sect. 5), motion discontinuity (Sect. 6), and illumination changes (Sect. 7).

2 Motion Estimator as Nonlinear Filters in Spatiotemporal Images

The basic term for a unified approach to low-level motion estimation in spatiotemporal images $g(\mathbf{x}, t)$ contains the product of derivatives into the different coordinate directions averaged over a certain window $w(\mathbf{x}, t)$ [Jähne, 1993]:

$$\begin{aligned} J_{pq}(\mathbf{x}) &= \int w(\mathbf{x} - \mathbf{x}', t - t') \partial_p g(\mathbf{x}', t') \partial_q g(\mathbf{x}', t') d^2 \mathbf{x}' dt' \\ &= \langle \partial_p g(\mathbf{x}, t) \partial_q g(\mathbf{x}, t) \rangle. \end{aligned} \quad (1)$$

In the lower part of the equation, the convolution integral is abbreviated with angle brackets. The operations in (1) can be performed as a cascade of linear convolution and (nonlinear) point operations:

$$\mathcal{B}(\mathcal{D}_p \cdot \mathcal{D}_q), \quad (2)$$

where $\mathcal{B}$ and $\mathcal{D}_p$ are a smoothing filter and a derivative filter into the direction p .

The standard least squares approach with the differential method results in the linear equation system for the optical flux $\mathbf{f} = (f_x, f_y)$:

$$\begin{bmatrix} J_{xx} & J_{xy} \\ J_{xy} & J_{yy} \end{bmatrix} \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} J_{xt} \\ J_{yt} \end{bmatrix}. \quad (3)$$

The tensor method additionally uses the component J_{tt} to form the symmetric tensor

$$\mathbf{J} = \begin{bmatrix} J_{xx} & J_{xy} & J_{xt} \\ J_{xy} & J_{yy} & J_{yt} \\ J_{xt} & J_{yt} & J_{tt} \end{bmatrix} \quad (4)$$

An eigenvalue analysis of this tensor yields the optical flow and a characterization of the spatial structure [Bigün and Granlund, 1987; Jähne, 1990 and 1993]. In case of a distributed spatial structure and a constant motion, only one eigenvalue of $\mathbf{J}$

is zero. The optical flow can then be computed from the corresponding eigenvector $\mathbf{e}_3 = (e_{3x}, e_{3y}, e_{3t})$ as [Karlholm, 1996]

$$\mathbf{f} = \frac{1}{e_{3t}} \begin{bmatrix} e_{3x} \\ e_{3y} \end{bmatrix} \quad (5)$$

For a spatially oriented structure (e.g. an edge), only the velocity component perpendicular to the oriented structure can be computed from the eigenvector to the largest eigenvalue $\mathbf{e}_1 = (e_{1x}, e_{1y}, e_{1t})$ as

$$|\mathbf{f}_\perp| = \sqrt{\frac{e_{1t}^2}{e_{1x}^2 + e_{1y}^2}} = \sqrt{\frac{e_{1t}^2}{1 - e_{1t}^2}}. \quad (6)$$

The phase method, introduced by *Fleet and Jepson* [1990], deviates in as much from the previous methods that only normal velocities are determined from images filtered by a set of directional quadrature filters (q_+ and q_-). Then the component of the optical into the direction of the directional quadrature filter can be computed by

$$f_\perp = -\frac{q_+ \partial_t q_- - q_- \partial_t q_+}{q_+ \partial_x q_- - q_- \partial_x q_+}. \quad (7)$$

The analysis performed in this paper is very general without making any specific assumptions about the types of gray value structures. Most importantly, it does not use Taylor expansions of the gray scale structure which give only approximate results. The following two classes of general continuous gray value structures are applied:

$$\begin{array}{ll} \text{planar wave, moving "edge"} & g(\mathbf{k}\mathbf{x} - \omega t) \\ \text{constant 2-D motion, moving "corner"} & g(\mathbf{x} - \mathbf{u}_0 t). \end{array} \quad (8)$$

These elementary classes will then be modified systematically by adding various types of noise and by including unsteady motion, motion discontinuities, and illumination changes.

It is important to note that the analytical study of continuous signals does not include errors due to discretization. Partitioning of the error analysis into an analytic and a discrete part has significant advantages. For example, it is possible to distinguish principal flaws in an approach from errors in the implementation due to an inadequate numerical method. Often these two types of errors are mixed up in the literature.

3 Exactness

In this section, we will analyze possible systematic errors of the various motion estimators discussed in Sect. 2. We will start with what appears to be trivial — but is not — and verify that the techniques indeed deliver exact results under ideal conditions. The computations also give further insight into the properties of the different techniques.

Figure 1: Systematic error in the velocity estimate as a function of the interframe displacement measured with a moving random pattern. Derivatives computed with **a** the symmetric difference filter $1/2 [1 \ 0 \ -1]$ **b** an optimized Sobel filter [Schar et al., 1997]. The error bars give the 1-sigma standard deviation when the velocity estimate is averaged over the whole image.

3.1 Differential Method

We take an arbitrary spatial gray value structure moving with a constant velocity, $g'(\mathbf{x}, t) = g(\mathbf{x} - \mathbf{u}t)$. The first-order partial derivatives are then given by $\nabla g'_x = \nabla g_x$ and $\partial_t g' = -\mathbf{u} \nabla g$ and we obtain

$$\begin{aligned}
 \mathbf{f} &= -\langle \nabla g' \nabla g'^T \rangle^{-1} \langle \partial_t g' \nabla g \rangle \\
 &= -\langle \nabla g \nabla g^T \rangle^{-1} \langle -\mathbf{u} (\nabla g \nabla g^T) \rangle \\
 &= \mathbf{u} \langle \nabla g \nabla g^T \rangle^{-1} \langle \nabla g \nabla g^T \rangle \\
 &= \mathbf{u},
 \end{aligned}$$

provided that the inverse of $\langle \nabla g \nabla g^T \rangle$ exists. ■

These computations are significant, since they do not depend at all on the *specific form* of the gray value structure. With $g(\mathbf{x} - \mathbf{u}t)$, we only made the assumption that an arbitrary spatial structure is moving with a constant speed $\mathbf{u}$.

It is a widespread misconception, which has also found its way into monographs on image sequence processing, that differential methods do not deliver accurate results if the spatial gray value structure cannot be adequately approximated by a first-order Taylor series (see for example [Singh, 1991]). Kearney et al. [1987] provided an error analysis of the gradient approach and concluded that it gives erroneous results as soon as second-order spatial derivatives become significant.

As shown above, correct results are obtained as soon as the linear equation system can be solved at all. Errors in the estimation of optical flow are only introduced by an inadequate discretization of the partial derivative operators to the extent that the

Figure 2: Systematic error in the velocity estimate as a function for a pure dilatational motion of a random pattern. The velocity is zero in the middle and ± 1 at the edges. Shown is the error in the velocity component in x direction. Derivatives computed with **a** the symmetric difference filter $1/2 [1 \ 0 \ -1]$ **b** an optimized 3×3 Sobel-type filter [Schar et al., 1997].

discrete differential operators deviate from the transfer function of an ideal derivative operator. With the standard symmetric difference filter $1/2 [1 \ 0 \ -1]$, deviations up than 0.1 pixels/frame occur Figure 1a. Interestingly, the errors cancel out at certain velocities (0 and 1 pixels/frame). This is due to the effect that in the solution of (3) only quotients of temporal and spatial derivatives occur. Thus if one of them is zero, e.g., at zero velocity, or temporal and spatial derivatives are equal, e.g. at 1 pixel/frame velocity, the errors cancel out. The error is maximal at a velocity of about 0.5 pixels/frame. The error of the estimate depends on the local wave number and frequency of the local neighborhood since the deviation of the derivative filters from the ideal transfer function of a derivative filter depends on the wave number.

Actually, it is not required that the transfer function of an optimal derivative filter for motion estimation approaches those of the ideal derivative filter, ik_q . As mentioned already above, the derivative filters occur only in quotients. This is the case for *all* low-level motion estimators based on derivative filters. Therefore the derivative filters in the different coordinate directions may contain an arbitrary *common* function:

$$\hat{D}_x = ik_x B(k, \omega), \quad \hat{D}_y = ik_y B(k, \omega), \quad \hat{D}_t = i\omega B(k, \omega). \quad (9)$$

For the function $B(k, \omega)$ typically an isotropic smoothing function is used in order to take advantage of the fact that the derivative filters can be regularized in this way as well. The optimization of this general type of derivative filters is discussed in detail by Schar et al. [1997]. It turns out that an optimized 3×3 Sobel-type filter is already sufficient to suppress the error of the velocity estimates well below 0.005 pixels/frame (Figure 1b).

For a dilatational motion the error is about the same for the simple symmetric derivative filter (Fig. 2a), while it is slightly higher for the optimized filter (Fig. 2b).

The slight increase in the error in this case becomes clear with the discussion in Sect. 6.

3.2 Tensor Method

To prove the results for the tensor method, we use a constant motion with and without a spatially oriented pattern and perform the eigenvalue analysis symbolically. For an oriented pattern (planar wave) $g'(\mathbf{x}, t) = g(\mathbf{k}\mathbf{x} - \omega t)$ moving along with the normal velocity $\mathbf{k}\omega / (k_x^2 + k_y^2)$,

$$\partial_x g' = k_x \partial_r g, \quad \partial_y g' = k_y \partial_r g, \quad \partial_t g' = -\omega \partial_r g, \quad (10)$$

where $\partial_r g = \partial g / \partial r$ with $r = \mathbf{k}\mathbf{x} - \omega t$. With (10), the six independent components of the structure tensor are

$$\begin{aligned} J_{xx} &= k_x^2 \langle g_r^2 \rangle & J_{yy} &= k_y^2 \langle g_r^2 \rangle & J_{xy} &= k_x k_y \langle g_r^2 \rangle \\ J_{xt} &= -k_x \omega \langle g_r^2 \rangle & J_{yt} &= -k_y \omega \langle g_r^2 \rangle & J_{tt} &= \omega^2 \langle g_r^2 \rangle. \end{aligned} \quad (11)$$

Leaving out the constant mean square gradient $\langle g_r^2 \rangle$, the structure tensor has the form

$$\begin{bmatrix} k_x^2 & k_x k_y & -k_x \omega \\ k_x k_y & k_y^2 & -k_y \omega \\ -k_x \omega & -k_y \omega & \omega^2 \end{bmatrix}. \quad (12)$$

It is too tedious to compute the eigenvalues and eigenvectors of the above symmetric tensor manually. But it can be performed easily with software for symbolic mathematical computations such as Mathematica from Wolfram Research. The result is that two eigenvalues are zero and that the eigenvector of the nonzero eigenvalue is of the form

$$\mathbf{e}_1 = \begin{bmatrix} -k_x \\ -k_y \\ \omega \end{bmatrix}. \quad (13)$$

With this equation we can verify (6) ■.

With a constant moving spatially distributed pattern $g'(\mathbf{x}, t) = g(\mathbf{x} - \mathbf{u}t)$ the derivatives are

$$\partial_x g' = \partial_x g, \quad \partial_y g' = \partial_y g, \quad \partial_t g' = -u_x \partial_x - u_y \partial_y. \quad (14)$$

With (1) and (14), the six independent components of the structure tensor are

$$\begin{aligned} J_{xx} &= \langle (\partial_x g)^2 \rangle \\ J_{yy} &= \langle (\partial_y g)^2 \rangle \\ J_{xy} &= \langle \partial_x g \partial_y g \rangle \\ J_{xt} &= -u_x \langle (\partial_x g)^2 \rangle - u_y \langle \partial_x g \partial_y g \rangle \\ J_{yt} &= -u_x \langle \partial_x g \partial_y g \rangle - u_y \langle (\partial_y g)^2 \rangle \\ J_{tt} &= u_x^2 \langle (\partial_x g)^2 \rangle + 2u_x u_y \langle \partial_x g \partial_y g \rangle + u_y^2 \langle (\partial_y g)^2 \rangle. \end{aligned} \quad (15)$$

Evaluating the eigensystem of this tensor with Mathematica shows that one eigenvalue is zero and that the eigenvector of the zero eigenvalue has the form

$$\mathbf{e}_3 = \frac{1}{u_x^2 + u_y^2 + 1} \begin{bmatrix} u_x \\ u_y \\ 1 \end{bmatrix}, \quad (16)$$

verifying (5). ■

3.3 Phase Method

To verify the phase method, we apply an arbitrary 1-D spatial gray value structure moving with constant velocity. We use

$$g'_+(x, t) = g_+(x - ut) \quad \text{and} \quad g'_-(x, t) = g_-(x - ut).$$

Then (7) results in

$$\begin{aligned} f &= -\frac{\langle -u g'_+ \partial_x g_- + u g'_- \partial_x g_+ \rangle}{\langle g'_+ \partial_x g_- - g'_- \partial_x g_+ \rangle} \\ &= u \frac{\langle g'_+ \partial_x g_- - g'_- \partial_x g_+ \rangle}{\langle g'_+ \partial_x g_- - g'_- \partial_x g_+ \rangle} \\ &= u, \end{aligned}$$

provided that $\langle g'_+ \partial_x g_- - g'_- \partial_x g_+ \rangle \neq 0$. ■

This result is surprising because we do not make any use of the fact that g_- is the Hilbert transform of g_+ . Consequently, the phase method gives accurate results with any pair of g_+ and g_- if only the requirement $\langle g'_+ \partial_x g_- - g'_- \partial_x g_+ \rangle \neq 0$ is met.

This result is also important considering the fact that it is difficult to design an effective Hilbert filter [Jähne, 1997a]. A rough approximation will already give good results.

4 Bias by Noise

Noise may introduce a systematic error into the flow estimate. We investigate this by adding isotropic zero-mean noise

$$g'(\mathbf{x}, t) = g(\mathbf{x} - \mathbf{u}t) + n(\mathbf{x}, t).$$

We assume that the partial derivatives of the noise function are not correlated with themselves or with the partial derivatives of the image patterns. Therefore we use the conditions:

$$\langle n \rangle = 0, \quad \langle (\partial_p n)^2 \rangle = \sigma_n^2, \quad \langle \partial_p n \partial_q n \rangle = 0, \quad p \neq q, \quad \text{and} \quad \langle \partial_p g \partial_q n \rangle = 0 \quad (17)$$

and the partial derivatives are

$$\nabla g' = \nabla g + \nabla n \quad \text{and} \quad \partial_t g' = -\mathbf{u} \nabla g + \partial_t n. \quad (18)$$

4.1 Differential Method

A noisy image with the conditions as set up above gives the optical flow estimate

$$\mathbf{f} = \mathbf{u}(\langle \nabla \mathbf{g} \nabla \mathbf{g}^T \rangle + \langle \nabla \mathbf{n} \nabla \mathbf{n}^T \rangle)^{-1} \langle \nabla \mathbf{g} \nabla \mathbf{g}^T \rangle. \quad (19)$$

The key to understanding this matrix equation is to observe that the noise matrix ($\nabla \mathbf{n} \nabla \mathbf{n}^T$) is diagonal in any coordinate system, because of the conditions set by (17). Therefore, we can transform the equation into the principal-axes coordinate system in which $\nabla \mathbf{g} \nabla \mathbf{g}^T$ is diagonal. Then we obtain

$$\mathbf{f} = \mathbf{u} \begin{bmatrix} \langle g_{x'}^2 \rangle + \sigma_n^2 & 0 \\ 0 & \langle g_{y'}^2 \rangle + \sigma_n^2 \end{bmatrix}^{-1} \begin{bmatrix} \langle g_{x'}^2 \rangle & 0 \\ 0 & \langle g_{y'}^2 \rangle \end{bmatrix}.$$

The inverse of the first matrix always exists when the variance of the noise is not zero. Therefore, we obtain

$$\mathbf{f} = \begin{bmatrix} u_x \frac{\langle g_{x'}^2 \rangle}{\langle g_{x'}^2 \rangle + \sigma_n^2} \\ u_y \frac{\langle g_{y'}^2 \rangle}{\langle g_{y'}^2 \rangle + \sigma_n^2} \end{bmatrix}. \quad (20)$$

This equation shows that in noisy images two unfortunate things happen. First, we *always* obtain a solution and can thus no longer distinguish constant neighborhoods from edges and corners. Second, the obtained estimate is erroneous. It is biased towards lower values. This is also verified by experiments as shown in Fig. 3a and b. If the variance of the noise is about the squared magnitude of the gradient, the estimated values are only about half of the true values. Thus the differential method is an example of a non-robust technique since it deteriorates in noisy image sequences.

4.2 Tensor Method

The key to an easy solution is the transformation of the structure tensor into the local principal-axes coordinate system. A spatiotemporal image consisting only of zero-mean normal-distributed and isotropic noise with the variance σ_n results in a diagonal inertia tensor in any coordinate system, $\sigma_n^2 \mathbf{I}$, where $\mathbf{I}$ is the unit diagonal tensor. The variance σ_n^2 of the noise depends on the degree of smoothing applied when calculating the elements of the structure tensor according to (1). With white noise, σ_n^2 is inversely proportional to the volume over which the elements are averaged, i.e., σ^3 of the smoothing filter.

The structure tensor of a noisy spatiotemporal image is therefore given by

$$\mathbf{J}' = \mathbf{J} + \sigma_n^2 \mathbf{I} = \begin{bmatrix} J_{xx} + \sigma_n^2 & J_{xy} & J_{xt} \\ J_{xy} & J_{yy} + \sigma_n^2 & J_{yt} \\ J_{xt} & J_{yt} & J_{tt} + \sigma_n^2 \end{bmatrix}.$$

Figure 3: Systematic error in the velocity estimate with a moving random pattern using the least squares gradient method (**a** and **b**) and the tensor method (**c** and **d**) for velocities in x direction from 0 to 1.4 pixels/frame. The signal-to-noise ratio is 10 (**a** and **c**) and 1, (**b** and **d**) respectively.

From this relation, we can immediately conclude that the eigenvectors for an ideal oriented structure with isotropic noise are the same as for the noise-free case. Therefore, an unbiased estimate of orientation is possible. This is no longer the case, however, if the noise is *directional*.

If $\{\lambda_1, \lambda_2, \lambda_3\}$ is the set of eigenvalues of J , the eigenvalues of J' are

$$\{\lambda_1 + \sigma_n^2, \lambda_2 + \sigma_n^2, \lambda_3 + \sigma_n^2\}. \quad (21)$$

This important relation allows us to check the influence of noise on the various coherency measures. The spatial coherency measure c_s [Jähne, 1997b] becomes

$$c_s = \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \right)^2 = \frac{(J_{yy} - J_{xx})^2 + 4J_{xy}^2}{(J_{xx} + J_{yy} + 2\sigma_n^2)^2}. \quad (22)$$

Note that the variance of the noise is contained only in the divisor. Consequently, noise reduces the coherency measure. If the noise and signal levels are about the same, the maximal coherency is reduced to about half the value of the noise-free

case. The effect is similar for the total coherency [Jähne, 1997b]:

$$c_t = \left(\frac{\lambda_1 - \lambda_3}{\lambda_1 + \lambda_3} \right)^2 = \left(\frac{\lambda_1 - \lambda_3}{\lambda_1 + \lambda_3 + 2\sigma_n^2} \right)^2. \quad (23)$$

Because the noise terms cancel in the difference of the eigenvalues, the noise term is again contained only in the divisor, reducing the coherency.

5 Unsteady Motion

In this section we check the sensitivity of the optical flow estimation to changes in the motion field. By way of example, the temporal changes are discussed. Since motion results in oriented structures in space-time images, changes in the motion lead to curved structures. In order to model an unsteady motion, we introduce an *acceleration*

$$\mathbf{a} = \frac{\partial \mathbf{u}}{\partial t}. \quad (24)$$

For an arbitrary gray scale structure we can then write

$$g'(\mathbf{x}, t) = g(\mathbf{x} - \mathbf{u}t - 1/2\mathbf{a}t^2). \quad (25)$$

It is obvious that this modified model will influence only the temporal derivatives. The temporal derivative now becomes

$$\begin{aligned} \partial_t g' &= -(u_x + a_x t)g_x - (u_y + a_y t)g_y \\ &= -(\mathbf{u} + \mathbf{a}t)\nabla g, \end{aligned} \quad (26)$$

where we use $\mathbf{a} = [a_x, a_y]^T$.

In all methods, weighted products of the derivatives in the form of (1) appear. For example,

$$\begin{aligned} \langle \partial_x g \partial_t g \rangle &= -\langle (\mathbf{u} + \mathbf{a}t)\nabla g \partial_x g \rangle \\ &= -\mathbf{u} \langle \nabla g \partial_x g \rangle - \mathbf{a} \langle t \nabla g \partial_x g \rangle \end{aligned} \quad (27)$$

and

$$\begin{aligned} \langle (\partial_t g)^2 \rangle &= \langle [-(\mathbf{u} + \mathbf{a}t)\nabla g]^2 \rangle \\ &= u_x^2 \langle g_x^2 \rangle + 2u_x u_y \langle g_x g_y \rangle + u_y^2 \langle g_y^2 \rangle \\ &\quad + 2u_x a_x \langle t g_x^2 \rangle + 2(u_x a_y + u_y a_x) \langle t g_x g_y \rangle \\ &\quad + 2u_y a_y \langle t g_y^2 \rangle \\ &\quad + a_x^2 \langle t^2 g_x^2 \rangle + 2a_x a_y \langle t^2 g_x g_y \rangle + a_y^2 \langle t^2 g_y^2 \rangle. \end{aligned} \quad (28)$$

Although these terms look discouragingly complex, they are easy to interpret. We can distinguish terms with linear and quadratic appearances of time t . First, we have terms in which the time appears linearly in the averaging together with the corresponding derivatives. In the ideal case, when the spatial derivatives are evenly distributed over the averaging area, they all cancel out. The expression $\langle t \rangle$ is zero since the center of the coordinate system has been placed at the center pixel in (26). Now consider the converse case that the whole neighborhood over which we average contains only a singular steep spatial derivative at time t_0 . Then, for instance, (27) reduces to

$$\langle g_x g_t \rangle = -(\mathbf{u} + \mathbf{a} t_0) \langle \nabla g \partial_x g \rangle. \quad (29)$$

This equation means that the velocity obtained is not the velocity at time 0 — as expected — but the velocity at time t_0 . Thus the worst that could happen is that the optical flow estimate is biased to a value which is at the edge of the averaging neighborhood. If the spatial gray values are more equally distributed, the shift is of course less.

The discussion so far is valid for *all* methods, since the effect of the accelerated motion is simply that the correct velocity $\mathbf{u}$ has to be replaced by a biased velocity $\mathbf{u} + \mathbf{a} t_0$. The upper limit for t_0 is the temporal extension of the averaging mask. It is obvious that the same train of thought is valid for spatial variations of the velocity field. The velocity in the neighborhood is biased at most to a velocity value at the edge of the averaging neighborhood. Generally, we can conclude that the estimate of the optical flow is pulled to the values of regions with the steepest gradients.

One term remains to be discussed which appears only in the tensor method. It is $J_{tt} = \langle (\partial_t g)^2 \rangle$ (28). If we neglect the terms linear in t , terms that are quadratic in t still remain and (28) reduces to

$$\begin{aligned} \langle (\partial_t g)^2 \rangle &= u_x^2 \langle (\partial_x g)^2 \rangle + 2u_x u_y \langle \partial_x g \partial_y g \rangle + u_y^2 \langle (\partial_y g)^2 \rangle \\ &+ a_x^2 \langle (t \partial_x g)^2 \rangle + 2a_x a_y \langle t^2 \partial_x g \partial_y g \rangle + a_y^2 \langle (t \partial_y g)^2 \rangle. \end{aligned}$$

The terms quadratic in t appear as soon as the motion is accelerated. They contain a product of the squared acceleration and a measure of the temporal variance weighted with the steepness of the gray value changes. Since the mathematics becomes quite complex for the 3-D tensor, we illustrate this effect of an accelerated motion with the 2-D case. Then

$$\langle (\partial_t g)^2 \rangle = u_x^2 \langle (\partial_x g)^2 \rangle + a_x^2 \langle (t \partial_x g)^2 \rangle \quad (30)$$

and the 2-D structure tensor is

$$\begin{bmatrix} \langle (\partial_x g)^2 \rangle & -u_x \langle (\partial_x g)^2 \rangle \\ -u_x \langle (\partial_x g)^2 \rangle & u_x^2 \langle (\partial_x g)^2 \rangle + a_x^2 \langle (t \partial_x g)^2 \rangle. \end{bmatrix} \quad (31)$$

In the limit $a_x^2 \langle (t \partial_x g)^2 \rangle \ll (1 + u_x^2) \langle (\partial_x g)^2 \rangle$, the eigenvalues are

$$\left\{ (1 + u_x^2) \langle (\partial_x g)^2 \rangle, a_x^2 \langle (t \partial_x g)^2 \rangle \right\}. \quad (32)$$

In this approximation the eigenvector $\mathbf{e}_2$ of the smallest eigenvalue is no longer simply $[u_x, 1]^T$ but

$$\mathbf{e}_2 = \left[u_x \left(1 + \frac{a_x^2 \langle (t \partial_x g)^2 \rangle}{(1 + u_x^2) \langle (\partial_x g)^2 \rangle} \right), 1 \right]^T. \quad (33)$$

Thus there is a small bias in the estimate of the optical flow which goes with the square of the acceleration. As an example we take

$$u_x = 1, \quad a_x = 0.1, \quad \text{and} \quad \langle (t \partial_x g)^2 \rangle / \langle (\partial_x g)^2 \rangle = 4.$$

This means that the motion is changing by $\pm 20\%$ within the temporal width of the averaging window with a radius of 2. Even with this large relative velocity variation, the relative bias in the estimate of the optical flow is only 2%.

Likewise the decrease in the coherency measure is only marginal:

$$c_s = \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \right)^2 \approx 1 - \frac{4a_x^2 \langle (t \partial_x g)^2 \rangle}{(1 + u_x^2) \langle (\partial_x g)^2 \rangle}. \quad (34)$$

For the above example the coherency decrease is just 8%.

We can draw two important conclusions. First, the coherency is not sensitive to *gradual* changes in the velocity. There are two other sources that decrease the coherency. We have already seen in Sect. 4 that the coherency also decreases with increasing noise level. This decrease is generally homogeneous in space and time. We will see in the next section that the coherency decreases significantly also at motion discontinuities in space and time.

Secondly, we can average the estimates in time even with accelerated motion. Thus much more reliable motion estimates can be determined than from only two consecutive images.

6 Motion Discontinuity

Next, we turn to the analysis of motion discontinuities. We discuss only the case of a one-dimensional flow estimate for the tensor and phase methods. The neighborhood is partitioned into two subareas with different velocities. Without loss of generality, we choose

$$g'(x, t) = g(x + ut)\Pi(-x) \quad \text{and} \quad g''(x, t) = g(x - ut)\Pi(x),$$

where $\Pi(x)$ is the step function (we can recover all other cases by rotation of the coordinate system).

6.1 Tensor Method

The tensor method gives the following estimate for the optical flow:

$$f \approx u \frac{\langle (\partial_x g')^2 \rangle - \langle (\partial_x g'')^2 \rangle}{\langle (\partial_x g')^2 \rangle + \langle (\partial_x g'')^2 \rangle}. \quad (35)$$

If the mean square gradient is equal in both regions, the estimated optical flow is zero, as expected. The coherency is

$$c^2 = 1 - 2\gamma \sin^2 \alpha/2, \quad (36)$$

where α is the angle between the spatiotemporal orientation in the two regions (α is related by $\tan(\alpha/2) = u$ to the velocity in the two regions) and γ is a measure that compares the mean square gradient in the two regions:

$$\gamma = 1 - \left(\frac{\langle (\partial_x g')^2 \rangle - \langle (\partial_x g'')^2 \rangle}{\langle (\partial_x g')^2 \rangle + \langle (\partial_x g'')^2 \rangle} \right). \quad (37)$$

The factor γ is one if both regions have the same mean square gradient and zero if the mean square gradient in one region is significantly larger than in the other. Let us first assume that the mean square gradients are equal, i.e., $\gamma = 1$. Then the expression for the coherency becomes even simpler: $c = \cos \alpha$. This tells us that the coherency is zero if the orientations in the two regions are orthogonal. If $\gamma < 1$, the coherency reaches a minimal value of $1 - \gamma$ for regions with orthogonal orientation. In this way, the coherency is a direct measure for the change of the orientation (and thus velocity). Because of the squared sine function in (36), it is quite insensitive for small differences in the orientation angle.

6.2 Phase Method

This paragraph on the response of the phase method to motion discontinuities will illuminate a general problem of how to weight estimates of the optical flow properly. We use the same approach as for the tensor method and obtain

$$f = \frac{u_x \langle g'_+ \partial_x g'_- - g'_- \partial_x g'_+ \rangle - u_x \langle g''_+ \partial_x g''_- - g''_- \partial_x g''_+ \rangle}{\langle g'_+ \partial_x g'_- - g'_- \partial_x g'_+ \rangle + \langle g''_+ \partial_x g''_- - g''_- \partial_x g''_+ \rangle}.$$

This result is very similar to the tensor method; only the weighting factors are different. With the phase method, we can avoid weighting with the magnitude or squared magnitude of the gradients. Then we must set some kind of empirical threshold, however, because the phase information obtained from low gray value gradients will not be reliable. More generally, we lose a confidence measure for the phase estimate, since we treat all estimates as if they had the same error.

In any case, a motion discontinuity at a straight line between two regions with random variations in the mean square gradient will not appear as a straight line in the optical flow image. It will rather become a wiggly line because of the bias caused by weighting. The deviation from the true discontinuity can be as much as half the mask size of the averaging filters. We seem to have reached a point in the analysis of motion that cannot be solved on the base of local information alone. At this point, global information must be considered, such as the smoothness of boundaries.

7 Illumination Change

Finally, we discuss the influence of illumination changes using $g'(x, t) = g(x - ut, t)$. Then $\partial_x g = \partial_x g$ and $\partial_t g' = -u \partial_x g + \partial_t g$, where $\partial_t g$ is the explicit temporal derivative.

7.1 Tensor Method

The estimate of the optical flow for the tensor method gives

$$f = \frac{u - \langle \partial_x g \partial_t g \rangle / \langle (\partial_x g)^2 \rangle}{1 - u^2 + 2u \langle \partial_x g \partial_t g \rangle / \langle (\partial_x g)^2 \rangle - \langle (\partial_t g)^2 \rangle / \langle (\partial_x g)^2 \rangle}.$$

Even if $\langle \partial_x g \partial_t g \rangle = 0$, the velocity estimate is biased towards higher velocities, as can be seen for $u \ll 1$:

$$f \approx \frac{u}{1 - \langle (\partial_t g)^2 \rangle / \langle (\partial_x g)^2 \rangle}. \quad (38)$$

This result is not surprising, since illumination changes occur as additional patterns with an orientation corresponding to infinite velocity.

7.2 Phase Method

From a phase-based method we would expect that global illumination changes are entirely suppressed since only the phase and not the amplitude of the signal is used for the determination of displacements.

After some algebra we obtain

$$f = u - \frac{\langle g_+ \partial_t g_- - g_- \partial_t g_+ \rangle}{\langle g_+ \partial_x g_- - g_- \partial_x g_+ \rangle}, \quad (39)$$

where $\partial_t g$ again means the *explicit* temporal derivative. The additional term vanishes when the explicit temporal derivative is directly proportional to the gray value. This is indeed the case for global illumination changes; they are directly proportional to the gray value. Thus the phase method suppresses global illumination changes under very general assumptions.

8 Conclusions

In this paper analytical techniques have been discussed to study the performance characteristics of low-level motion estimators in continuous space-time images. This technique proved to be powerful because it allows to separate principal characteristics of a technique from errors introduced by the numerical implementation and because it gives conclusive results under very general conditions.

Still, a lot remains to be done. Interesting topics of further research are the study of statistical error propagation, a comparative study of gradient-based techniques with quadrature-filter set techniques and correlation-based techniques, and the inclusion of regularization techniques.

Acknowledgements

Financial support of this research by the Deutsche Forschungsgemeinschaft through the interdisciplinary research unit (Forschergruppe) “Image Sequence Processing to Investigate Dynamic Processes” (Jae 395/6-1) is gratefully acknowledged.

References

- [1] Barron, J. L., D. J. Fleet, and S. Beauchemin. Performance of optical flow techniques. *Int. J. Computer Vision*, 12 (1):43–77, 1994.
- [2] Bigün, J., and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proceedings ICCV’87, London 1987*, pages 433–438, IEEE Computer Society Press, Los Alamitos, 1987.
- [3] Fleet, D. J., and A. D. Jepson. Computation of component image velocity from local phase information. *Int. J. Comp. Vision*, 5:77–104, 1990.
- [4] Haglund, L., and D. J. Fleet. Stable estimation of image orientation. In *Proceedings ICIP’94, Austin 1994, Vol. III*, pages 68–72, IEEE Computer Society Press, Los Alamitos, 1994.
- [5] Haußecker, H., and B. Jähne. A tensor approach for precise computation of dense displacement vector fields, *Proc. Mustererkennung 1997*, Braunschweig, 15–17. September 1997, F. Wahl und E. Paulus (Eds.), Informatik Aktuell, Springer, Berlin, in press, 1997.
- [6] Jähne, B. Motion determination in space-time images, *Proc. Computer Vision — ECCV’90*, O. Faugeras, ed., Lecture Notes Computer Science 427, pp. 161–173, 1990.
- [7] Jähne, B. *Spatio-Temporal Image Processing, Theory and Scientific Applications*, Lecture Notes in Computer Science, Vol. 751, Springer, Berlin, 1993.
- [8] B. Jähne. *Handbook of digital image processing for scientific applications*. CRC Press, Boca Raton, Florida, 1997a.
- [9] Jähne, B. *Digital Image Processing — Concepts, Algorithms, and Scientific Applications*, 4th completely revised edition, Springer, Berlin, 1997b.
- [10] Jähne, B., H. Haußecker, F. Hering, G. Balschbach, J. Klinke, M. Lell, D. Schmundt, M. Schultz, U. Schurr, M. Stitt, and U. Platt. The role of active vision in exploring growth, transport, and exchange processes. *Proc. Aktives Sehen in technischen und biologischen Systemen*, Workshop der GI-Fachgruppe 1.0.4. Bildverstehen Hamburg, 3–4. December 1996, B Mertsching (Hrsg), *Proc. in Artificial Intelligence 4*, infix, Sankt Augustin, pp. 194–202, 1996.
- [11] Karlholm, J. *Efficient Spatiotemporal Filtering and Modelling*. Linköping Studies in Science and Technology, Thesis No. 562, Linköping University, Sweden, 1996.
- [12] Kearney, J. K., W. B. Thompson, and D. L. Boley. Optical flow estimation: an error analysis of gradient-based methods with local optimization. *IEEE Trans. PAMI*, 9 (2):229–244, 1987.

- [13] Scharr, H., S. Körkel, and B. Jähne. Numerische Isotropieoptimierung von FIR-Filtern mittels Querglättung, Proc. *Mustererkennung 1997*, Braunschweig, 15–17. September 1997, F. Wahl und E. Paulus (Eds.), Informatik Aktuell, Springer, Berlin, in press, 1997.
- [14] Singh, A. *Optic flow computation: a unified perspective*. IEEE Computer Society Press, Los Alamitos, CA, 1991.