On the Estimation of Optical Flow: Relations between Different Approaches and Some New Results

Hans-Hellmut Nagel
Fraunhofer-Institut für Informations- und Datenverarbeitung, 7500 Karlsruhe 1, Fed. Rep. Germany; and Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Fed. Rep. Germany

Recommended by M. Brady and A. Pentland

ABSTRACT

A common basis is suggested for the optical flow estimation approaches of Nagel (1983), Haralick and Lee (1983) and Tretiak and Pastor (1984). Based on a discussion of these approaches, an exact solution for the system of partial differential equations proposed by Horn and Schunck (1981) is given at gray value corners and extrema. The insight gained by this solution results in a modification of the "oriented smoothness" approach of Nagel (1983) which thereby becomes considerably simpler. In addition, the optical flow estimation approach of Hildreth (1983, 1984) can be shown to represent a kind of special case of this modified "oriented smoothness" approach in a more direct manner than discussed in Nagel (1984).

1. Introduction

It is well known that the temporal changes of a gray value structure recorded by an imaging sensor contain information which may allow to infer—in combination with certain assumptions—the relative three-dimensional motion between the sensor and its environment as well as the spatial structure of this environment. Many animals as well as man exploit this information in order to move around, to detect prey and to monitor their environment regarding threat from predators or moving objects. Koenderink and van Doorn [14] present an especially illustrative exposition of some ideas how the relevant information could be extracted. The corresponding ability of "biological information processing systems" attracted quite naturally the attention of the vision community (see, for example, the work by Ullman [27], Horn and Schunck [11],

Artificial Intelligence **33** (1987) 299–324

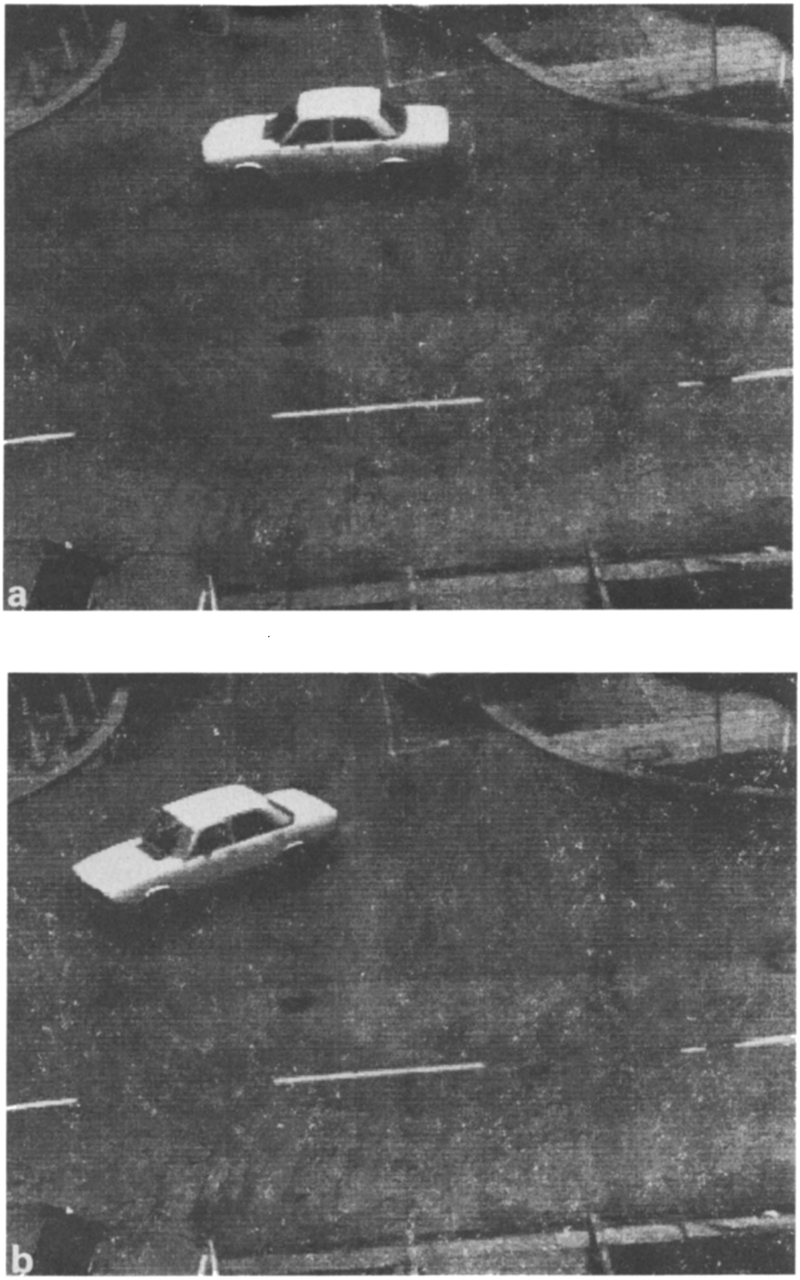


FIG. 1. The first (a) and seventh (b) frame of a short image sequence recorded by a stationary TV-camera from a street scene (by permission from Enkelmann [5]).

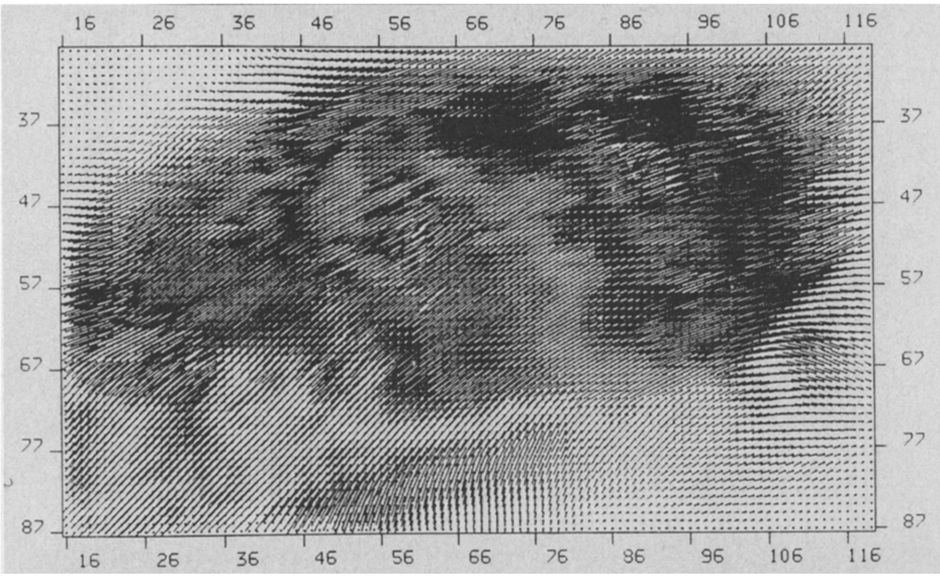


FIG. 2. The optical flow field for the shift of gray value structures between the sixth and the seventh frame for a window around the moving car in the sequence shown in Fig. 1, estimated according to the method described by Enkelmann [5].

and Hildreth [9, 10]). A more recent survey of technical aspects can be found in Nagel [18, 19].

Relative motion between an imaging sensor and its environment will generally result in shifts of gray value structures in the image plane of the sensor. One may conceptually divide the task of extracting the desired information into two steps:

(1) to estimate the shift of gray value structures (also denoted as optical flow), and

(2) to interpret the estimated optical flow field.

In order to illustrate the concept of optical flow, Figs. 1(a) and (b) present the first and last frame of a short image sequence showing a moving car in a street scene taken from [5] (see also [7]). For each pixel of a window around the image of the moving car, Fig. 2 shows an estimate of the shift of the gray value structure between the preceding frame and the one shown in Fig. 1(b). Selected optical flow vectors for prominent gray value structures are superimposed on a reproduction of the image window from the next to last frame of this sequence in Fig. 3(a) and likewise from the last frame in Fig. 3(b). Optical flow vectors such as those shown in Fig. 3 have been used by Westphal and Nagel [28] to construct a 3-D description for a moving object (see also [1, 2, 4]).

The remainder of this contribution will concentrate on the estimation of an

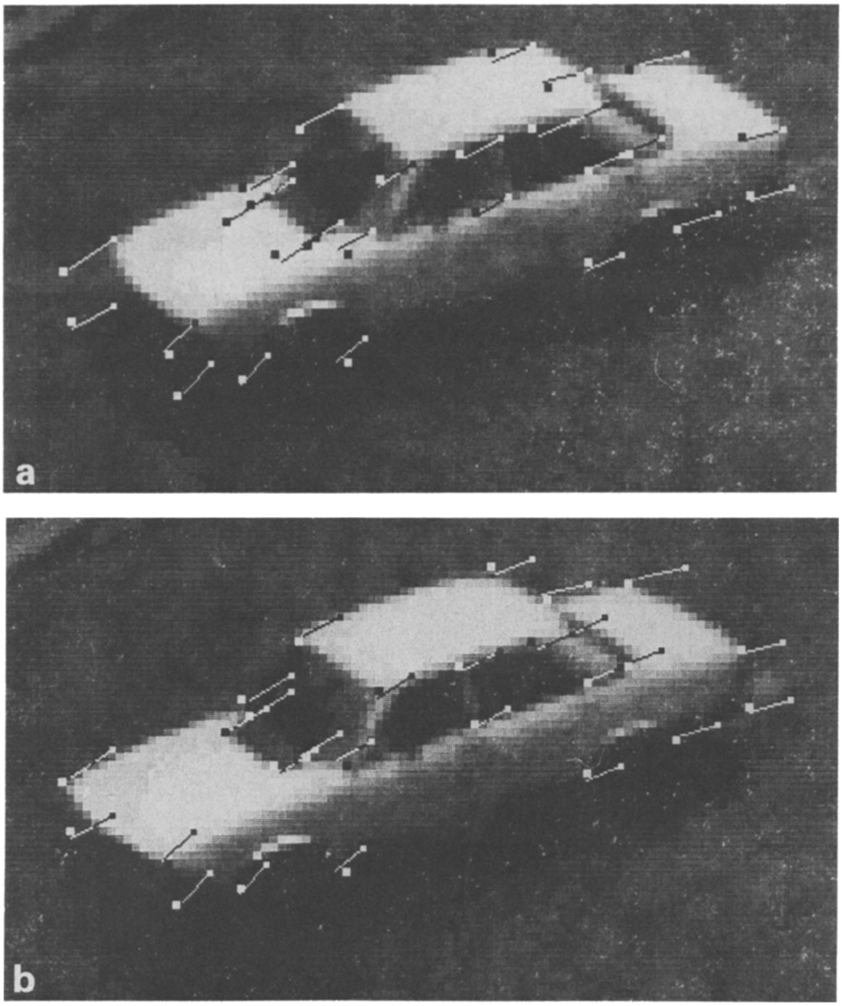


FIG. 3. Selected optical flow vectors, superimposed on the window from the sixth frame (a) and from the seventh frame (b).

optical flow field. Section 2 recapitulates an equation which is employed during the estimation of optical flow vector fields from image sequences. This equation does not provide enough constraints, however, to fully determine the optical flow. Additional assumptions have to be introduced, therefore, which provide further constraints on the optical flow field. Alternatively, one might attempt to exploit more fully the details of a gray value structure in order to estimate both components of its shift between consecutive image frames. A combination of both these approaches appears to be necessary in general. This leads to two questions:

(1) Which information in a local gray value structure can be exploited for the estimation of both components of an optical flow vector?

(2) What are the minimal assumptions required for this goal?

Section 3 recapitulates the concept of a gray value corner and how this idea about a specific gray value structure can be combined with the assumption that the optical flow vector is constant within a small image region in order to estimate both of its components. Sections 4 and 5 describe two approaches which start from apparently different premises but can be shown to result in expressions very similar to those derived in Section 2. A specific way to relax the assumption of constant optical flow within the image window exploited for the estimation is then discussed in Section 6, namely assuming a linear variation of the optical flow field as a function of the image plane coordinates. This is not quite sufficient to fully determine the optical flow, although the resulting expressions are of the same form as those discussed in Section 5. They need, however, to be supplemented with another, weaker, assumption in order to recover both components of the optical flow. The technique developed in Section 6 is used in Section 7 for an investigation of how the smoothness assumption of Horn and Schunck [11] can be combined with assumptions about the local gray value structure in order to determine both components of the optical flow field. The resulting insights lead to a significant simplification for the formulation of an "oriented smoothness requirement" suggested earlier by Nagel [16] and investigated by Nagel and Enkelmann [21–23] as well as by Enkelmann [5–7]. The exposition of these ideas in Section 8 is followed by a discussion regarding how the approach of Hildreth [9, 10] can be understood as a special case of this simplified formulation of an "oriented smoothness requirement."

2. The "Motion Constraint Equation"

In order to simplify the subsequent exposition, the following discussion will be restricted to the situation where a single imaging sensor moves within an otherwise stationary environment.

Let us assume that a surface element in the environment is imaged at time t onto the image plane element $\mathrm{d}x\,\mathrm{d}y$ at location $\boldsymbol{x} = (x, y)^{\mathrm{T}}$ with gray value $g(\boldsymbol{x}, t)$. Let the instantaneous motion vector of the sensor relative to its three-dimensional (3-D) environment be essentially parallel to the image plane. Let us assume further that the 3-D surface element imaged at $\boldsymbol{x}$ is more or less parallel to the image plane and that its illumination does not vary substantially from time t to time $(t + \mathrm{d}t)$. Under these assumptions, the gray value $g(\boldsymbol{x} + \mathrm{d}\boldsymbol{x}, t + \mathrm{d}t)$ for the image of this surface element at time $(t + \mathrm{d}t)$ remains essentially constant during the time $\mathrm{d}t$, i.e.

$$dg(\boldsymbol{x}, t) = 0 = \frac{\partial g(\boldsymbol{x}, t)}{\partial x}\, dx + \frac{\partial g(\boldsymbol{x}, t)}{\partial y}\, dy + \frac{\partial g(\boldsymbol{x}, t)}{\partial t}\, dt\,. \tag{1}$$

Schunck and Horn [25] as well as Schunck [24] have shown that equation (1) can be considered valid even in the vicinity of strong gray value variations where a first-order Taylor approximation does not appear to be justified a priori. Schunck [24] discusses, too, a generalization of (1) for the case where some of the above-mentioned assumptions are not fully applicable. Such aspects, however, will not be pursued further in this contribution.

Denoting partial derivatives by subscripts, equation (1) may be written in the form

$$g_x\, dx + g_y\, dy + g_t\, dt = 0\,, \tag{1a}$$

or

$$g_x \frac{dx}{dt} + g_y \frac{dy}{dt} + g_t = 0\,. \tag{1b}$$

The vector $d\boldsymbol{x}/dt = (dx/dt, dy/dt)^{\mathrm{T}}$ describes the instantaneous displacement velocity for the image of the surface element which is depicted at time t at the location $\boldsymbol{x}$ in the image plane. Provided the assumptions mentioned above are justified, this image plane displacement velocity is related to spatio-temporal changes of the recorded gray value structure according to (1b)

The way in which (1b) has been introduced implies that the image position of the surface element in question is known as a function of time. This, however, will not be true in general. The only available knowledge consists in the observed spatio-temporal gray value structure $g(\boldsymbol{x}, t)$. In order to distinguish between the generally unknown displacement of depicted surface elements in the image plane and the observable positional shifts of prominent gray value structures, the term "optical flow" is used for the latter one. The two-dimensional time-dependent vector field

$$\boldsymbol{u}(\boldsymbol{x}, t) = \begin{pmatrix} u(\boldsymbol{x}, t) \\ v(\boldsymbol{x}, t) \end{pmatrix} \tag{2}$$

is introduced as an instantaneous mapping of the gray value structure $g(\boldsymbol{x}, t)$ observed at time t onto another one observed at time $(t + dt)$:

$$g(\boldsymbol{x}, t + dt) = g(\boldsymbol{x} - \boldsymbol{u}(\boldsymbol{x}, t)\, dt, t)\,. \tag{3}$$

Development of (3) into a Taylor series up to first order yields

$$g_t\, dt = -\,(\nabla g)^{\mathrm{T}}\, \boldsymbol{u}\, dt\,, \tag{4a}$$

or

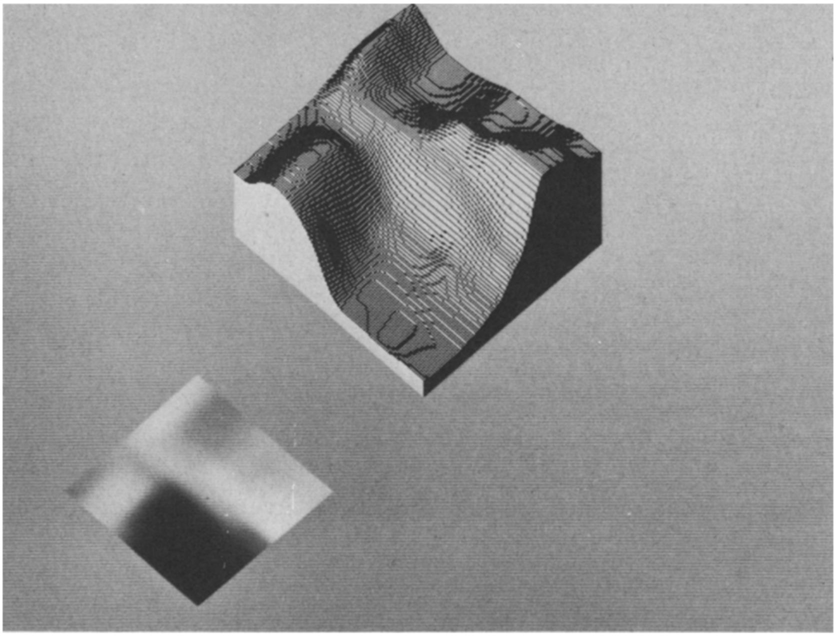
$$(\nabla g)^{\mathrm{T}} \boldsymbol{u} + g_t = 0\,, \tag{4b}$$

which formally corresponds to (1b). Equation (3) alone, however, does not provide enough constraints in general to determine $\boldsymbol{u}(\boldsymbol{x}, t)$ uniquely for all image plane locations $\boldsymbol{x}$. Equation (4b) contains the unknown vector function $\boldsymbol{u}(\boldsymbol{x}, t)$ explicitly, but represents only a single relation for the two unknown components $u(\boldsymbol{x}, t)$ and $v(\boldsymbol{x}, t)$ of $\boldsymbol{u}(\boldsymbol{x}, t)$ at each image plane location $\boldsymbol{x}$. Additional assumptions about $\boldsymbol{u}(\boldsymbol{x}, t)$ are required in order to facilitate its estimation.

3. Estimation of Optical Flow at Gray Value Corners

Based on earlier investigations [1–4], Nagel [15] noticed the relation between a heuristic definition of a certain type of a prominent gray value structure [1–4] and a similar one first discussed by Kitchen and Rosenfeld [12, 13]. This led to the characterization of a "gray value corner" as the location of maximum planar curvature in the locus line of steepest gray value slope (see Figs. 4–6). This characterization can be formulated quantitatively if we assume that the local coordinate system is aligned with the directions of principal curvature of the gray value structure $g(\boldsymbol{x}, t)$ from which it follows that $g_{xy} = 0$:

FIG. 4. Enlarged section of a digitized image from an image sequence recording a parking lot scene. The two image subsections represented in Figs. 5 and 6 have been marked.



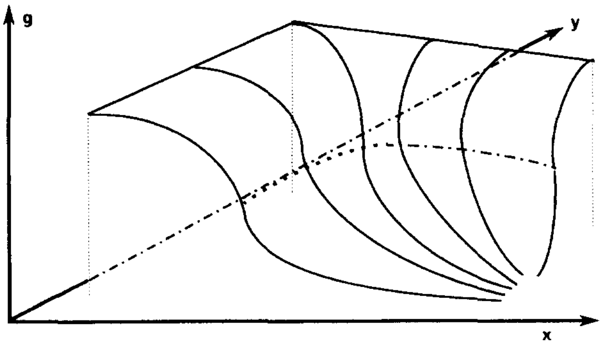


FIG. 5. (a) The gray values from the window corner on the bright car in Fig. 4 shown as a function of the image plane coordinates. Dark gray areas correspond to low values, bright ones to high values. The subsection taken from Fig. 4 has been rotated by 45 degrees. The same data are shown as a gray value representation. (b) Idealized sketch of $g(x, y)$ for a "gray value corner" such as the one depiced in Fig. 5(a).

$$g_x(\boldsymbol{x}, t) = \text{extremum} \neq 0 \,, \tag{5a}$$

$$g_{xx}(\boldsymbol{x}, t) = 0 \,, \tag{5b}$$

$$g_y(\boldsymbol{x}, t) = 0 \,, \tag{5c}$$

$$g_{yy}(\boldsymbol{x}, t) = \text{extremum} \neq 0 \,. \tag{5d}$$

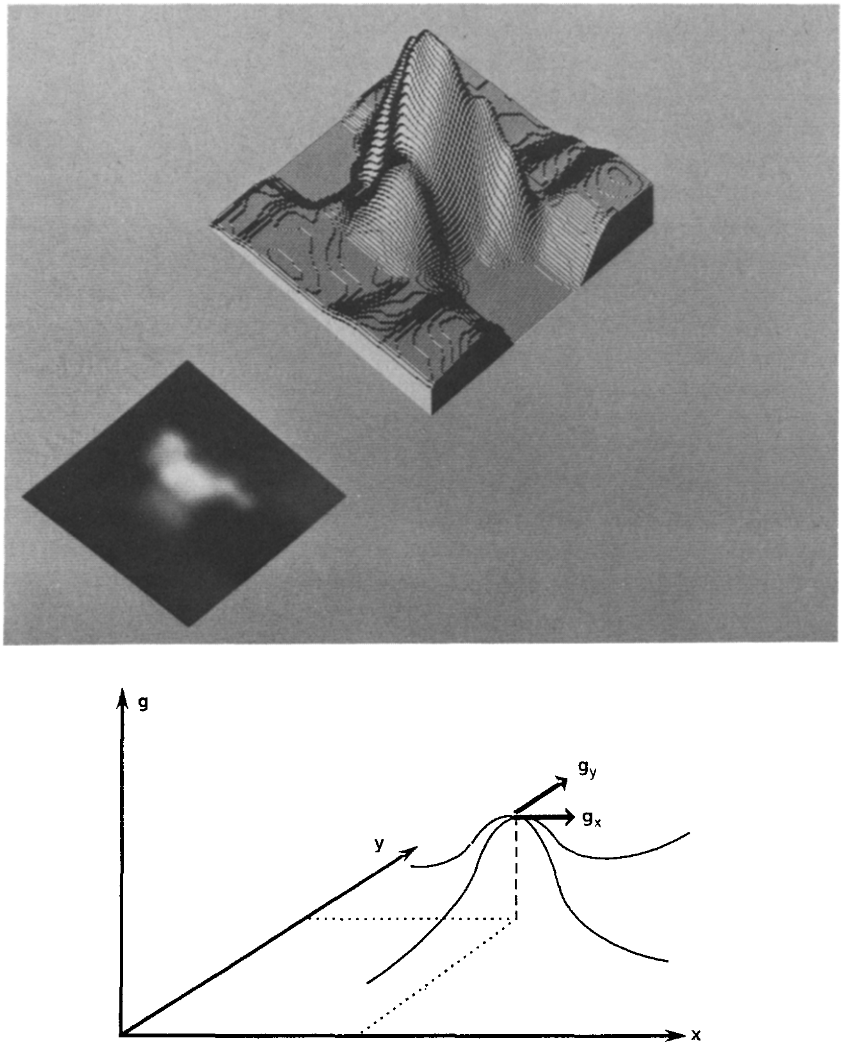


FIG. 6. (a) The gray values from the lamp post in Fig. 4 shown in analogy to Fig. 5(a) as an example for a gray value extremum. (b) Idealized sketch of $g(x, y)$ for a "gray value extremum" such as the one depicted in Fig. 6(a).

Equation (5a) expresses the requirement that the gray value slope in x-direction is extremal which implies that the second partial derivative of $g(\boldsymbol{x}, t)$ with respect to x must cross zero at this location (see (5b)). If the gradient is oriented along the x-axis, the first partial derivative of $g(\boldsymbol{x}, t)$ with respect to y must be zero (see (5c)), i.e. the locus line of maximal slope is locally an iso-intensity curve. The curvature of this planar curve is proportional to the second partial derivative of $g(\boldsymbol{x}, t)$ with respect to y. This curvature should be an extremum according to equation (5d).

Whereas (4b) does not provide enough constraints to estimate both components of the optical flow vector $\boldsymbol{u}$, Nagel [15] realized that this becomes possible at gray value corners. His approach is based on the assumption that a local combination of gray values is displaced as a rigid structure from time t_0 to time $t_1 = (t_0 + \mathrm{d}t)$ by a displacement $\boldsymbol{u}$. This displacement, which is assumed to be constant for the entire gray value structure, has to be estimated by minimization of the squared differences between the gray value structures observed at times t_0 and t_1:

$$\int\int \mathrm{d}x\,\mathrm{d}y[g(\boldsymbol{x}, t_1) - g(\boldsymbol{x} - \boldsymbol{u}\,\mathrm{d}t, t_0)]^2 = \text{minimum}\,. \tag{6}$$

The dependence of $g(\boldsymbol{x} - \boldsymbol{u}\,\mathrm{d}t, t_0)$ on $\boldsymbol{u}$ is made explicit by developing $g(\boldsymbol{x} - \boldsymbol{u}\,\mathrm{d}t, t_0)$ into a Taylor series up to second-order terms. Equating the partial derivatives of the integral in (6) with respect to the unknown components u and v to zero results in a system of two nonlinear equations in u and v. Nagel [15] could show that this system of nonlinear equations can be solved in closed form at gray value corners with the result, writing $g_0(\boldsymbol{x})$ for $g(\boldsymbol{x}, t_0)$ and $g_1(\boldsymbol{x})$ for $g(\boldsymbol{x}, t_1)$,

$$u = -\frac{(\overline{g_1} - \overline{g_0} - \frac{1}{2} g_{0,yy} v^2)}{g_{0,x}}\,, \tag{7a}$$

$$v = -\frac{g_y(\boldsymbol{x}_0, t_1)}{g_{yy}(\boldsymbol{x}_0, t_0)} = -\frac{g_{1,y}}{g_{0,yy}}\,. \tag{7b}$$

If the time interval $\mathrm{d}t = (t_1 - t_0)$ between the two frame times is set to unity, one obtains the relation

$$\overline{g_1} - \overline{g_0} = g_t \Delta t \Rightarrow \overline{g_1} - \overline{g_0} = g_t\,, \tag{8}$$

where the overbar indicates averaging of the gray values within the image region chosen to comprise the gray value structure in question. In order to relate this result to earlier ones, let us assume that the optical flow points essentially along the x-axis, i.e. v is assumed to be negligible. Using (8), equation (7a) will then simplify to

$$u = -\frac{(\overline{g_1} - \overline{g_0} - \frac{1}{2} g_{0,yy} v^2)}{g_{0,x}} \simeq -\frac{\overline{g_1} - \overline{g_0}}{g_{0,x}} \simeq -\frac{g_t}{g_{0,x}} . \tag{9}$$

Under the same assumptions, equation (4b) will just yield the result given by the right-hand side of equation (9).

The result given by (7) for $\boldsymbol{u}$ at a gray value corner could be used as a starting value for an iterative solution approach (see also [20]) which extends the solution of the nonlinear systems of equations for u and v into image areas surrounding this gray value corner. Nagel [15] linearized these equations in order to estimate a correction vector

$$\boldsymbol{u}^{(k+1)} = \boldsymbol{u}^{(k)} + \mathrm{d}\boldsymbol{u} \tag{10}$$

with the result

$$\mathrm{d}\boldsymbol{u} = - C^{-1}\overline{(g_1(\boldsymbol{x}) - g_0(\boldsymbol{x} - \boldsymbol{u}^{(k)}\,\mathrm{d}t))\, \nabla g_0(\boldsymbol{x} - \boldsymbol{u}^{(k)}\,\mathrm{d}t)} \tag{11}$$

where the matrix C is given by

$$C = (\nabla g_0)(\nabla g_0)^{\mathrm{T}} + \overline{x^2}(\nabla\nabla g_0)(\nabla\nabla g_0)^{\mathrm{T}} , \tag{12a}$$

or

$$C = \begin{pmatrix} g_{0,x}^2 & g_{0,x} g_{0,y} \\ g_{0,x} g_{0,y} & g_{0,y}^2 \end{pmatrix} + \overline{x^2} \begin{pmatrix} g_{0,xx}^2 + g_{0,xy}^2 & g_{0,xy}(g_{0,xx} + g_{0,yy}) \\ g_{0,xy}(g_{0,xx} + g_{0,yy}) & (g_{0,xy}^2 + g_{0,yy}^2) \end{pmatrix} . \tag{12b}$$

As it has been shown by Nagel [16–18], the matrix C embodies important information about the gray value structure. It should be noted that with vanishing partial derivatives of second order, the solutions given by (7) as well as by (11) become undefined—a version of the matrix C containing only first-order partial derivatives is singular! At a gray value corner, the solution for the correction vector $\mathrm{d}\boldsymbol{u}$ according to equation (11) with starting vector $\boldsymbol{u}^{(k)} = 0$ will be equivalent to the solution given by (7) as it has been shown by Nagel [15].

4. The Approach of Haralick and Lee

Based on the facet model for low-level image processing, Haralick and Lee [8] extended the interpretation of the optical flow estimation approach outlined in connection with equation (4b). These authors interpret this equation as the intersection line of the iso-intensity contour plane at time t_0 with the gray value structure from a successive image frame at time t_1. In order to single out a match point along this intersection line, they require that

(i) the gray values match: this leads back to equation (4b);

(ii) the first derivatives of $g(\boldsymbol{x}_0 - \boldsymbol{u}\,\mathrm{d}t, t)$ match those of $g(\boldsymbol{x}_0, t + \mathrm{d}t)$;

(iii) the partial derivatives of third and higher order are negligible.

Requirement (ii) implies

$$g_x(\boldsymbol{x}_0, t + \mathrm{d}t) = g_x(\boldsymbol{x}_0 - \boldsymbol{u}\,\mathrm{d}t, t)\,, \tag{13a}$$

$$g_y(\boldsymbol{x}_0, t + \mathrm{d}t) = g_y(\boldsymbol{x}_0 - \boldsymbol{u}\,\mathrm{d}t, t)\,, \tag{13b}$$

$$g_t(\boldsymbol{x}_0, t + \mathrm{d}t) = g_t(\boldsymbol{x}_0 - \boldsymbol{u}\,\mathrm{d}t, t)\,. \tag{13c}$$

In order to obtain the dependencies on $\mathrm{d}t$ explicitly, the expressions for the normal vectors at locations $\boldsymbol{x}_0$ and $(\boldsymbol{x}_0 - \boldsymbol{u}\,\mathrm{d}t)$ are developed into first-order Taylor series in $\mathrm{d}t$ which yields

$$g_{xt}(\boldsymbol{x}_0, t)\,\mathrm{d}t = -\,\{g_{xx}(\boldsymbol{x}_0, t)\,u + g_{xy}(\boldsymbol{x}_0, t)\,v\}\,\mathrm{d}t\,, \tag{14a}$$

$$g_{yt}(\boldsymbol{x}_0, t)\,\mathrm{d}t = -\,\{g_{yx}(\boldsymbol{x}_0, t)\,u + g_{yy}(\boldsymbol{x}_0, t)\,v\}\,\mathrm{d}t\,, \tag{14b}$$

$$g_{tt}(\boldsymbol{x}_0, t)\,\mathrm{d}t = -\,\{g_{tx}(\boldsymbol{x}_0, t)\,u + g_{ty}(\boldsymbol{x}_0, t)\,v\}\,\mathrm{d}t\,, \tag{14c}$$

or

$$g_{xx}u + g_{xy}v + g_{xt} = 0\,, \tag{15a}$$

$$g_{yx}u + g_{yy}v + g_{yt} = 0\,, \tag{15b}$$

$$g_{tx}u + g_{ty}v + g_{tt} = 0\,. \tag{15c}$$

Equations (15), together with equation (4b), form an overdetermined system of four linear equations for the two unknown components u and v of $\boldsymbol{u}(\boldsymbol{x}_0, t)$. This system of equations can be solved for u and v based on the pseudo-inverse formalism as it will be discussed in the following section.

5. The Approach of Tretiak and Pastor

Apparently without knowledge about the work by Haralick and Lee [8], Tretiak and Pastor [26] differentiated equation (4b) with respect to x and y, by only regarding the gray value derivatives g_x and g_y. Tretiak and Pastor [26] did not give any argument for not differentiating the components of $\boldsymbol{u}$, i.e. they implicitly considered $\boldsymbol{u}$ to be locally constant. They thus obtained two new equations, namely,

$$g_{xx}u + g_{yx}v + g_{tx} = 0 , \tag{16a}$$

$$g_{xy}u + g_{yy}v + g_{ty} = 0 , \tag{16b}$$

in addition to equation (4b)

$$g_x u + g_y v + g_t = 0 .$$

It is immediately seen that (16a) and (16b) are equivalent to (15a) and (15b) as derived by Haralick and Lee. In addition to using just two of the equations (4b), (16a) and (16b) to determine u and v, Tretiak and Pastor, too, suggested to employ the pseudo-inverse formalism in order to solve the overdetermined set of equations (4b), (16a), and (16b):

$$\begin{pmatrix} g_x & g_y \\ g_{xx} & g_{yx} \\ g_{xy} & g_{yy} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{pmatrix} g_t \\ g_{tx} \\ g_{ty} \end{pmatrix} . \tag{17}$$

The pseudo-inverse solution for this system of equations yields

$$\begin{pmatrix} u \\ v \end{pmatrix} = - \left(\begin{pmatrix} g_x & g_{xx} & g_{xy} \\ g_y & g_{yx} & g_{yy} \end{pmatrix} \begin{pmatrix} g_x & g_y \\ g_{xx} & g_{yx} \\ g_{xy} & g_{yy} \end{pmatrix} \right)^{-1} \begin{pmatrix} g & g_{xx} & g_{xy} \\ g_y & g_{yx} & g_{yy} \end{pmatrix} \begin{pmatrix} g_t \\ g_{tx} \\ g_{ty} \end{pmatrix} . \tag{18}$$

The close similarity between (18) and the approach discussed in Section 3 will become even more obvious if the matrix which has to be inverted on the right-hand side of equation (18) is written explicitly:

$$\begin{aligned} &\left(\begin{pmatrix} g_x & g_{xx} & g_{xy} \\ g_y & g_{yx} & g_{yy} \end{pmatrix} \begin{pmatrix} g_x & g_y \\ g_{xx} & g_{yx} \\ g_{xy} & g_{yy} \end{pmatrix} \right)^{-1} \\ &= \begin{pmatrix} g_x^2 + g_{xx}^2 + g_{xy}^2 & g_x g_y + g_{xy}(g_{xx} + g_{yy}) \\ g_x g_y + g_{xy}(g_{xx} + g_{yy}) & g_y^2 + g_{xy}^2 + g_{yy}^2 \end{pmatrix}^{-1} . \end{aligned} \tag{19}$$

Apart from the average of x^2 which appears as a factor for those terms containing second partial derivatives of g, the matrix C defined in (12) corresponds to the matrix which has to be inverted on the right-hand side of (19). A specialization of (18) for the case of a gray value corner is discussed in

Appendix A in order to further illustrate the relation between the approach of Haralick and Lee [8] as well as that of Tretiak and Pastor [26] with the solution presented in Section 3.

6. A Common Basis for the Approaches Discussed in the Preceding Sections

Sections 3–5 reported different ways to formulate the basic requirement that the gray value structures observed at time t_0 around $(\boldsymbol{x}_0 - \boldsymbol{u}\,\mathrm{d}t)$ should match the one observed at time $t_1 = (t_0 + \mathrm{d}t)$ around $\boldsymbol{x}_0$. If the gray value structure around $\boldsymbol{x}_0$ is articulated well enough—i.e. if the gradient of $g(\boldsymbol{x})$ changes rapidly enough in the vicinity of $\boldsymbol{x}_0$, expressed by the fact that not all second partial derivatives of $g(\boldsymbol{x})$ with respect to x and y vanish at $\boldsymbol{x}_0$—then both components u and v of $\boldsymbol{u}$ can be estimated. All approaches required so far that the optical flow $\boldsymbol{u}$ is constant within the image region exploited for the estimation of $\boldsymbol{u}$.

It is suggested here to relax this requirement in favor of the assumption that the optical flow varies at most linearly with the image plane coordinates x, y within the image region considered. This assumption is equivalent to the neglect of higher than linear terms in the Taylor development of $\boldsymbol{u}(\boldsymbol{x})$ around the center $\boldsymbol{x}_0$ of the region. One may consider this assumption as a special kind of smoothness requirement. It will be supplemented by the explicit introduction of knowledge about the gray value variation around $\boldsymbol{x}_0$: the entities ∇g and g_t in the relation expressed by equation (4b) should vary at most linearly as a function of $\mathrm{d}\boldsymbol{x}$ for all locations $\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}$ within an entire environment $\mathrm{d}\boldsymbol{x}$ around $\boldsymbol{x}_0$:

$$\nabla g(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x})^{\mathrm{T}}\, \boldsymbol{u}(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}) + g_t(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}) = 0\,. \tag{20}$$

The dependence of ∇g, g_t, and $\boldsymbol{u}$ on $\mathrm{d}\boldsymbol{x}$ is made explicit by developing these entities into a Taylor series:

$$g_x(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}) = g_x + g_{xx}\,\mathrm{d}x + g_{xy}\,\mathrm{d}y + \mathrm{O}(\mathrm{d}^2)\,, \tag{21a}$$

$$g_y(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}) = g_y + g_{yx}\,\mathrm{d}x + g_{yy}\,\mathrm{d}y + \mathrm{O}(\mathrm{d}^2)\,, \tag{21b}$$

$$g_t(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}) = g_t + g_{tx}\,\mathrm{d}x + g_{ty}\,\mathrm{d}y + \mathrm{O}(\mathrm{d}^2)\,. \tag{21c}$$

All partial derivatives on the right-hand sides of (21) are taken at $\boldsymbol{x}_0$. $\mathrm{O}(\mathrm{d}^2)$ denotes terms which are at least of second order in the components of $\mathrm{d}\boldsymbol{x} = (\mathrm{d}x, \mathrm{d}y)^{\mathrm{T}}$. Analogously, we have

$$u(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}) = u + u_x\,\mathrm{d}x + u_y\,\mathrm{d}y + \mathrm{O}(\mathrm{d}^2)\,, \tag{22a}$$

$$v(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x}) = v + v_x\,\mathrm{d}x + v_y\,\mathrm{d}y + \mathrm{O}(\mathrm{d}^2)\,. \tag{22b}$$

Inserting (21) and (22) into (20) and ordering by increasing powers of $\mathrm{d}x$ and $\mathrm{d}y$ yields

$$\begin{aligned}0 = {} & \{g_x u + g_y v + g_t\} \\ & + \{g_{xx}u + g_{yx}v + g_x u_x + g_y v_x + g_{tx}\}\,\mathrm{d}x \\ & + \{g_{xy}u + g_{yy}v + g_x u_y + g_y v_y + g_{ty}\}\,\mathrm{d}y \\ & + \{g_{xx}u_x + g_{yx}v_x\}\,\mathrm{d}x^2 \\ & + \{g_{xy}u_x + g_{xx}u_y + g_{yy}v_x + g_{yx}v_y\}\,\mathrm{d}x\,\mathrm{d}y \\ & + \{g_{xy}u_y + g_{yy}v_y\}\,\mathrm{d}y^2\,. \end{aligned} \tag{23}$$

Since the coefficients of this binomial in $(\mathrm{d}x, \mathrm{d}y)$ are constants—namely the partial derivatives of g and $\boldsymbol{u}$ taken at $\boldsymbol{x}_0$—the right-hand side of (23) can only vanish for all $\mathrm{d}\boldsymbol{x}$ if each coefficient in braces is zero. We thus obtain six equations for the six unknown values of u, v, and their partial derivatives u_x, u_y, v_x and v_y:

$$g_x u + g_y v + g_t = 0\,, \tag{24a}$$

$$g_{xx}u + g_{yx}v + g_x u_x + g_y v_x + g_{tx} = 0\,, \tag{24b}$$

$$g_{xy}u + g_{yy}v + g_x u_y + g_y v_y + g_{ty} = 0\,, \tag{24c}$$

$$g_{xx}u_x + g_{yx}v_x = 0\,, \tag{24d}$$

$$g_{xy}u_x + g_{xx}u_y + g_{yy}v_x + g_{yx}v_y = 0\,, \tag{24e}$$

$$g_{xy}u_y + g_{yy}v_y = 0\,. \tag{24f}$$

The coefficient matrix of this system of linear equations does not have a full rank of six, however, as it can be easily seen if the coordinate system is aligned with the directions of principal curvature. In this case, $g_{xy} = 0$ and we obtain

$$\begin{aligned}\det\begin{pmatrix} g_x & g_y & 0 & 0 & 0 & 0 \\ g_{xx} & 0 & g_x & 0 & g_y & 0 \\ 0 & g_{yy} & 0 & g_x & 0 & g_y \\ 0 & 0 & g_{xx} & 0 & 0 & 0 \\ 0 & 0 & 0 & g_{xx} & g_{yy} & 0 \\ 0 & 0 & 0 & 0 & 0 & g_{yy} \end{pmatrix} \\ = -\,g_{xx}g_{yy}[g_x(g_y g_{xx} g_{yy}) - g_y(g_x g_{xx} g_{yy})] = 0\,.\end{aligned} \tag{25}$$

In order to determine all unknowns, we might add an additional requirement for $\boldsymbol{u}$, for example that

$$(\mathrm{rot}\ \boldsymbol{u})_z = 0$$

which implies

$$v_x - u_y = 0\,. \tag{26}$$

We then obtain from (24e) with $g_{xy} = 0$:

$$g_{xx}u_y + g_{yy}v_x = (g_{xx} + g_{yy})\,u_y = 0 \tag{27}$$

from which we may conclude for $g_{xx} + g_{yy} \neq 0$ that $u_y = v_x = 0$. Since $g_{xx} + g_{yy} \neq 0$ implies that at least one of these two second partial derivatives must be different from zero, equation (24d) or (24f) will yield, in view of $g_{xy} = 0$, the additional results $u_x = 0$ or $v_y = 0$ or both.

If we consider these results, the following conclusion emerges. The approaches discussed in Sections 3–5 selected a small, finite image region from one frame and described the structural gray value variation within this region as a low-order bivariate polynomial in the image plane coordinates. Assuming a constant displacement for the entire region, one had to search for a matching structural gray value variation in an adjacent image frame at a position not too far away from the center $\boldsymbol{x}_0$ of the original image region. These approaches are not concerned with potential local variations of the optical flow, but emphasize the match between structural gray value variations from adjacent image frames.

The approach described by equations (20) through (24) admits a limited variability for the optical flow field. As it turns out in (25), even a considerable gray value structure—i.e. nonzero first and second partial derivatives of the picture function $g(x, y)$ with respect to the image plane coordinates x and y—does not allow to determine $\boldsymbol{u}(\boldsymbol{x}_0 + \mathrm{d}\boldsymbol{x})$ in the form of $\boldsymbol{u}(\boldsymbol{x}_0)$ and the first partial derivatives of $\boldsymbol{u}$ with respect to x and y at $\boldsymbol{x}_0$. A weak additional assumption, however, is sufficient to fix the values of $\nabla \boldsymbol{u}(\boldsymbol{x}_0)$ to zero and, thereby, turn the underdetermined system of equations (24) for six variables into an overdetermined system of three equations resulting from (24a)–(24c) for the two unknowns $u(\boldsymbol{x}_0)$ and $v(\boldsymbol{x}_0)$.

The assumptions made here are weaker than the requirement of constant $\boldsymbol{u}$, but they have the same effect:

(1) $\boldsymbol{u} = \mathrm{const}$ follows from these assumptions as a solution.

(2) The resulting equations are equivalent to equations (15a) and (15b) or equations (16) in combination with equation (4b). At a gray value corner,

these equations yield a solution very similar to the one discussed on the basis of different premises in Section 3.

7. The Smoothness Requirement of Horn and Schunck

This section will extend the investigation started in the preceding section by substituting the smoothness requirement of Horn and Schunck [11] for the assumption that the optical flow varies at most linearly as a function of the image plane coordinates x and y within the image region under consideration. The assumptions about the local gray value structure are again introduced by neglecting terms of higher than second order in the Taylor expansion of the gray value at the center $\boldsymbol{x}_0$ of the image region, or, alternatively, by a first-order Taylor expansion of $\nabla g(\boldsymbol{x})$ and $g_t(\boldsymbol{x})$ at $\boldsymbol{x}_0 = 0$. Since the dependency of $\boldsymbol{u}(\boldsymbol{x}, t)$ on t is not investigated in this context, this dependency will not be shown explicitly in the subsequent discussion.

Horn and Schunck [11] introduced the following smoothness requirement for $\boldsymbol{u}(\boldsymbol{x})$:

$$\iint \mathrm{d}x\,\mathrm{d}y\{(\nabla g^{\mathrm{T}}\boldsymbol{u} + g_t)^2 + \alpha^2(u_x^2 + u_y^2 + v_x^2 + v_y^2)\} = \text{minimum}\,. \tag{28}$$

In this equation, the partial derivatives of the gray value as well as $\boldsymbol{u}$ and its partial derivatives are functions of $\boldsymbol{x} = (x, y)^{\mathrm{T}}$. Equation (28) supplements the minimization requirement for the square of the left-hand side of (4b) by the smoothness requirement expressed as

$$\alpha^2(u_x^2 + u_y^2 + v_x^2 + v_y^2)\,. \tag{29}$$

This smoothness term limits the variation of $\boldsymbol{u}(\boldsymbol{x})$ as a function of $\boldsymbol{x}$. The requirement expressed by (20) had been formulated with the same intention, although it is not strong enough to fully determine $\boldsymbol{u}(\boldsymbol{x})$.

In order to simplify the subsequent discussion, we assume—without loss of generality—that the coordinate system has been aligned with the directions of principal curvatures of $g(0)$, i.e. $g_{xy} = 0$. Introduction of the first-order Taylor expansion for $\nabla g(\boldsymbol{x})$ and $g_t(\boldsymbol{x})$ into (28) yields

$$\iint \mathrm{d}x\,\mathrm{d}y\{[(g_x + g_{xx}x)\,u + (g_y + g_{yy}y)\,v + (g_t + g_{tx}x + g_{ty}y)]^2 + \alpha^2(u_x^2 + u_y^2 + v_x^2 + v_y^2)\} = \text{minimum}\,. \tag{30}$$

It will now be shown that there exist solutions for this minimization problem which yield the value zero for the integral in (30). If it is possible to find constant values for u and v which let the expression in the square brackets vanish identically, then the problem is already solved, because the smoothness term then vanishes identically, too.

The first integrand of (30) represents a bivariate polynomial in the image plane coordinates x and y which should vanish identically. This is only possible, in general, if all coefficients of this bivariate polynomial vanish. We obtain, therefore, the following equations:

$$0 = (g_x u + g_y v + g_t)\,, \tag{31a}$$

$$0 = (g_{xx} u + g_{tx})\,, \tag{31b}$$

$$0 = (g_{yy} v + g_{ty})\,. \tag{31c}$$

If all second partial derivatives of $g(\boldsymbol{x})$ vanish at $\boldsymbol{x} = 0$, we encounter the same problem discussed with respect to (4b), i.e. there is not sufficient information in the gray value structure to estimate both components u and v of $\boldsymbol{u}$. We assume, therefore, that at least one of the second partial derivatives g_{xx} or g_{yy} is nonzero. The possible solutions are now obtained by a case analysis:

(1) Both g_x as well as g_{xx} are zero simultaneously. In this case, there is no gray value variation up to second order in the x direction and it is impossible to estimate u. An analogous observation can be made regarding the partial derivatives with respect to y. In accordance with our intuition, in these situations there is no solution for $\boldsymbol{u}$ of the kind we are currently interested in.

(2) At least one of the partial derivatives of $g(\boldsymbol{x})$ with respect to x is different from zero at $\boldsymbol{x} = 0$:

(i) If $g_x = 0$, then g_{xx} must be nonzero according to our assumptions, and we obtain from (31b)

$$u = -\, g_{tx}/g_{xx}\,. \tag{32}$$

(ii) If $g_x \neq 0$, but $g_{xx} = 0$, then g_{yy} must be different from zero according to our assumptions. Equation (31c) will yield

$$v = -\, g_{ty}/g_{yy}\,. \tag{33}$$

Since v is thereby known, we can use (31a) to determine u:

$$u = -\frac{g_y v + g_t}{g_x} = \frac{g_y g_{ty} - g_t g_{yy}}{g_{yy} g_x}\,. \tag{34}$$

Analogous considerations with respect of g_y and g_{yy} will result again in the solutions given by (32) and (33). Instead of (34) we obtain

$$v = -\frac{g_x u + g_t}{g_y} = \frac{g_x g_{tx} - g_t g_{xx}}{g_{xx} g_y} . \tag{35}$$

In the case of a gray value extremum, both first partial derivatives with respect to x and y vanish and the solution is given by (32) and (33). Equation (31a) then implies that $g_t = 0$ which is nothing but a consequence of the assumption that $g(\boldsymbol{x}_0)$ is a gray value extremum: if the first-order change of g with respect to the spatial coordinates should vanish and if the change of g with respect to time should be due to a shift of the gray value structure, there can be no first-order change of g with respect to time either.

In the case of a gray value corner, see equations (5), we obtain the solutions in the form discussed in Appendix A. This implies that $g_{tx} = 0$ which is reasonable since $g_{tx} = g_{xt}$ and we may interpret $g_{xt} = 0$ as the requirement that a maximal slope of g in x-direction will not exhibit first-order changes with time.

We thus see that gray value structures which are sufficiently well localized to support the expectation of a well defined optical flow vector, namely the gray value corner and the gray value extremum, lead indeed to an estimate for both components of the optical flow vector. The values for $\boldsymbol{u}$ which we obtain in these situations are equivalent to the results obtained earlier. Thus it has been shown that the smoothness requirement of Horn and Schunk [11] fully determines the optical flow vector without any additional approximations, provided the local gray value structure is characteristic enough.

8. Consequences for the "Oriented Smoothness" Requirement

The smoothness requirement of Horn and Schunck [11] does not take into account the consideration that a gray value transition front might be the image of an occluding edge across which the optical flow field needs not exhibit a smooth variation. Nagel [16] modified the smoothness requirement by the introduction of a weight matrix which depends on the gray value changes in such a way that the smoothness requirement would be retained essentially only for the optical flow component perpendicular to strong gray value transitions. The smoothness requirement for the optical flow component in the direction of the gray value change would be suppressed since this component could be determined from the gray value change itself.

The formulation originally suggested by Nagel [16] was

$$\iint \mathrm{d}x\, \mathrm{d}y \{(g_1(\boldsymbol{x}) - g_0(\boldsymbol{x} - \boldsymbol{u}\, \mathrm{d}t))^2 + \alpha^2 \operatorname{trace}((\nabla \boldsymbol{u})^{\mathrm{T}} C^{-1}(\nabla \boldsymbol{u}))\} = \text{minimum} , \tag{36}$$

where the weight matrix C^{-1} is the inverse of the matrix C given by (12). Extensive numerical experiments with the system of partial differential equations derived as Euler–Lagrange equations from (36) have shown that it is advantageous to use a different weight matrix [5–7, 21–23]:

$$\iint \mathrm{d}x\, \mathrm{d}y \{ (g_1(\boldsymbol{x}) - g_0(\boldsymbol{x} - \boldsymbol{u}\, \mathrm{d}t))^2 + \alpha^2 \operatorname{trace}((\boldsymbol{\nabla u})^{\mathrm{T}} W^*(\boldsymbol{\nabla u})) \} = \text{minimum} \tag{37}$$

with

$$W^* = \frac{w^* + I\gamma}{\operatorname{trace}(w^* + I\gamma)} \tag{38a}$$

where γ is a constant, I represents the 2×2 identity matrix and

$$w^* = \left\{ \begin{pmatrix} g_y \\ -g_x \end{pmatrix} \begin{pmatrix} g_y \\ -g_x \end{pmatrix}^{\mathrm{T}} + \overline{x^2} \begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix} \begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix}^{\mathrm{T}} \right\} . \tag{38b}$$

w^* is related to C by

$$C^{-1} = \frac{w^*}{\det(w^*)} . \tag{39}$$

This formulation had been guided by the desire to incorporate the knowledge about the importance of second-order derivatives for the determination of optical flow vector fields as discussed in Sections 3–6 into the minimization approach. According to the insight gained in Section 7, we do not need to develop the term $g(\boldsymbol{x} - \boldsymbol{u}\, \mathrm{d}t)$ in (37) into a Taylor series including up to second-order terms just to make sure that gray value corners can be handled appropriately. The structure of a gray value corner or extremum will enforce appropriate estimates for both flow vector components. As a consequence, we may write instead of (37)

$$\iint \mathrm{d}x\, \mathrm{d}y \{ (\boldsymbol{\nabla} g^{\mathrm{T}} \boldsymbol{u} + g_t)^2 + \alpha^2 \operatorname{trace}((\boldsymbol{\nabla u})^{\mathrm{T}} W^*(\boldsymbol{\nabla u})) \} = \text{minimum} . \tag{40}$$

The same argument leading from equation (37) to (40) suggests to drop the terms containing second-order derivatives of g from W^* as well. We then obtain the following formulation:

$$\iint \mathrm{d}x\, \mathrm{d}y \{ (\boldsymbol{\nabla} g^{\mathrm{T}} \boldsymbol{u} + g_t)^2 + \alpha^2 \operatorname{trace}((\boldsymbol{\nabla u})^{\mathrm{T}} W(\boldsymbol{\nabla u})) \} = \text{minimum} \tag{41}$$

with the weight matrix W given by

$$W = \frac{w + I\gamma}{\operatorname{trace}(w + I\gamma)}, \tag{42a}$$

where I represents the 2×2 identity matrix and

$$w = \left\{ \begin{pmatrix} g_y \\ -g_x \end{pmatrix} \begin{pmatrix} g_y \\ -g_x \end{pmatrix}^{\mathrm{T}} \right\}. \tag{42b}$$

The new version of the weight matrix may be written in a more concise way

$$W = \frac{\begin{pmatrix} g_y^2 + \gamma & -g_x g_y \\ -g_x g_y & g_x^2 + \gamma \end{pmatrix}}{g_x^2 + g_y^2 + 2\gamma}. \tag{43}$$

This form of the weight matrix exhibits the basic idea behind the "oriented smoothness" requirement in an especially clear manner. If we set $\gamma = 0$—i.e. omit the term which has been introduced just to make sure that the weight matrix specializes to the identity matrix if the gray value structure becomes locally constant—the smoothness term in (41) represents the square of the projection of the optical flow vector onto the direction perpendicular to the gradient:

$$\begin{aligned}
&\operatorname{trace}((\nabla \boldsymbol{u})^{\mathrm{T}} W (\nabla \boldsymbol{u})) \\
&= \operatorname{trace}\left(\begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix}^{\mathrm{T}} W \begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix} \right) \\
&= \frac{\begin{pmatrix} u_x \\ u_y \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} g_y^2 + \gamma & -g_x g_y \\ -g_x g_y & g_x^2 + \gamma \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix}}{g_x^2 + g_y^2 + 2\gamma} \\
&\quad + \frac{\begin{pmatrix} v_x \\ v_y \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} g_y^2 + \gamma & -g_x g_y \\ -g_x g_y & g_x^2 + \gamma \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix}}{g_x^2 + g_y^2 + 2\gamma}
\end{aligned} \tag{44a}$$

$$\begin{aligned}
&= \frac{\left[\begin{pmatrix} u_x \\ u_y \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} g_y \\ -g_x \end{pmatrix} \right]^2}{\|\nabla g\|^2 + 2\gamma} + \frac{\left[\begin{pmatrix} v_x \\ v_y \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} g_y \\ -g_x \end{pmatrix} \right]^2}{\|\nabla g\|^2 + 2\gamma} \\
&\quad + \gamma \frac{u_x^2 + u_y^2 + v_x^2 + v_y^2}{\|\nabla g\|^2 + 2\gamma}.
\end{aligned} \tag{44b}$$

Since the norm of the tangent vector $(g_y, -g_x)^{\mathrm{T}}$ to a gray value transition is identical to the norm of the gradient vector $(g_x, g_y)^{\mathrm{T}}$, the first term of (44b) is equal to the square of the projection of $\nabla u = (u_x, u_y)^{\mathrm{T}}$ onto the tangent direction if we neglect γ. The second term expresses the equivalent projection of $\nabla v = (v_x, v_y)^{\mathrm{T}}$. Equation (44b) illuminates the way in which a vanishing gray value gradient will transform the "oriented smoothness" requirement into a general, unoriented one as introduced originally by Horn and Schunck [11].

The Euler–Lagrange equations for the minimization problem of (41) can be written in the following form:

$$(\nabla g^{\mathrm{T}} \boldsymbol{u} + g_t)\, g_x + \alpha^2 \frac{\nabla g^{\mathrm{T}}\left\{\begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix} + 2\begin{pmatrix} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{pmatrix} W\right\}\begin{pmatrix} u_x \\ u_y \end{pmatrix}}{g_x^2 + g_y^2 + 2\gamma}$$

$$- \alpha^2 \operatorname{trace}(W \nabla\nabla u) = 0\,, \tag{45a}$$

$$(\nabla g^{\mathrm{T}} \boldsymbol{u} + g_t)\, g_y + \alpha^2 \frac{\nabla g^{\mathrm{T}}\left\{\begin{pmatrix} g_{yy} & -g_{xy} \\ -g_{xy} & g_{xx} \end{pmatrix} + 2\begin{pmatrix} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{pmatrix} W\right\}\begin{pmatrix} v_x \\ v_y \end{pmatrix}}{g_x^2 + g_y^2 + 2\gamma}$$

$$- \alpha^2 \operatorname{trace}(W \nabla\nabla v) = 0\,. \tag{45b}$$

In Nagel [16], the weight matrix W^* comprised second-order partial derivatives of the picture function $g(\boldsymbol{x})$. Since the derivative of this weight matrix with respect to x and y would have comprised third-order partial derivatives of $g(\boldsymbol{x})$, the initial evaluation of this approach threatened to become a bit too complicated. Therefore, this derivative of the weight matrix had been dropped at that time. The approach formulated in (41), however, contains only first-order partial derivatives of $g(\boldsymbol{x})$ even in the weight matrix W. The partial derivatives of this weight matrix, therefore, contain only partial derivatives of $g(\boldsymbol{x})$ of second order. The formulation of (41) has thus the added advantage that no terms have to be neglected just in order to restrain the required computations.

It is seen immediately that the solutions of equations (30) given by (32)–(35) are also solutions of (41): for constant values of u and v, the oriented smoothness term vanishes just like the smoothness term of Horn and Schunck [11] vanishes in (30).

A final consideration will show that the minimization problem formulated by Hildreth [9, 10] can be considered as a special case of (41). Hildreth formulated the estimation of optical flow vectors as a one-dimensional minimization problem along a closed zero-crossing contour:

$$\oint_{\text{zero-crossing contour}} \mathrm{d}s\left\{\beta(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{n} - u^{\perp})^2 + \left(\frac{\partial u}{\partial s}\right)^2 + \left(\frac{\partial v}{\partial s}\right)^2\right\} = \text{minimum}\,. \tag{46}$$

Here, $\boldsymbol{n}$ denotes the unit length normal vector to the zero-crossing contour at a point specified by the arclength s. $u^{\perp}$ denotes the estimate of the optical flow vector component perpendicular to the zero-crossing contour. This estimate is obtained by Hildreth in a manner analogous to (4b) from the spatial and temporal derivatives of $g(\boldsymbol{x})$ convolved with the Laplacian of a Gaussian. If we use (4b) directly instead of the first term in the integrand of (46), we obtain:

$$\oint_{\substack{\text{zero-crossing}\\ \text{contour}}} \mathrm{d}s\left\{(\nabla g^{\mathrm{T}}\boldsymbol{u} + g_t)^2 + \frac{1}{\beta}\left[\left(\frac{\partial u}{\partial s}\right)^2 + \left(\frac{\partial v}{\partial s}\right)^2\right]\right\} = \text{minimum}\,. \tag{47}$$

If we assume that all partial derivatives of $g(\boldsymbol{x})$ are negligible except at the zero-crossing contours, then the integrand of (41) will yield essential contributions only along a zero-crossing contour, apart from the smoothness terms proportional to γ. In this respect, the approach of Hildreth [9, 10] can be considered as a special case of (41). The second term in the integrand of (47) expresses the smoothness requirement for the component of the optical flow vector parallel to the zero-crossing contour. This discussion of the oriented smoothness requirement in connection with (44) shows the specialization of the more general requirement incorporated into (41) to the situation where the gray value gradient is taken to be constant along the zero-crossing contour as in the formulation of Hildreth.

9. Conclusion

The discussion started with the observation by Nagel [15] that the change in gradient direction, i.e. second-order partial derivatives of $g(\boldsymbol{x})$, has to be taken into account in order to estimate locally both components of the optical flow field. The approaches formulated later by Haralick and Lee [8] as well as Tretiak and Pastor [26] have been related quantitatively to results obtained by Nagel [15]. All these estimation approaches are based on the assumption that the optical flow vector $\boldsymbol{u}$ is constant within an image region around the location of a gray value corner or extremum.

Section 6 investigated whether one might substitute a less stringent requirement for that of constant $\boldsymbol{u}$: would it be sufficient to require that $\boldsymbol{u}$ is at most a linear function of the image plane coordinates, provided the gray value structure is well enough articulated within the region under consideration? It turned out that this is not quite sufficient to determine both componens of $\boldsymbol{u}$, but a relatively weak additional assumption allowed to recover $\boldsymbol{u}$ together with a set of equations which are equivalent to those treated in preceding sections.

The weaker assumption of Horn and Schunck [11], namely that the optical flow vector field is not constant, but exhibits only a smooth variation, resulted in a system of partial differential equations for $\boldsymbol{u}(\boldsymbol{x})$. Horn and Schunck presented an approximate solution for this system of partial differential equations.

The main result of Section 7 from this contribution is the fact that an exact solution to the equations derived by Horn and Schunck [11] yields a locally constant optical flow vector field at gray value corners or extrema. Thereby, the approach of Horn and Schunck could be unified with the observations by Nagel [15]. This result suggested to modify the "oriented smoothness" approach proposed by Nagel [16] in such a manner that the resulting system of equations becomes much simpler. It turns out that this simplification has two additional benefits:

(1) no terms have to be dropped from the Euler–Lagrange equations resulting from (41), and

(2) the approach by Hildreth [9, 10] can be seen as a kind of special case for equation (41) in an even more direct way than discussed previously by Nagel [17].

Multigrid methods especially developed by Enkelmann [5–7] for the solution of the Euler–Lagrange equations resulting from the original "oriented smoothness" proposal by Nagel [16] will be used in order to explore the new system of partial differential equations (45) derived in this contribution.

Appendix A. The Approach of Tretiak and Pastor Analyzed at a Gray Value Corner

We may specialize the result of equation (18) to the situation where the coordinate system has been aligned with the directions of principal curvature, i.e. $g_{xy} = 0$, and to a gray value corner as characterized by (5):

$$\begin{pmatrix} u \\ v \end{pmatrix} = -\left(\begin{pmatrix} g_x & 0 & 0 \\ 0 & 0 & g_{yy} \end{pmatrix} \begin{pmatrix} g_x & 0 \\ 0 & 0 \\ 0 & g_{yy} \end{pmatrix} \right)^{-1} \begin{pmatrix} g_x & 0 & 0 \\ 0 & 0 & g_{yy} \end{pmatrix} \begin{pmatrix} g_t \\ g_{tx} \\ g_{ty} \end{pmatrix}, \tag{A.1a}$$

or

$$\begin{pmatrix} u \\ v \end{pmatrix} = -\begin{pmatrix} g_x^2 & 0 \\ 0 & g_{yy}^2 \end{pmatrix}^{-1} \begin{pmatrix} g_x g_t \\ g_{yy} g_{ty} \end{pmatrix} = -\begin{pmatrix} g_t / g_x \\ g_{ty} / g_{yy} \end{pmatrix}. \tag{A.1b}$$

This corresponds to the result obtained in (7b) as will be shown by the following considerations. We write

$$\begin{aligned} g_y(x, y, t + \mathrm{d}t) = g_y(0, 0, t) &+ g_{yx}(0, 0, t)\, x \\ &+ g_{yy}(0, 0, t)\, y + g_{yt}(0, 0, t)\, \mathrm{d}t\,. \end{aligned} \tag{A.2}$$

Since we assume that $g_{xy} = 0$ due to the alignment of the coordinate axes with the directions of principal curvature at $\boldsymbol{x}_0 = 0$ and that $g_y(0, 0, t) = 0$ according to the requirements for a gray value corner, we have for $\mathrm{d}t = 1$

$$g_y(0, 0, t+1) = g_{yt}(0, 0, t) , \tag{A.3a}$$

which can be written for $(t + \mathrm{d}t) = t_1$ as

$$\begin{aligned} g_y(0, 0, t+1) &= g_y(\boldsymbol{x}_0, t_1) \\ &= g_{1,y} = g_{yt}(0, 0, t) = g_{ty} . \end{aligned} \tag{A.3b}$$

The lower half of the vector equation (A.1b) is, therefore, equal to (7b). The upper half of the vector equation (A.1b) corresponds to the expected solution for u in the case where the gradient of g is aligned with the x-axis. The term with v^2 in (7a) is due to the nonlinear approach and does not show up in (A.1b).

ACKNOWLEDGMENT

I thank W. Enkelmann, A. Korn, and Th. Krämer for comments on a draft version of this contribution, W. Enkelmann for his permission to use Figs. 1–3 from his dissertation, and Th. Krämer for making the gray value images for Figs. 4–6 available.

REFERENCES

1. Dreschler, L., Ermittlung markanter Punkte auf den Bildern bewegter Objekte und Berechnung einer 3-D Beschreibung auf dieser Grundlage, Dissertation, Fachbereich Informatik der Universität Hamburg, F.R.G., 1981.
2. Dreschler, L. and Nagel, H.-H., Volumetric model and 3D trajectory of a moving car derived from monocular TV frame sequences of a street scene, in: *Proceedings IJCAI-81*, Vancouver, BC (1981) 692–697.
3. Dreschler, L. and Nagel, H.-H., On the selection of critical points and local curvature extrema of region boundaries for interframe matching, in: *Proceedings International Conference on Pattern Recognition*, München, F.R.G. (1982) 542–544.
4. Dreschler, L. and Nagel, H.-H., Volumetric model and 3D trajectory of a moving car derived from monocular TV frame sequences of a street scene, *Comput. Vision Graph. Image Process.* **20** (1982) 199–228.
5. Enkelmann, W., Mehrgitterverfahren zur Ermittlung von Verschiebungsvektorfeldern in Bildfolgen, Dissertation, Fachbereich Informatik der Universität Hamburg, Hamburg, F.R.G., 1985.
6. Enkelmann, W., Ein Mehrgitterverfahren zur Ermittlung von Verschiebungsvektorfeldern in Bildfolgen, in: H. Niemann (Ed.), *7. DAGM-Symposium Mustererkennung*, *Informatik-Fachberichte* **107** (Springer, Berlin, 1985) 97–101.
7. Enkelmann, W., Investigation of multigrid algorithms for the estimation of optical flow fields in image sequences, in: *Proceedings Workshop on Motion*: *Representation and Analysis*, Charleston, SC (1986) 81–87.
8. Haralick, R.M. and Lee, J.S., The facet approach to optic flow, in: L.S. Baumann (Ed.), *Proceedings Image Understanding Workshop* (Science Applications, Arlington, VA, 1983) 84–93.
9. Hildreth, E.C., The measurement of visual motion, Ph. D. Dissertation, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1983.
10. Hildreth, E.C., Computations underlying the measurement of visual motion, *Artificial Intelligence* **23** (1984) 309–354.

11. Horn, B.K.P. and Schunck, B.G., Determining optical flow, *Artificial Intelligence* **17** (1981) 185-203.
12. Kitchen, L. and Rosenfeld, A., Gray-level corner detection, TR-887, Computer Science Center, University of Maryland, College Park, MD, 1980.
13. Kitchen, L. and Rosenfeld, A., Gray-level corner detection, *Pattern Recogn. Lett.* **1** (1982) 95-102.
14. Koenderink, J.J. and van Doorn, A.J., How an ambulant observer can construct a model of the environment from the geometrical structure of the visual inflow, in: G. Hauske and E. Butenandt (Eds.), *Kybernetik 1977* (Oldenbourg, München, 1977).
15. Nagel, H.-H., Displacement vectors derived from second order intensity variations in image sequences, *Comput. Vision Graph. Image Process.* **21** (1983) 85-117.
16. Nagel, H.-H., Constraints for the estimation of displacement vector fields from image sequences, in: *Proceedings IJCAI-83*, Karlsruhe, F.R.G. (1983) 945-951.
17. Nagel, H.-H., Recent advances in image sequence analysis, in: *Premier Colloque Image: Traitement, Synthèse, Technologie et Applications*, Biarritz, France (1984) 545-558.
18. Nagel, H.-H., Analyse und Interpretation von Bildfolgen, Teil I, *Inf. Spektrum* **8** (1985) 178-200; Teil II, *Inf. Spektrum* **8** (1985) 312-327.
19. Nagel, H.-H., Image sequences—Ten (octal) years—from phenomenology towards a theoretical foundation, in: *Proceedings Eighth International Conference on Pattern Recognition*, Paris, France (1986) 1174-1185.
20. Nagel, H.-H. and Enkelmann, W., Investigation of second order gray value variations to estimate corner point displacements, in: *Proceedings International Conference on Pattern Recognition*, München, F.R.G. (1982) 768-773.
21. Nagel, H.-H. and Enkelmann, W., Towards the estimation of displacement vector fields by "oriented smoothness" constraints, in: *Proceedings International Conference on Pattern Recognition*, Montreal, Que. (1984) 6-8.
22. Nagel, H.-H. and Enkelmann, W., Berechnung von Verschiebungsvektorfeldern in Bildbereichen mit linienhaften oder partiell homogenen Grauwertverteilungen, in: W. Kropatsch. *6. DAGM-Symposium Mustererkennung, Informatik-Fachberichte* **87** (Springer, Berlin, 1984) 154-160.
23. Nagel, H.-H. and Enkelmann, W., An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences, *IEEE Trans. Pattern Anal. Mach. Intell.* **8** (1986) 565-593.
24. Schunck, B.G., The motion constraint equation for optical flow, in: *Proceedings International Conference on Pattern Recognition*, Montréal, Que. (1984) 20-22.
25. Schunck, B.G. and Horn, B.K.P., Constraints on optical flow computation, in: *Proceedings IEEE Conference on Pattern Recognition and Image Processing PRIP-81*, Dallas, TX (1981) 205-210.
26. Tretiak, O. and Pastor, L., Velocity estimation from image sequences wih second order differential operators, in: *Proceedings International Conference on Pattern Recognition*, Montréal, Que. (1984) 16-19.
27. Ullman, S., *The Interpretation of Visual Motion* (MIT Press, Cambridge, MA, 1979).
28. Westphal, H. and Nagel, H.-H., Toward the derivation of three-dimensional descriptions from image sequences for nonconvex moving objects, *Comput. Vision Graph. Image Process.* **24** (1986) 302-320.

Received February 1986; revised version received January 1987