COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 21, 85–117 (1983)

Displacement Vectors Derived from Second-Order Intensity Variations in Image Sequences

HANS-HELLMUT NAGEL

Fachbereich Informatik, Universitaet Hamburg, Schlueterstrasse 70, 2000 Hamburg 13 Germany

Received March 11, 1982; revised April 7, 1982

A local approach for interframe displacement estimates is developed by minimization of the squared differences between a second-order Taylor expansion of gray values from one frame and the observed gray values within the same window from the next frame. If the second-order terms in the Taylor expansion are significant, a system of two coupled nonlinear equations for the two unknown components of the displacement vector can be derived. In the special case of "gray value corners," these equations can be simplified to facilitate a closed form solution. An iterative refinement procedure is developed to extend these estimates for image regions which do not exhibit exactly the properties of "gray value corners." The minimization approach is generalized in such a way that the approach of Horn and Schunck (*Artif. Intell.* 17, 1981, 185–203) can be recognized as a special case of this generalized form which should be applicable even across occluding edges. It thus appears to be an interesting model for the local computation of optical flow.

1. INTRODUCTION

The frame-to-frame displacement of characteristic spatial gray value variations can yield information about the three-dimensional structure of objects and their motion relative to the sensor which records an image sequence from a scene with these objects (see [1–12] and further references quoted there).

Stereopsis can be considered as a special case where the number of frames is restricted to two and the displacement direction is known [13–15].

Attempts to improve bandwidth compression by motion-compensated prediction for interframe coding provide another reason to study the determination of image displacement vector fields (see, e.g., [16–18] for reviews of this area).

Against this background of wide-spread interest, numerous ideas have been investigated to determine image displacement vector fields. One might distinguish three categories among these investigations. Approaches in the first category define primitive features characterizing significant local gray value variations. Such features are extracted from each image and matched from frame to frame. The second category comprises approaches which construct complex nonlocal descriptors for image patches by combining several primitive features. Heuristic (dis-) similarity functions are employed to grade candidate matches between complex descriptors from different frames. The so-called "nonmatching" approaches toward the determination of displacement vector fields are assigned to a third category. An explicit simple parametric function of the image coordinates is used to approximate a local gray value variation by appropriate choice of the parameter values. Methods from analysis are employed to map the domain of this function into a neighborhood

0734-189X/83/010085-33$3.00/0
Copyright © 1983 by Academic Press, Inc.
All rights of reproduction in any form reserved.

from a preceding or subsequent frame such that this fully specified function represents an acceptable approximation to the gray values recorded there.

Examples for the second category are the tracking experiments in blocks-world scenes by Roach and Aggarwal [19] or Radig and his students [20], for rocket scenes by Gilbert *et al.* [21], and for street scenes by Radig [22, 23]. Additional examples can be found in the comprehensive survey of application-oriented literature about image sequences by Nagel [24]. Interframe matching techniques applied in this context have recently been surveyed by Aggarwal *et al.* [25]. Since the design of complex descriptors and (dis-) similarity grading functions is usually based on scene-dependent heuristics which are difficult to analyze in general, this category will not be pursued further at this point.

The "interest operator" of Moravec [26] and the "directed variances" described by Hannah [27] are examples of feature detectors employed in the first category. They represent heuristics to search for gray value variations with a steeply decreasing autocorrelation function. Failures of Moravec's interest operator (see [28, 29]) caused Dreschler and Nagel [5] to scrutinize the gray value variation around visually acceptable cornerpoints in digitized TV images of street scenes. Based on this experience they formulated an algorithm to determine local descriptors suitable for an interframe match. First results have been presented in [29–31]. Independently, Kitchen and Rosenfeld [32] developed a gray level corner detector. Upon closer inspection it turns out that the concepts behind these two approaches are identical and can be justified in a rather fundamental way.

The contribution of this paper is an analysis which shows that the properties of feature locations selected in this manner facilitate the determination of both components for the associated displacement vector without recourse to approximate solutions of nonlinear equations. This result provides a common conceptual foundation for the approaches of the first and third categories discussed in the preceding paragraphs. Relations between this method and the directionally selective operator investigated by Marr and Ullman [33] will be pointed out.

Section 2 will introduce the notation and discuss the relation between the corner detector of Kitchen and Rosenfeld [32] and that of Dreschler and Nagel [5] followed by a discussion of previous attempts to determine the displacement vector based on "nonmatching" approaches. Section 4 will employ the concept of a "gray value corner" to determine the displacement vector of such a corner between consecutive frames. An error analysis of this method provides the basis to extend this approach in the Section 5 by iterative approximation to an environment around such feature locations.

2. MODELING GRAY VALUE VARIATIONS FOR A CORNER DETECTOR

Figure 1 presents an image from a TV frame sequence studied by Dreschler and Nagel [5]. The gray value variation around the upper left rear window of the white cab in the center of Fig. 1 is shown as a 3-D plot in Fig. 2. An idealized version of this gray value variation is given in Fig. 3 with coordinate axes for the image plane coordinates x, y and the gray value g indicated.

To simplify the notation, the origin of the coordinate system is assumed to be centered at the point of interest, i.e., $\mathbf{X} = (x, y)'$, where the prime indicates transposition. As is well known from elementary differential geometry, a surface may be locally approximated by a second-order function of the coordinate vector

FIG. 1. The first half frame of frame 1 from a TV frame sequence recorded from a street scene.

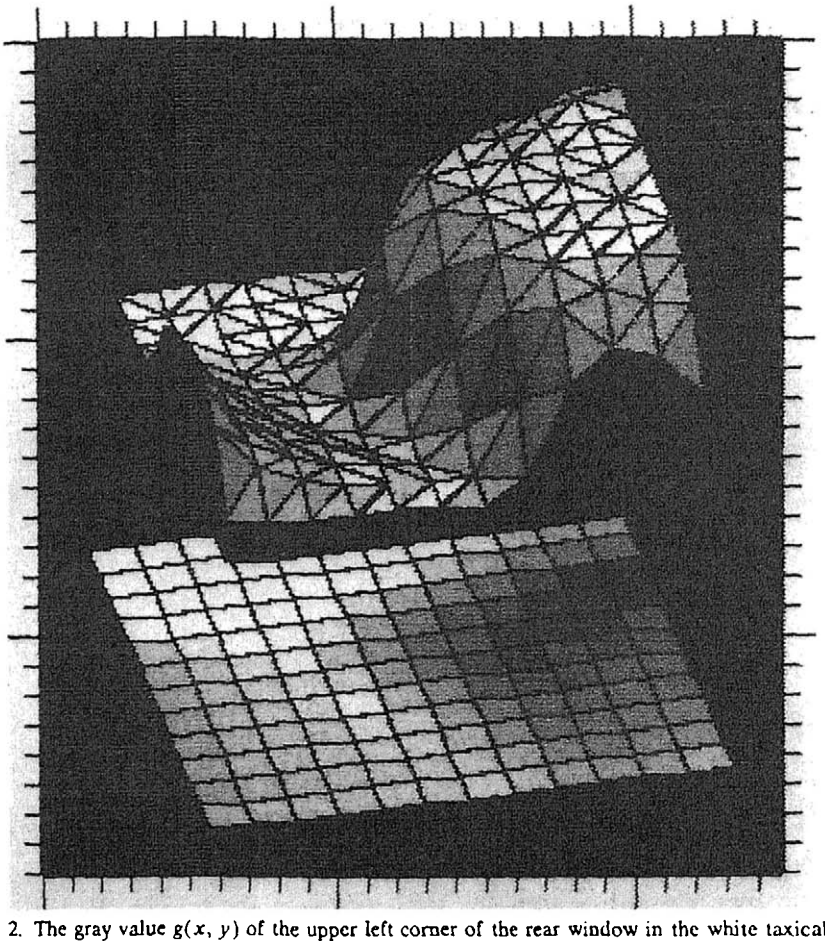

FIG. 2. The gray value $g(x, y)$ of the upper left corner of the rear window in the white taxicab from Fig. 1, plotted as a function of the raster coordinates x and y. The selected area from Fig. 1 can be seen below the pseudo 3-D plot.

$\mathbf{X} = (x, y)'$,

$$g(\mathbf{X}) = g(\mathbf{X}_0) + (\nabla g)' * \mathbf{X} + \tfrac{1}{2}\mathbf{X}'(\nabla\nabla g)\mathbf{X} + \epsilon \tag{1a}$$

$$= g(\mathbf{X}_0) + g_x x + g_y y + \tfrac{1}{2}g_{xx}x^2 + g_{xy}xy + \tfrac{1}{2}g_{yy}y^2 + \epsilon, \tag{1b}$$

where g_x and g_y stand for the partial derivatives of $g(\mathbf{X})$ with respect to x and y, respectively, and g_{xx}, g_{xy}, g_{yy} represent the corresponding second partial derivatives, all taken at the point $\mathbf{X}_0 = (0, 0)'$. The symmetric matrix of second derivatives

$$(\nabla\nabla g) = \begin{pmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix} \quad \text{with} \quad g_{xy} = g_{yx} \tag{2a}$$

specifies the planar curvature at $\mathbf{X}_0$ for the intersection curve between the surface $g(\mathbf{X})$ and a plane containing the surface normal at $\mathbf{X}_0$. This matrix can be

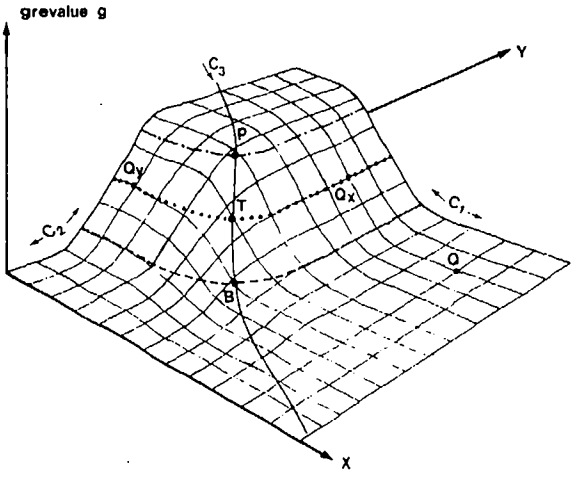

FIG. 3. Idealized sketch of $g(x, y)$ versus x, y for a "gray value corner" such as the one depicted in Fig. 2. The intersection curve $C1$ is parallel to the x axis and $C2$ is parallel to the y axis. The intersection curve $C3$ connects the minimum of the Gaussian curvature at the "bottom" point B with the turning point T of the slope and the maximum of the Gaussian curvature at the "promontory" point P. At Q_x the main curvature along the x direction crosses zero, and analogously for Q_y. At the point Q both main curvatures are zero.

transformed into a diagonal one by a suitable rotation of the coordinate system

$$\begin{pmatrix} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{pmatrix} \Rightarrow \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}, \tag{2b}$$

where the elements in the diagonal are denoted as the principle curvatures of the surface at the point $\mathbf{X}_0$. Since the trace and determinant of a matrix are invariant to rotations of the coordinate system, one obtains

$$\text{Laplacian:} \quad \nabla^2 g = g_{xx} + g_{yy} = \kappa_1 + \kappa_2 = 2((\kappa_1 + \kappa_2)/2), \tag{3a}$$

$$\text{Hessian:} \quad \det(\nabla \nabla g) = g_{xx} g_{yy} - g_{xy}^2 = \kappa_1 \kappa_2, \tag{3b}$$

i.e. the Laplacian of g yields twice the value of the average curvature and the determinant of the second derivatives ("Hessian") yields the product of the principle curvatures, called the "Gaussian curvature."

2.1 The Approach of Dreschler and Nagel

For all points along the intersection curve C1 in Fig. 3, the principal curvature corresponding to g_{yy} vanishes identically. Therefore, the Gaussian curvature is likewise identically zero along $C1$ and, by analogy, along $C2$. If we consider the intersection curve $C3$, however, the Gaussian curvature will be nonzero around the "promontory" point P and the "bottom" point B. At the promontory point P, one principal curvature represents the flattening of g from the steep ascent to the constant plateau. The second principal curvature at P corresponds to the turning orientation of the gradient projection into the xy plane, i.e., to the corner. Since the center of the osculating circle for the intersection curves corresponding to both

principal curvatures are located on the same side of the surface, the Gaussian curvature will be positive. At the bottom point B, one principal curvature again corresponds to the turning orientation of the gradient projection and will thus have the same sign as at P. The other principal curvature, however, corresponds to the flattening from descent to the constant base value. The center of its osculating circle is located at the viewer's side of the surface. Since the centers of the osculating circles are on different sides of the surface, the principal curvatures have opposite sign at this point and the Gaussian curvature will be negative.

The algorithm of Dreschler and Nagel [5] basically consists of the following rules:

(i) Determine Gaussian curvature.

(ii) Select locations of extremal—positive as well as negative—Gaussian curvature.

(iii) Match a location of maximum positive Gaussian curvature such as P with a location of extreme negative Gaussian curvature such as B provided that the directions of those principal curvatures which have opposite sign at B and P are approximately aligned.

(iv) Select the point T (i.e., the turning point of the S-like curve increasing in this direction from the lower to the higher gray value plateau), where this principal curvature crosses zero. This corresponds to the point of maximum slope on the curve between B and P.

2.2 The Approach of Kitchen and Rosenfeld

Kitchen and Rosenfeld [32] consider a unit vector tangent to a curve of constant gray value passing through the point of interest. Such a curve $\mathbf{X}(s)$ can be implicitly defined as a function of its arc length s by

$$g(\mathbf{X}(s)) = g(\mathbf{X}_0) \tag{4}$$

provided $\mathbf{X}_0$ does not specify the position for a local extremum of the gray value surface $g(\mathbf{X})$. Under these conditions, the tangent vector $(dx/ds, dy/ds)'$ to this curve can be obtained from

$$dg/ds = 0 = g_x\, dx/ds + g_y\, dy/ds \tag{5a}$$

which yields (taking into account the requirement that the tangent vector should have unit length)

$$dx/ds = -g_y \Big/ \sqrt{g_x^2 + g_y^2} \quad \text{and} \quad dy/ds = g_x \Big/ \sqrt{g_x^2 + g_y^2}\,. \tag{5b}$$

The rate of change of this unit tangent vector, i.e., its turning angle, is given by the curvature κ of $\mathbf{X}(s)$ at $\mathbf{X}$. This curvature can be determined according to

$$\kappa = \left(d^2y/dx^2\right)/\left(1 + (dy/dx)^2\right)^{3/2} \tag{6}$$

by implicit differentiation of (4), namely,

$$g(\mathbf{X}(s)) - g(\mathbf{X}_0) = g(x, y) - g(x_0, y_0) = 0. \tag{4a}$$

The result for κ, multiplied by the gradient magnitude, yields

$$\kappa\sqrt{g_x^2 + g_y^2} = -\left(g_{xx}g_y^2 - 2g_{xy}g_xg_y + g_{yy}g_x^2\right)/\left(g_x^2 + g_y^2\right). \tag{6a}$$

Kitchen and Rosenfeld found that local maxima of this gradient weighted planar curvature isolated corners, especially if the heuristic of nonmaximum suppression along the gradient direction is applied to the gradient magnitude prior to its multiplication with the curvature κ.

2.3 Comparison of Both Approaches

To see the equivalence of both approaches explicitly, we shall discuss the locations where the gradient magnitude or (equivalently) the squared gradient magnitude takes on its maximum value.

$$\nabla\left(g_x^2 + g_y^2\right) = 0 \tag{7a}$$

or

$$\begin{pmatrix} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{pmatrix}\begin{pmatrix} g_x \\ g_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{7b}$$

At each point **X** we may align the coordinate axes with the main curvature directions, i.e., $g_{xy} = 0$. The requirement of vanishing derivatives for the expressions to be maximized then yields

$$g_{xx}g_x + g_{xy}g_y = 0 \Rightarrow \kappa_1 g_x = 0 \tag{7c}$$

and

$$g_{xy}g_x + g_{yy}g_y = 0 \Rightarrow \kappa_2 g_y = 0. \tag{7d}$$

There are four basic situations satisfying these conditions:

(a) Both principal curvatures vanish identically. This corresponds to a planar gray value variation, either constant as at point Q in Fig. 3 or a "sloping plane." There is no extreme of the gradient magnitude.

(b) Both principal curvatures are nonzero. Thus, both gradient components have to vanish simultaneously and we have a local minimum of the gradient magnitude as it occurs at gray value extrema, namely, "pits" and "peaks."

(c) Only one principal curvature vanishes identically, whereas the other one is different from zero or at most crosses zero at the point **X** considered. If this other main curvature crosses zero, the gradient component along its direction will take on an extremum, as at point Q_x (κ_1 crossing zero) or Q_y (κ_2 crossing zero) in Fig. 3. This corresponds to a "straight line" gray value transition.

(d) The last case is the one of most interest here. One principal curvature crosses zero (equivalent to an extremum of the gradient component along this direction), whereas the other principal curvature is nonzero, forcing the corresponding gradient component to vanish. This implies that the gradient is oriented along

the direction of the principal curvature which crosses zero, resulting in a local maximum of this gradient component and therefore the gradient magnitude. An example for this "corner" situation is the turning point T on the intersection curve $C3$ in Fig. 3.

When Kitchen and Rosenfeld apply their nonmaximum suppression heuristic to the gradient component perpendicular to a curve of constant gray value in the surface $g(x, y)$, they effectively select a curve connecting local turning points of the gray value slope such as the curve comprising the points Q_y, T, and Q_x in Fig. 3. Selecting the point with largest planar gradient-weighted curvature given by Eq. (6a) picks the corner point corresponding to T in Fig. 3. If the coordinate system is aligned with the principal curvature directions, Eq. (6a) yields

$$\kappa\sqrt{g_x^2 + g_y^2} \Rightarrow -\left(\kappa_1 g_y^2 + \kappa_2 g_x^2\right)/\left(g_x^2 + g_y^2\right). \tag{8a}$$

The nonmaximum suppression heuristic effectively enforces one principal curvature, say κ_2, to cross zero, thus resulting in

$$\kappa\sqrt{g_x^2 + g_y^2} \Rightarrow -\kappa_1 g_y^2/\left(g_x^2 + g_y^2\right) = \kappa_1 \tag{8b}$$

because the zero-crossings of κ_2 implies that the gradient exhibits a maximum and is aligned with this principal curvature direction. They then select the largest value for this expression, that is the point of steepest gray value slope where simultaneously the rate of change in the gradient orientation attains a maximum.

2.4 Curved Zero-Crossing Contours of the Laplacian

This discussion has another consequence. Zero-crossing of the Laplacian applied to a Gaussian-filtered gray value window have become prominent as descriptors to be used in computational theories of stereopsis (e.g., Marr and Hildreth [34], Grimson [14], Mayhew and Frisby [15]; see also Baker and Binford [35]). Since the displacement direction is assumed to be available in stereopsis, there is no need to distinguish between cases (c) and (d) above (i.e., between straight line edges and corners), provided they have a significant gradient component along the expected displacement direction.

In the analysis of unrestricted displacement vector fields, it is preferable to isolate and track corners in order to estimate both components of the displacement vector. Otherwise, supplementary assumptions are required in order to determine the missing component; see the discussion in the following section.

If corners are selected on zero-crossing lines as implied, for example, by the approach of Yam and Davis [36] (see also [25]), then it follows from eq. (3a) that

$$g_{xx} + g_{yy} = 0 = \kappa_1 + \kappa_2 \quad \text{or} \quad \kappa_1 = -\kappa_2. \tag{9}$$

If one principal curvature is different from zero on a bending zero-crossing contour, so must the other one. More or less sharp corners on zero-crossing contours of the Laplacian are therefore not indicative of steepest gray value slope, whereas this is the

case on straight line stretches with $\kappa_1 = -\kappa_2 = 0$. Potential consequences of this observation are being investigated [37].

The concepts for this corner detector in [5, 29, 32] have been discussed in detail because they form the basis for a method to determine both components of the displacement vector for such corner points in the following section. In practice, the required first and second derivatives of the gray value distribution are determined using the operators of Beaudet [38]. The derivation of these operators including closed formulas for quadratic windows with odd numbers of pixels along each side are presented in Appendix 1 in the manner used for subsequent discussions in this contribution.

3. FUNCTIONAL APPROXIMATION FOR THE ESTIMATION OF DISPLACEMENT VECTOR FIELDS

Starting from empirical considerations, Limb and Murphy [39, 40] investigated an approach to estimate the displacement vector component within a TV line. Cafforio and Rocca [41, 42] developed a cluster analysis approach to determine both components of the displacement vector, possibly for more than a single moving object image. Independently, Fukinuki *et al.* [43] as well as Fennema and Thompson [44] reported results obtained by a basically similar approach.

It is assumed that the image gray value can be approximated sufficiently well by a Taylor series (see Eq. (1)) without recourse to higher than linear terms:

$$g(\mathbf{X}) = g(\mathbf{X}_0) + g_x(x - x_0) + g_y(y - y_0). \tag{10}$$

Neglecting boundary effects, the temporal change of gray value inside an area is exclusively attributed to a shift $\mathbf{U} = (u, v)'$ of the sloping gray value plane, i.e.,

$$g_t = \partial g(\mathbf{X}_0)/\partial t = -(g_x u + g_y v) = -(\nabla g)' * \mathbf{U}. \tag{11}$$

The time derivative is approximated by the interframe gray value difference

$$g_t \cong Dg/Dt = \{g(\mathbf{X}, t2) - g(\mathbf{X}, t1)\}/Dt, \tag{12}$$

where Dt denotes the time interval between consecutive frames. Since Dt is usually constant, one may set it to unity and write for the ith area

$$Dg_i = -(\nabla g_i)' * \mathbf{U}. \tag{11a}$$

Since each area is characterized by a fixed value for the gray value gradient ∇g_i, this gives one equation relating the two unknown components u and v of the displacement vector $\mathbf{U}$ to the measured values of ∇g_i and Dg_i. As an additional relation, one may assume that one component is known (e.g., to be equal to zero as in the case of binocular stereo vision).

Both components of the displacement vector may be obtained by the assumption that $\mathbf{U}$ is constant across at least two different areas with linear independent values for the gradient ∇g_i. In this case, at least two independent equations (11a) are available to determine both u and v. A standard minimum squared error approach

may be used if the constancy of **U** is assumed for more than two areas [41, 43, 45] (see also [18, 24, 46, 47] for additional discussion).

In general, it will not be known which image areas experience the same displacement vector **U**. Therefore, Horn and Schunck [48] investigated a weaker assumption, namely, that the displacement vector **U** should vary smoothly with the image plane coordinates. They used a minimization approach to combine their requirement with Eq. (11a). As a result, they obtained two equations relating u and v to their average values in a local environment and to the measured values of Dg as well as g_x and g_y at each location **X**. Yachida [49] extended this approach by averaging the displacement estimates not only in a spatial neighborhood of the current frame, but in addition in the same neighborhoods in the preceding and the subsequent frames.

These approaches, however, do not account for the possibility that the displacement can change abruptly across occluding or shadow edges. Examples of this effect have been given in Dinse *et al.* [50]. Schunck and Horn [51] show in a careful discussion that Eq. (11a) remains valid across such edges. The smoothness assumption or even the stronger one of constant **U**, however, is no longer satisfied in such situations. Schunck and Horn [51] present heuristics based on probabilistic estimates by which they attempt to detect unrealistic displacement estimates which might possibly be due to occlusion or shadow edges.

A linear model of gray value variations is too simplistic around edges or corners. Snyder *et al.* [52] used the Taylor expansion of Eq. (1) to express the interframe gray value difference as a function of the displacement vector and the first- as well as second-order derivatives of the gray value distribution in one frame. This does not solve the problem, however, because they again obtain a single equation which is, moreover, nonlinear in the two unknown displacement vector components u and v. Prazdny [53] discussed the special case of pure sensor translation with already known focus of expansion. The interframe displacement vectors are then aligned with radii emanating from the focus of expansion. Since the direction of the displacement vector is known in this case, only its magnitude remains to be determined. Prazdny [53] suggested using the single second-order equation given by Snyder *et al.* [52] for the determination of this displacement vector length. This is a nonlinear analog of the work of Limb and Murphy [39, 40] where special knowledge reduces the number of unknown displacement components to one.

Webb [12] used a Taylor expansion according to Eq. (1) to estimate the coefficients of an affine mapping between local neighborhoods around image locations in different frames. He had to assume, however, that the proper displacement between the corresponding locations was known or could be deduced by an exhaustive search.

Neither of these approaches is suited to estimate the displacement vector using local information only. There are situations where this is possible—as demonstrated by interframe matching of isolated local descriptors; for example, using a nearest neighbor approach. The gray value corner discussed in Section 2 represents such a localized descriptor. The next section will show that a minimization approach can be formulated which exploits only local information to derive two coupled nonlinear equations for the two unknown displacement components u and v. In the special case of a gray value corner, these equations can be reduced to yield two equations. One of them is linear in one displacement vector component and no longer contains the other unknown component.

4. ESTIMATING THE DISPLACEMENT OF GRAY VALUE CORNERS

Assume that frame 1 of a sequence has been recorded at time $t1$ and frame 2 at time $t2$. $g(\mathbf{X}_0, t1)$ denotes the gray value observed at the location $\mathbf{X}_0$ in frame 1 and $g(\mathbf{X}_0, t2)$ the gray value observed at the same location in frame 2. A local environment around $\mathbf{X}_0$ is assumed to have been displaced between time $t1$ and $t2$ by a vector $\mathbf{U} = (u, v)'$. We want to determine $\mathbf{U}$ from the requirement that

$$MD = [g(\mathbf{X}, t2) - g(\mathbf{X} - \mathbf{U}, t1)]^2 \Rightarrow \text{minimum} \tag{13}$$

summed over all $\mathbf{X}$ from the local environment around $\mathbf{X}_0$. If the time interval $Dt = t2 - t1$ is small enough so that $\mathbf{U}$ remains well within the chosen environment around $\mathbf{X}_0$, we may use a Taylor series expansion according to Eq. (1) in order to represent $g(\mathbf{X} - \mathbf{U}, t1)$. The best fit approximation to the first- and second-order partial derivatives determined according to Appendix 1 will exhibit the frame time $t1$ in the notation

$$\mathbf{P}1 = (f1_0, f1_x, f1_y, f1_{xx}, f1_{xy}, f1_{yy})'. \tag{14}$$

The axes of the local coordinate system for the environment around $\mathbf{X}_0$ should be aligned with the principal curvature directions such that $f1_{xy} = 0$. Moreover, we initially neglect the measurement uncertainties and their effect on $\mathbf{P}1$.

The expression to be minimized by a suitable choice of the unknown displacement vector components u and v can now be written as

$$\begin{aligned} MD = \sum \big[& g(\mathbf{X}, t2) - f1_0 - f1_x(x-u) - f1_y(y-v) \\ & - \tfrac{1}{2} f1_{xx}(x-u)^2 - \tfrac{1}{2} f1_{yy}(y-v)^2 \big]^2. \end{aligned} \tag{15}$$

Setting to zero the partial derivatives with respect to u and v yields

$$\begin{aligned} 0 = \sum \big[f1_x + f1_{xx}(x-u) \big] \big[& g(\mathbf{X}, t2) - f1_0 - f1_x(x-u) - f1_y(y-v) \\ & - \tfrac{1}{2} f1_{xx}(x-u)^2 - \tfrac{1}{2} f1_{yy}(y-v)^2 \big] \end{aligned} \tag{16a}$$

$$\begin{aligned} 0 = \sum \big[f1_y + f1_{yy}(y-v) \big] \big[& g(\mathbf{X}, t2) - f1_0 - f1_x(x-u) - f1_y(y-v) \\ & - \tfrac{1}{2} f1_{xx}(x-u)^2 - \tfrac{1}{2} f1_{yy}(y-v)^2 \big]. \end{aligned} \tag{16b}$$

In analogy to Eq. (A12) we define

$$Sk_{pqr} = \sum_{ij} x_i^p y_j^q g(x_i, y_j, t_k)^r \qquad \text{for} \quad k = 1, 2, \tag{17}$$

but still use definition (A12) without time index if $r = 0$ since Sk_{pq0} does not depend on the frame time. Evaluating the sums and exploiting the symmetry of the

environment around $\mathbf{X}_0 = 0$ yields

$$0 = [fl_x - fl_{xx}u]\big[S2_{001} - fl_0 S_{000} - \tfrac{1}{2}fl_{xx}S_{200} - \tfrac{1}{2}fl_{yy}S_{020} + S_{000}u\big(fl_x - \tfrac{1}{2}fl_{xx}u\big) + S_{000}v\big(fl_y - \tfrac{1}{2}fl_{yy}v\big)\big] + \big[S2_{101} - (fl_x - fl_{xx}u)S_{200}\big]fl_{xx} \tag{18a}$$

$$0 = \big[fl_y - fl_{yy}v\big]\big[S2_{001} - fl_0 S_{000} - \tfrac{1}{2}fl_{xx}S_{200} - \tfrac{1}{2}fl_{yy}S_{020} + S_{000}u\big(fl_x - \tfrac{1}{2}fl_{xx}u\big) + S_{000}v\big(fl_y - \tfrac{1}{2}fl_{yy}v\big)\big] + \big[S2_{011} - (fl_y - fl_{yy}v)S_{020}\big]fl_{yy}. \tag{18b}$$

Using Eqs. (A19) together with (A15) it can be shown that

$$fl_0 S_{000} + \tfrac{1}{2}fl_{xx}S_{200} + \tfrac{1}{2}fl_{yy}S_{020} = S1_{001}. \tag{19}$$

With

$$\overline{g1} = \overline{g(\mathbf{X}, t1)} = S1_{001}/S_{000} \tag{20a}$$

$$\overline{g2} = \overline{g(\mathbf{X}, t2)} = S2_{001}/S_{000} \tag{20b}$$

and $\overline{x^2} = S_{200}/S_{000}$ as well as $\overline{y^2} = S_{020}/S_{000}$ we can write (see (A19))

$$0 = [fl_x - fl_{xx}u]\big[\overline{g2} - \overline{g1} + u\big(fl_x - \tfrac{1}{2}fl_{xx}u\big) + v\big(fl_y - \tfrac{1}{2}fl_{yy}v\big)\big] + \overline{x^2}fl_{xx}\big[f2_x - (fl_x - fl_{xx}u)\big] \tag{21a}$$

$$0 = \big[fl_y - fl_{yy}v\big]\big[\overline{g2} - \overline{g1} + u\big(fl_x - \tfrac{1}{2}fl_{xx}u\big) + v\big(fl_y - \tfrac{1}{2}fl_{yy}v\big)\big] + \overline{y^2}fl_{yy}\big[f2_y - (fl_y - fl_{yy}v)\big]. \tag{21b}$$

If the second partial derivatives can be neglected, both Eqs. (21) reduce to the linear version (11a) provided the first derivatives are different from zero.

A significant second partial derivative exerts its influence through two effects. First, the gradient will not be constant within the environment and, therefore, the gradient components fl_x or fl_y determined at the center of the environment must be corrected in order to obtain the gradient components at the off-center position given by U.

The factors $\overline{x^2}fl_{xx}$ or $\overline{y^2}fl_{yy}$, respectively, in front of the second term of Eq. (21) take into account the nonvanishing average gray value deviation from the center value, caused by the curvature of the gray value distribution. Corresponding terms can be recognized in Eq. (19). If $fl_x - fl_{xx}u$ is different from zero, we may write Eq. (21a) in the form

$$\overline{g2} - \overline{g1} + u\big(fl_x - \tfrac{1}{2}fl_{xx}u\big) + v\big(fl_y - \tfrac{1}{2}fl_{yy}v\big) + \overline{x^2}fl_{xx}(f2_x - (fl_x - fl_{xx}u))/(fl_x - fl_{xx}u) = 0. \tag{22}$$

The second term of Eq. (21a) thus represents the relative difference between the gradients at $\mathbf{X}_0$ in frame 2 and $\mathbf{X}_0 - \mathbf{U}$ in frame 1, multiplied by the average gray value deviation in the x direction caused by the curvature of the gray value distribution. We may write Eq. (22) in the form

$$\begin{aligned}\overline{g2} - \overline{g1} + \overline{x^2} f1_{xx}(f2_x - (f1_x - f1_{xx}u))/(f1_x - f1_{xx}u) \\ = -u(f1_x - \tfrac{1}{2}f1_{xx}u) - v(f1_y - \tfrac{1}{2}f1_{yy}v),\end{aligned} \tag{22a}$$

i.e., this term modifies the average temporal gray value difference $\overline{g2} - \overline{g1}$ which approximates the partial derivative of the gray value with respect to time. Equation (22a) thus can be looked at as a generalization of Eq. (11).

Now let $\mathbf{X}_0$ denote the position of a gray value corner in frame 1 as discussed in Section 2. The axes of the local coordinate system should be aligned with the principal curvature directions, i.e., $f1_{xy} = 0$. The gradient passes through a maximum at $\mathbf{X}_0$, i.e.,

$$f1_x = \text{extremum} \qquad \text{and} \qquad f1_y = 0 \tag{23a}$$

and

$$f1_{xx} = 0 \qquad \text{with} \qquad f1_{yy} = \text{extremum} \neq 0. \tag{23b}$$

Since $f1_y$ and $f1_{xx}$ will rarely vanish exactly at a raster position, the following somewhat weaker conditions will allow the same simplification of Eqs. (21) as the exact corner conditions (23a) + (23b):

$$f1_{xx}u \ll f1_x \qquad \text{and} \qquad f1_y \ll f1_{yy}v \tag{23c}$$

for all relevant values of u and v. With (23) we obtain from (21)

$$0 = f1_x\left[\overline{g2} - \overline{g1} + uf1_x - \tfrac{1}{2}f1_{yy}v^2\right] + \overline{x^2}f1_{xx}[f2_x - f1_x] \tag{24a}$$

$$0 = -f1_{yy}v\left[\overline{g2} - \overline{g1} + uf1_x - \tfrac{1}{2}f1_{yy}v^2\right] + \overline{y^2}f1_{yy}[f2_y + f1_{yy}v]. \tag{24b}$$

Equation (24a) is linear in u and can be employed to eliminate the factor containing the square of v in Eq. (24b). This turns Eq. (24b) into an equation linear in the remaining unknown v. Assuming a square neighborhood so that $\overline{x^2} = \overline{y^2}$, we obtain the result

$$v = -f2_y/\left(f1_{yy} + f1_{xx}(f2_x/f1_x - 1)\right). \tag{25a}$$

This result can now be inserted to compute

$$u = -\frac{1}{f1_x}\left\{\overline{g2} - \overline{g1} - \tfrac{1}{2}f1_{yy}v^2 + \overline{x^2}f1_{xx}(f2_x/f1_x - 1)\right\}. \tag{25b}$$

The more stringent requirements (23a) and (23b) result in the simple expressions

$$v = -f2_y/f1_{yy} \tag{26a}$$

$$u = -\left(\overline{g2} - \overline{g1} - \tfrac{1}{2}f1_{yy}v^2\right)/f1_x. \tag{26b}$$

These expressions emphasize in an especially clear manner the importance of the two main characteristics for "gray value corners": the gradient component $f1_x$ along one principal curvature direction and the second principal curvature $f1_{yy}$ (i.e., perpendicular to the gradient) must attain local maxima which implies that they are both different from zero.

Experience has shown that the gray value variations characteristic for a "corner" are well localized so that a 5×5 environment may be already too large to satisfy conditions (23). Corner point localization has been improved by a two-stage approach, using 3×3 operators for precision positioning [29, 5]. Such small environments may be insufficient, however, to enclose the characteristic gray value variation in the second frame. It therefore seems advantageous to localize a gray value corner in frame 1 based on 3×3 operators, but to use 5×5 operators to determine the parameters required for Eqs. (25) based on (23c). The results obtained in this manner may then be considered as suitable starting values for an iterative improvement of the displacement vector estimation, as discussed in the next section.

5. ITERATIVE IMPROVEMENT OF THE DISPLACEMENT VECTOR ESTIMATE

We assume that an initial estimate $\mathbf{U}_0$ for the displacement vector around $\mathbf{X}_0$ has been obtained. Moreover, the gray value distribution $g(\mathbf{X} - \mathbf{U}_0, t1)$ should be adequately represented within a neighborhood around $\mathbf{X}_0$ by the best fit approximation according to Appendix 1:

$$g(\mathbf{X} - \mathbf{U}_0, t1) = f1(\mathbf{X} - \mathbf{U}_0; \mathbf{P}1). \tag{27}$$

We now want to determine a correction $\mathbf{DU}$ for the displacement estimate $\mathbf{U}_0$ by requiring that the weighted sum of squared differences

$$dg = g(\mathbf{X}, t2) - f1(\mathbf{X} - \mathbf{U}_0 - \mathbf{DU}; \mathbf{P}1) \tag{28}$$

should be minimized by a suitable choice of $\mathbf{DU}$:

$$MG = (\mathbf{dG})' W_G (\mathbf{dG}) \Rightarrow \text{minimum}. \tag{29}$$

Here $\mathbf{dG}$ represents the column vector of differences according to Eq. (28), where the implicit variation of $\mathbf{X}$ within the $(2k + 1) * (2k + 1)$ neighborhood around $\mathbf{X}_0$ is made explicit by introduction of raster locations $\mathbf{X}_i = (x_i, y_i)'$ defined in Eq. (A2). The $N * N$ weight matrix W_G is the inverse of a covariance matrix which reflects both the measurement uncertainties of $g(\mathbf{X}, t2)$ as well as those of $g(\mathbf{X}, t1)$ propagated forward to $f1(\mathbf{X}; \mathbf{P}1)$. The covariance matrix for $g(\mathbf{X}, t2)$ is given by the inverse of W introduced in Eq. (A6). The covariance matrix for $f_1(\mathbf{X}; \mathbf{P}1)$ has been derived in Appendix 1. Its elements are given in Eq. (A20). We assume that the measurement errors for $g(\mathbf{X}, t2)$ and $f1(\mathbf{X}; \mathbf{P}1)$ are uncorrelated. The joint covariance matrix for the differences dg is then given by the sum of the covariance matrices, i.e.,

$$\left(W_G^{-1}\right)_{ij} = \sigma^2 \delta_{ij} + \left(W_F^{-1}\right)_{ij}. \tag{30}$$

The weight matrix required in Eq. (29) can be obtained by inversion of the matrix defined in Eq. (30). This inversion can be considerably simplified if the symmetries of W_G are exploited as discussed in Appendix 2.

If $\mathbf{U}_0$ is close to the true displacement vector minimizing MG, **DU** will be small and higher than linear terms in the components (Du, Dv) of **DU** can be neglected. Using this approximation yields the following expression for the partial derivatives of dg with respect to Du and Dv:

$$B_{i,Du} = \partial\, dg(x_i, y_i)/\partial\, Du = fl_x + fl_{xx}(x_i - u_0) + fl_{xy}(y_i - v_0) \quad (31a)$$

$$B_{i,Dv} = \partial\, dg(x_i, y_i)/\partial\, Dv = fl_y + fl_{xy}(x_i - u_0) + fl_{yy}(y_i - v_0) \quad (31b)$$

for $i = 1, \ldots, N$. In this approximation, equating the partial derivatives of Eq. (29) to zero yields

$$\partial MG/\partial\, DU = 0 = B'W_G\{g(\mathbf{X}, t2) - fl(\mathbf{X} - \mathbf{U}_0; \mathbf{P}1) + B\,\mathbf{DU}\} \quad (32)$$

with the solution

$$\mathbf{DU} = -(B'W_GB)^{-1}B'W_B\{g(x_i, y_i, t2) - fl(x_i - u_0, y_i - v_0; \mathbf{P}1)\}. \quad (33)$$

The term in braces on the right-hand side of Eq. (33) represents a column vector with the row index i ($= 1, \ldots, N$) specifying the raster point location according to Eq. (A2). It should be noted that the components of the $N * 2$ matrix B depend on $\mathbf{U}_0$ as given by Eqs. (31). The corrected displacement vector is given by

$$\mathbf{U} = \mathbf{U}_0 + \mathbf{DU}. \quad (34)$$

The covariance matrix for **DU** and thus **U** is given by (cf. Appendix 1, Eq. (A17))

$$W_U^{-1} = (B'W_GB)^{-1}. \quad (35)$$

Another iteration can be attempted, this time using the result for **U** from Eq. (34) as the new initial value, until the square of the corrections Du and Dv become comparable to the corresponding diagonal elements of W_U^{-1}.

6. DISCUSSION OF THE ITERATIVE CORRECTION FORMULA FOR THE DISPLACEMENT ESTIMATE

It is instructive to evaluate Eq. (33) for the example of a 3×3 environment as depicted in Fig. A1 with $N = (2k + 1)^2 = 9$. The required matrix W_G is given explicitly in (A27). We introduce an abbreviation to denote the best fit gradient components corrected for the displacement $\mathbf{U}_0$:

$$hl_x = fl_x - fl_{xx}u_0 - fl_{xy}v_0 \quad (36a)$$

$$hl_y = fl_y - fl_{xy}u_0 - fl_{yy}v_0. \quad (36b)$$

Using this notation, Eqs. (31) can be written

$$B_{i,Du} = hl_x + fl_{xx}x_i + fl_{xy}y_i \quad (37a)$$

$$B_{i,Dv} = hl_y + fl_{xy}x_i + fl_{yy}y_i. \quad (37b)$$

Due to the internal symmetries of W_G, one can express $(B'W_G)$ in the simple form

$$(B'W_G)_{Du,i} = (1/2\sigma^2)(hl_x + fl_{xx}x_i + fl_{xy}y_i) \tag{38a}$$

$$(B'W_G)_{Dv,i} = (1/2\sigma^2)(hl_y + fl_{xy}x_i + fl_{yy}y_i). \tag{38b}$$

A straightforward calculation will yield the intermediate results

$$\begin{aligned}\det(B'W_GB) = (N/2\sigma^2)^2\Big\{&\overline{x^2y^2}\left(fl_{xx}fl_{yy} - fl_{xy}^2\right)^2 \\ &+\overline{x^2}\left[v_0\left(fl_{xx}fl_{yy} - fl_{xy}^2\right) - \left(fl_yfl_{xx} - fl_xfl_{xy}\right)\right]^2 \\ &+\overline{y^2}\left[u_0\left(fl_{xx}fl_{yy} - fl_{xy}^2\right) + \left(fl_yfl_{xy} - fl_xfl_{yy}\right)\right]^2\Big\}\end{aligned} \tag{39}$$

and

$$\begin{aligned}(B'W_GB)^{-1} = \frac{(N/2\sigma^2)}{\det(B'W_GB)}\Bigg\{&\begin{pmatrix} hl_y^2 & -hl_xhl_y \\ -hl_xhl_y & hl_x^2 \end{pmatrix} \\ &+\begin{pmatrix} fl_{yy} & -fl_{xy} \\ -fl_{xy} & fl_{xx} \end{pmatrix}\begin{pmatrix} \overline{y^2} & 0 \\ 0 & \overline{x^2} \end{pmatrix}\begin{pmatrix} fl_{yy} & -fl_{xy} \\ -fl_{xy} & fl_{xx} \end{pmatrix}\Bigg\}.\end{aligned} \tag{40}$$

Similar computations result in the following representation for the two-component vector:

$$\begin{aligned}&B'W_G\{g(\mathbf{X}, t2) - fl(\mathbf{X} - \mathbf{U}_0; \mathbf{P}1)\} \\ &= (N/2\sigma^2)\Bigg\{\left[\overline{g2} - \overline{g1} + fl_xu_0 + fl_yv_0 - \tfrac{1}{2}(u_0\,v_0)\begin{pmatrix} fl_{xx} & fl_{xy} \\ fl_{xy} & fl_{yy} \end{pmatrix}\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}\right]\begin{pmatrix} hl_x \\ hl_y \end{pmatrix} \\ &\quad+\begin{pmatrix} fl_{xx} & fl_{xy} \\ fl_{xy} & fl_{yy} \end{pmatrix}\begin{pmatrix} \overline{x^2} & 0 \\ 0 & \overline{y^2} \end{pmatrix}\begin{pmatrix} f2_x - hl_x \\ f2_y - hl_y \end{pmatrix}\Bigg\}.\end{aligned} \tag{41}$$

The desired correction $\mathbf{DU}$ for the displacement estimate $\mathbf{U}_0$ can be written in the form

$$\begin{aligned}\mathbf{DU} = \begin{pmatrix} Du \\ Dv \end{pmatrix} = -\frac{(N/2\sigma^2)^2}{\det(B'W_GB)}\Bigg\{&\left[\overline{g2} - \overline{g1} + fl_xu_0 + fl_yv_0 - \frac{1}{2}\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}'\begin{pmatrix} fl_{xx} & fl_{xy} \\ fl_{xy} & fl_{yy} \end{pmatrix}\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}\right] \\ &\times\begin{pmatrix} fl_{yy} & -fl_{xy} \\ -fl_{xy} & fl_{xx} \end{pmatrix}\begin{pmatrix} \overline{y^2} & 0 \\ 0 & \overline{x^2} \end{pmatrix}\begin{pmatrix} fl_{yy} & -fl_{xy} \\ -fl_{xy} & fl_{xx} \end{pmatrix}\begin{pmatrix} hl_x \\ hl_y \end{pmatrix} \\ &+\begin{pmatrix} hl_y^2 & -hl_xhl_y \\ -hl_xhl_y & hl_x^2 \end{pmatrix}\begin{pmatrix} fl_{xx} & fl_{xy} \\ fl_{xy} & fl_{yy} \end{pmatrix}\begin{pmatrix} \overline{x^2} & 0 \\ 0 & \overline{y^2} \end{pmatrix}\begin{pmatrix} f2_x - hl_x \\ f2_y - hl_y \end{pmatrix} \\ &+\overline{x^2y^2}\left(fl_{xx}fl_{yy} - fl_{xy}^2\right)\begin{pmatrix} fl_{yy} & -fl_{xy} \\ -fl_{xy} & fl_{xx} \end{pmatrix}\begin{pmatrix} f2_x - hl_x \\ f2_y - hl_y \end{pmatrix}\Bigg\}.\end{aligned} \tag{42}$$

This expression for $\mathbf{DU} = (Du, Dv)'$ will now be specialized to the situation of a gray value corner as discussed in the preceding sections, including the alignment of the local coordinate axes with the principal curvature directions:

$$f1_{xy} = 0, \quad f1_x = \text{extremum} \neq 0; \quad f1_y = 0,$$
$$f1_{xx} = 0, \quad f1_{yy} = \text{extremum} \neq 0.$$

The Gaussian curvature vanishes in this case

$$f1_{xx} f1_{yy} - f1_{xy}^2 = 0$$

and the expression for the determinant in Eq. (39) reduces to

$$\det(B' W_G B) = (N/2\sigma^2)^2 \overline{y^2} f1_x^2 f1_{yy}^2 \tag{43}$$

and the expression for $\mathbf{DU}$ in Eq. (42) to

$$\mathbf{DU} = \begin{pmatrix} Du \\ Dv \end{pmatrix} = -\frac{1}{\overline{y^2} f1_x^2 f1_{yy}^2} \left\{ \left[\overline{g2} - \overline{g1} + f1_x u_0 - \tfrac{1}{2} f1_{yy} v_0^2 \right] \begin{pmatrix} f1_{yy}^2 \overline{y^2} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h1_x \\ h1_y \end{pmatrix} \right.$$
$$\left. + \overline{y^2} f1_{yy} \begin{pmatrix} 0 & -h1_x h1_y \\ 0 & h1_x^2 \end{pmatrix} \begin{pmatrix} f2_x - h1_x \\ f2_y - h1_y \end{pmatrix} \right\}. \tag{44}$$

Under these conditions, Eqs. (36) reduce to

$$h1_x = f1_x \quad \text{and} \quad h1_y = -f1_{yy} v_0 \tag{45}$$

and we obtain for the components of $\mathbf{DU}$,

$$Du = -(1/f1_x)\{\overline{g2} - \overline{g1} + f1_x u_0 - \tfrac{1}{2} f1_{yy} v_0^2 + (f2_y + f1_{yy} v_0) v_0\} \tag{46a}$$
$$Dv = -(1/f1_{yy})\{f2_y + f1_{yy} v_0\}. \tag{46b}$$

If we use an arbitrary value v_0 in order to compute v, we see that v_0 drops out from the result:

$$v = v_0 + Dv = v_0 - \{(f2_y/f1_{yy}) + v_0\} = -f2_y/f1_{yy}. \tag{47a}$$

In the case of Eq. (46a), a bit more care is required. We exploit the assumption that the correction Dv is small compared to the approximately correct value v_0. Then we can write

$$v^2 = (v_0 + Dv)^2 \approx v_0^2 + 2 v_0 Dv$$

and together with $(f2_y + f1_{yy} v_0) = -f1_{yy} Dv$ based on Eq. (46b),

$$u = u_0 + Du = -(1/f1_x)\{\overline{g2} - \overline{g1} - \tfrac{1}{2} f1_{yy} v^2\}. \tag{47b}$$

The initial value u_0 drops out from the result for u, just as in the case of v. In this form, Eqs. (47) are identical to Eqs. (26) derived earlier.

The correction **DU** for the displacement estimate will become unreliable if $\det(B'W_G B)$ threatens to vanish; see Eq. (42). It is therefore desirable to derive a threshold for $\det(B'W_G B)$. If its value turns out to be comparable to or smaller than this threshold, the results from Eq. (33) would be largely determined by noise.

To simplify the discussion, we assume that the coordinate system is aligned with the principal curvature directions, i.e., $fl_{xy} = 0$. Then we can find two situations where the determinant given by Eq. (39) will vanish:

(i) both principal curvatures vanish,

(ii) only one principal curvature, say fl_{xx}, and the gradient component fl_x along this principal curvature direction vanish.

The first case corresponds to a locally planar gray value variation and the latter one to a locally straight line edge, see the discussion in Section 2.3. If the Gaussian curvature represented here by $(fl_{xx}fl_{yy} - fl_{xy}^2)$ is different from zero, the two last terms in Eq. (39) may nevertheless vanish for suitable values of u_0 and v_0. We concentrate, therefore, on the first term in the expression for $\det(B'W_G B)$ and suppress the frame index. Let

$$K = f_{xx}f_{yy} - f_{xy}^2 \tag{48}$$

represent the estimate for the Gaussian curvature. Using propagation of small errors, we may write for the variance of K

$$(\Delta K)^2 = \left(\frac{\partial K}{\partial f_{xx}}\right)^2 (\Delta f_{xx})^2 + \left(\frac{\partial K}{\partial f_{yy}}\right)^2 (\Delta f_{yy})^2 + \left(\frac{\partial K}{\partial f_{xy}}\right)^2 (\Delta f_{xy})^2 \tag{49a}$$

$$= (f_{yy})^2(\Delta f_{xx})^2 + (f_{xx})^2(\Delta f_{yy})^2 + (2f_{xy})^2(\Delta f_{xy})^2. \tag{49b}$$

The variances on the right-hand side of Eq. (49) are given by the corresponding diagonal elements in the covariance matrix (A17) for the parameters appearing in Eq. (48) defining K. Since we are interested in a threshold value, we replace the squared parameters by their variances which represent a reasonable lower limit for them. With

$$(\Delta f_{xx})^2 = (\Delta f_{yy})^2 = 4\sigma^2/(S_{400} - S_{220}) \tag{50a}$$

$$(\Delta f_{xy})^2 = \sigma^2/S_{220} \tag{50b}$$

we obtain

$$(\Delta K)^2 = 2\left(4\sigma^2/(S_{400} - S_{200})\right)^2 + 4\left(\sigma^2/S_{220}\right)^2. \tag{51}$$

Since the determinant contains the square of K, we write (neglecting the second and third term)

$$(\Delta \det(B'W_G B))^2 = (\partial \det(B'W_G B)/\partial K)^2 (\Delta K)^2$$

$$= \left[(N/2\sigma^2)^2 \overline{x^2 y^2} 2K\right]^2 (\Delta K)^2. \tag{52}$$

To get a reasonable threshold, we replace K in Eq. (52) by the square root of its variance which indicates our uncertainty about the exact value of K. Taking the square root yields a lower bound to the standard deviation for the determinant

$$\Delta \det(B'W_G B) = (N/2\sigma^2)^2 \overline{x^2 y^2} 2(\Delta K)^2. \tag{53}$$

In the case of a square environment with $N = (2K + 1)^2 = S_{000}$ pixels, we have $\overline{x^2} = \overline{y^2} = S_{200}/S_{000}$. We thus obtain, by using Eq. (51),

$$\Delta \det(B'W_G B) = S_{200}^2 \{[4/(S_{400} - S_{220})]^2 + 2(1/S_{220})^2\}. \tag{54}$$

The entities appearing on the right-hand side of Eq. (54) can be evaluated using their closed form representation given in Eqs. (A14). In the special case of a 3×3 environment (i.e., $k = 1$), we obtain

$$\Delta \det(B'W_G B) = 9(16 + \tfrac{1}{8}) \cong 144.$$

If the determinant given by Eq. (39) becomes smaller than the threshold given by Eq. (54), the environment does not exhibit a gray value variation which is suitable to reliably estimate a displacement vector.

7. FROM ISOLATED DISPLACEMENT ESTIMATES TO DISPLACEMENT VECTOR FIELDS

The formalism for iterative refinement of a displacement estimate developed in the preceding sections can be employed, too, to estimate displacement vectors for image locations around a gray value corner. The basic idea is to use the final displacement estimate obtained at a gray value corner as an initial value for image points in its neighborhood and to refine it iteratively based on Eq. (42).

Such an approach raises two questions:

(1) Where does it have to stop?

(2) Will the result for other locations be independent of the initial value at the gray value corner?

The answer to the first question can be found in the discussion at the end of the preceding section: the propagation must be terminated when the gray value variation in the environment around another point is insufficient to yield a determinant according to Eq. (39) which exceeds the threshold given by Eq. (53).

The second question is more problematical. Gray value corners may occur on occluding boundaries. Their displacement will be representative for the moving occluding object but not for the background which is occluded or uncovered by the displacement of the occluding boundary. The displacement estimate obtained at a gray value corner may therefore be inappropriate as a starting value for part of the environment. This problem could be controlled by propagating the displacement estimate from each gray value corner until either no further propagation is possible due to the threshold given by Eq. (53) or until conflicting estimates are encountered. It should be noted that the approach developed here uses basically local information. The initial value employed in the iterative procedure described in the preceding sections could be derived in principle by random search as is often done in complex

minimization tasks. Propagation from neighboring pixels—based on the assumption of "almost always" smooth displacement variations—could be considered merely as a way to reduce the time spent in search.

This situation has to be distinguished from that encountered in other "nonmatching" approaches discussed in Section 3 (e.g., the approach of Horn and Schunck [48] or that discussed by Marr and Ullman [33]). The restriction to linear models of gray value variation enforces additional nonlocal assumptions such as smooth variation of displacement vectors in order to compensate for the fact that a linear model does not provide enough information for the local determination of both displacement vector components.

This analysis suggests a solution to overcome both problems indicated by the two questions from which this discussion started. Smooth displacement variations should only be postulated for image regions which do not exhibit sufficient structure to estimate both components of the displacement vector. The minimization approach introduced in Section 4 should be supplemented by the requirement that the change of displacement components weighted by the gray value curvature has to be minimized simultaneously.

$$\begin{aligned} MT = \int\int dx\,dy \Big\{ & [g(\mathbf{X}, t2) - g(\mathbf{X} - \mathbf{U}, t1)]^2 \\ & + \beta^2 (u_x\ u_y) \begin{pmatrix} \gamma + g_{xx} & g_{xy} \\ g_{xy} & \gamma + g_{yy} \end{pmatrix}^{-1} \begin{pmatrix} u_x \\ u_y \end{pmatrix} \\ & + \beta^2 (v_x\ v_y) \begin{pmatrix} \gamma + g_{xx} & g_{xy} \\ g_{xy} & \gamma + g_{yy} \end{pmatrix}^{-1} \begin{pmatrix} v_x \\ v_y \end{pmatrix} \Big\} \end{aligned} \tag{55}$$

with

$$u_x = \partial u/\partial x, \qquad u_y = \partial u/\partial y, \qquad v_x = \partial v/\partial x, \qquad v_y = \partial v/\partial y. \tag{56}$$

The integration has to be extended over the entire image. Vector $\mathbf{U}$ is considered to be a function of the image location $\mathbf{X}$, i.e., $\mathbf{U} = (u, v)' = \mathbf{U}(\mathbf{X})$. The first term on the right-hand side of Eq. (55) corresponds to the requirement that the gray value at location $\mathbf{X}$ in frame 2 is optimally approximated by the gray value at location $\mathbf{X} - \mathbf{U}$ in frame 1. A Taylor expansion of $g(\mathbf{X}, t1)$ up the second order is used to make the dependence of $g(\mathbf{X} - \mathbf{U}, t1)$ on $\mathbf{U}$ explicit in analogy to Eq. (15).

The second and third terms on the right-hand side of Eq. (55) express the demand that larger changes in the displacement vector can be admitted only where the gray value distribution exhibits sufficient variation. This can be seen more easily if we assume that the coordinate system is aligned with the principal curvature directions of $g(\mathbf{X}, t1)$ at $\mathbf{X}$, i.e., $g_{xy} = 0$. The second term may then be written in the form

$$(u_x\ v_y) \begin{pmatrix} \gamma + g_{xx} & 0 \\ 0 & \gamma + g_{yy} \end{pmatrix}^{-1} \begin{pmatrix} u_x \\ u_y \end{pmatrix} = \frac{u_x^2}{\gamma + g_{xx}} + \frac{u_y^2}{\gamma + g_{yy}}. \tag{57}$$

The constant γ has been introduced to prevent the determinant from becoming

singular in homogeneous gray value regions. A reasonable choice for γ will be the square root of the standard deviation for the Gaussian curvature, i.e., the fourth root of Eq. (51).

If the gray value curvature g_{xx} in the x direction becomes large at location $\mathbf{X}$, larger values for the partial derivative of u along the x-direction will nevertheless not contribute much to the expression (55) which has to be minimized, i.e., large values of g_{xx} diminish the constraint on variation of u_x along the x direction. Analogous considerations can be made for u_y and the partial derivatives of v.

In the limit of vanishing gray value curvature Eq. (55) specializes to the expression introduced by Horn and Schunck [48]. Let $g_{xx} = g_{xy} = g_{yy} = 0$ so that $g(\mathbf{X} - \mathbf{U}, t1) = g(\mathbf{X}, t1) - \mathbf{U} * \nabla g(\mathbf{X}, t1)$. With $g_t \Delta t = (\partial g/\partial t)\Delta t \approx g(\mathbf{X}, t2) - g(\mathbf{X}, t1)$ for $\Delta t = t2 - t1$ taken to be the unit time interval we obtain with $\alpha^2 = \beta^2/\gamma$

$$MT = \int\int dx\, dy \left([g_t + g_x u + g_y v]^2 + \alpha^2 \left(u_x^2 + u_y^2 + v_x^2 + v_y^2 \right) \right). \tag{58}$$

The approach represented by Eq. (55) can therefore be considered to be a generalization of both the approaches of Horn and Schunck suitable for image regions with slowly varying gray values and the approach outlined in the preceding sections for gray value corners.

If we have a straight line gray value transition (e.g., only $g_{yy} = g_{xy} = 0$) the displacement vectors should vary smoothly along the y direction, but may exhibit considerable changes along the x-direction, i.e., perpendicular to the edge. This case corresponds to the situations where displacement estimates are derived from moving edges (see, e.g., Korn and Kories [54], Kories [55], Thompson [56], or Haynes and Jain [57]).

8. DISCUSSION

Some of the theoretical developments discussed in this paper have been implemented. Preliminary results obtained on the basis of Eqs. (25) and (26) have been reported by Nagel and Enkelmann [58]. Figures 4a and b illustrate displacement estimates obtained on the basis of two half frames from a TV frame sequence. These two half-frames were taken at a 40-msec interval. Figures 5a and b illustrate the estimates obtained from the first and second half frame, i.e., 20 msec apart, of the same TV frame. The gray value corners are those selected by the (partially heuristic) algorithm reported by Dreschler [29] and Dreschler and Nagel [5]. Closer inspection has shown that the conditions for the use of Eq. (25) or (26) are not met at image points where unsatisfactory estimates can be seen. This observation has encouraged investigations to improve the selection of gray value corners, using the more stringent requirements (23). A more thorough investigation of these questions including results derived by iterative refinement and an evaluation of Eq. (55) will be reported in the future.

The approach developed in this paper appears to have a conceptual significance beyond its practical application. It provides a mathematical basis for a computation which may determine displacement vectors based on local information. It thus appears possible to determine optical flow without some of the more global assumptions required so far. This opens the possibility to determine displacement vectors even in the vicinity of optical flow discontinuities which are important clues

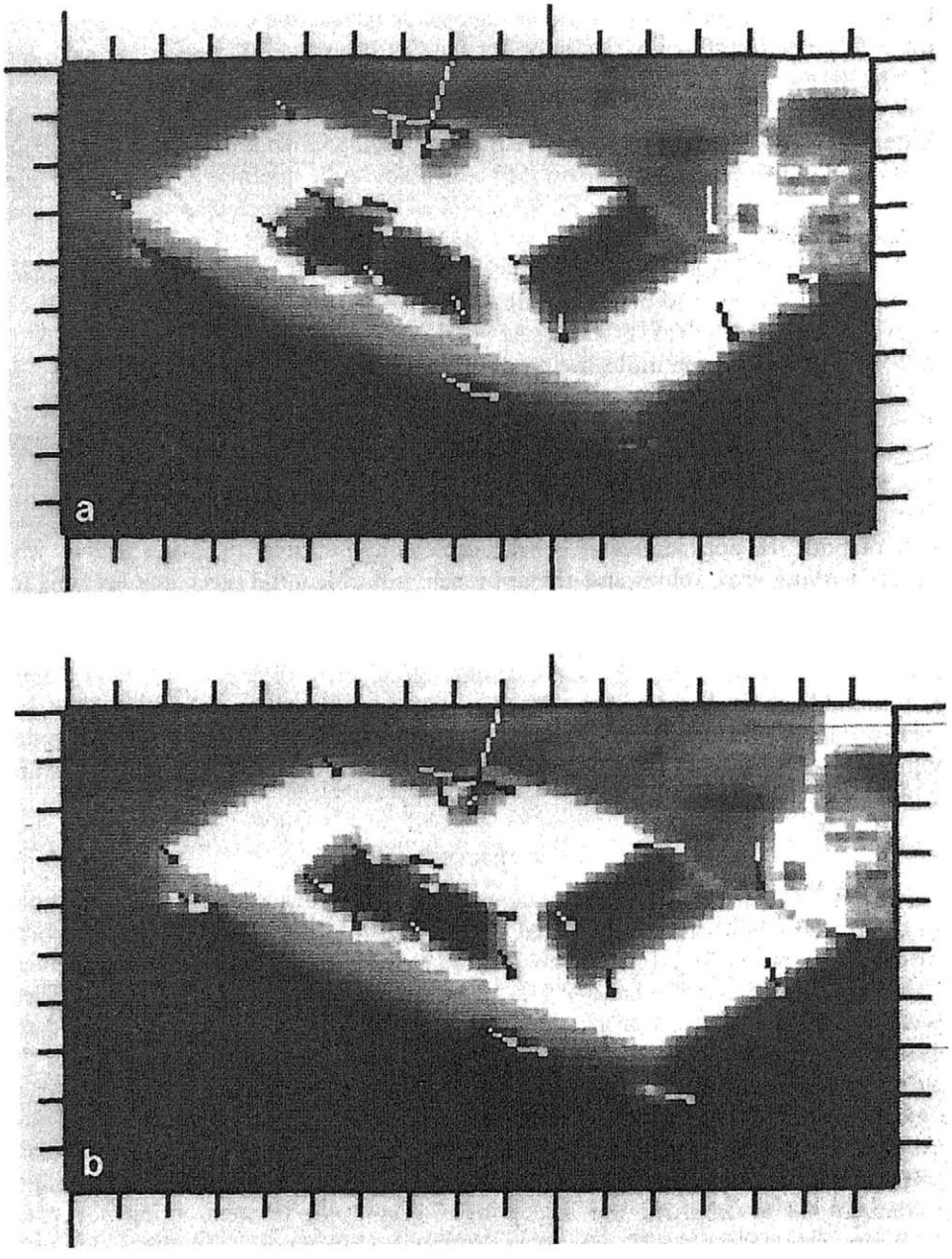

FIG. 4. Enlarged section around the white taxicab from Fig. 1 with displacement vectors estimated according to Eq. (26). The "gray value corners" have been selected by the algorithm of Dreschler [29] and Dreschler and Nagel [5]. (a) The first half frame has been compared with (b) the first half frame of the subsequent TV frame, recorded 40 msec later. Detailed inspection has shown that the characteristics assumed for "gray value corners" are not fully met by those points for which unsatisfactory estimates are shown. This insight has fostered investigations to improve the corner point selection in the program of Dreschler and Nagel. It led, moreover, to the iterative refinement procedure as discussed in the text.

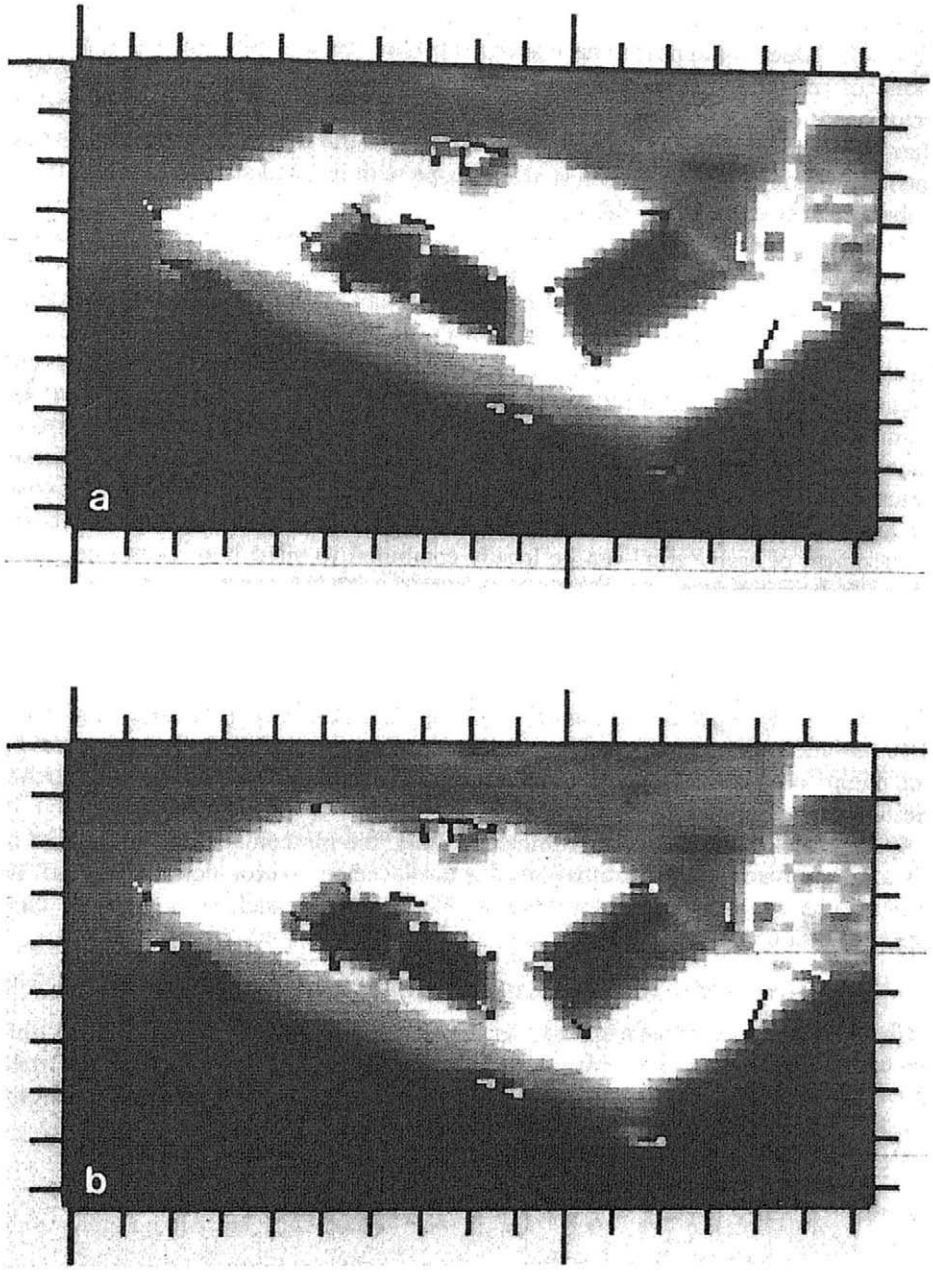

FIG. 5. Analogous to Fig. 4., (a) the first and (b) second half frames of the first TV frame have been compared. The time interval is only 20 msec long.

for 3-D interpretation. The directionally selective operator discussed by Marr and Ullman [33] does not appear to have such properties. Batali and Ullman [59] studied a kind of relaxation approach to determine both components of the displacement vector based on the output of the zero-crossing displacement operator suggested by Marr and Ullman [33]. They essentially assume that the displacement vector varies smoothly although their approach is able to cope with isolated strong discontinuities in the displacement vector field.

If we neglect the fact that a curved zero-crossing contour deviates from the locus of maximum gray value slope as discussed in Section 2.4, then we may compare the approach developed here to that of Marr and Ullman [33] in the following manner: Marr and Ullman determine only the sign of the displacement for a small zero-crossing segment. They therefore have to combine this piece of information from several neighboring zero-crossing contour segments in order to determine the displacement vector. This implies the assumption that the displacement vector varies smoothly—supplemented by a facility to detect discontinuities in the displacement vector field. The concept of a "gray value corner," however, comprises a zero-crossing segment which exhibits sufficient curvature so that the statements about the displacement of each "arm" can be locally combined to yield both components of the displacement vector.

It will be an interesting question to study the influence of a Gaussian filter applied to the gray value distribution prior to the determination of the displacement vector field according to the method developed here. Moreover, this approach has been formulated in an asymmetric manner (see, e.g., Eq. (15) which describes the gray value variation in frame 2 by the approximating function derived from frame 1). One might formulate this approach in a more symmetric manner. These and other questions have to be studied.

Apart from the encouraging preliminary results, the mathematical formulation of this local approach for the estimation of a displacement vector field as well as its generalization to include the approach of Horn and Schunck as a limiting case appeal to intuition.

APPENDIX 1: APPROXIMATING QUADRIC AND ITS ERRORS

Given a quadratic domain of side length $(2k + 1)$ for the digitized gray value function $g(x, y)$ with $x, y = -k, -k + 1, \ldots, -1, 0, 1, \ldots, k$. It is further assumed that the measurement errors for $g(x, y)$ are independent of each other and are distributed according to a zero-mean normal distribution $N(0, \sigma)$.

We want to determine the free parameters in the quadric

$$f(x, y) = f_0 + f_x x + f_y y + \tfrac{1}{2} f_{xx} x^2 + f_{xy} xy + \tfrac{1}{2} f_{yy} y^2 \tag{A1}$$

in such a way that the sum of the squared differences between the measurement and the approximating expression—weighted by the (diagonal) inverted covariance matrix of the measurements—becomes a minimum.

Since it will be required later on, we introduce an explicit linear order among the raster points in the given domain.

$$x_i = (i - 1)\mathrm{mod}(2k + 1) - k \tag{A2a}$$

$$y_i = k - (i - 1)\mathrm{div}(2k + 1) \tag{A2b}$$

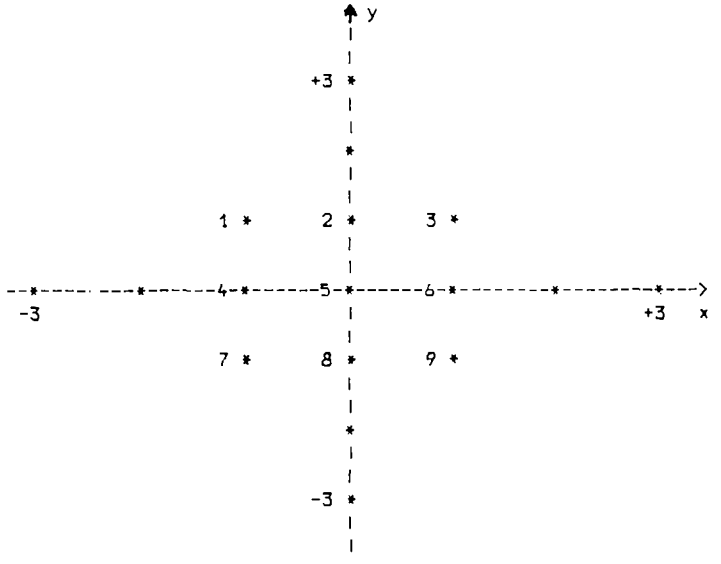

FIG. A1. The digit beside the inner nine raster points indicates the chosen sequence number for a square domain of side length $2k + 1$ for $k = 1$.

for $i = 1, 2, \ldots, N$ with $N = (2k + 1)^2$, where mod represents the modulo operation and div truncating integer division. This serialization starts with the leftmost pixel in the top row and proceeds row by row from left to right as indicated for the example of a 3×3 domain in Fig. A1.

We use this order to introduce column vectors with N components for measurement values

$$\mathbf{G} = g(x_i, y_i), \qquad i = 1, 2, \ldots, N \tag{A3}$$

and the approximation function

$$\mathbf{F}(\mathbf{P}) = f(x_i, y_i; \mathbf{P}), \qquad i = 1, 2, \ldots, N, \tag{A4}$$

where $\mathbf{P}$ represents the column vector of parameters

$$\mathbf{P} = (P_1, P_2, P_3, P_4, P_5, P_6)' = (f_0, f_x, f_y, f_{xx}, f_{xy}, f_{yy})'. \tag{A5}$$

The weight matrix W is given by

$$W = (W_{ij}) = \left(\frac{1}{\sigma^2}\delta_{ij}\right) \qquad \text{with} \quad \delta_{ij} = 1 \quad \text{for} \quad i = j, \qquad = 0 \quad \text{otherwise}, \tag{A6}$$

for $i, j = 1, 2, \ldots, N$. With these conventions, the expression to be minimized can be written

$$M = (\mathbf{F}(\mathbf{P}) - \mathbf{G})'W(\mathbf{F}(\mathbf{P}) - \mathbf{G}). \tag{A7}$$

Let A denote the matrix of derivatives of F with respect to the components of $\mathbf{P}$, i.e.,

$$A_{il} = \partial f(x_i, y_i; \mathbf{P})/\partial P_l, \qquad i = 1, 2, \ldots, N, \quad l = 1, 2, \ldots, 6, \tag{A8a}$$

$$A_i = \left(1, x_i, y_i, \tfrac{1}{2}x_i^2, x_i y_i, \tfrac{1}{2}y_i^2\right) \qquad i = 1, \ldots, N. \tag{A8b}$$

Then $\mathbf{F}$ can be written as

$$\mathbf{F} = A\mathbf{P} \tag{A9}$$

and we have

$$M = (A\mathbf{P} - \mathbf{G})'W(A\mathbf{P} - \mathbf{G}). \tag{A10}$$

The requirement to minimize this expression by appropriate choice of parameters yields

$$\partial M/\partial \mathbf{P} = 0 = A'W(A\mathbf{P} - \mathbf{G}) \tag{A11a}$$

or

$$\mathbf{P} = (A'WA)^{-1}A'W\mathbf{G}. \tag{A11b}$$

Following Kitchen and Rosenfeld [32], we define

$$S_{pqr} = \sum_{i=1}^{N} x_i^p y_i^q g_i^r. \tag{A12}$$

Due to the symmetry of the domain, terms with $r = 0$ and odd powers of x_i or y_i vanish. Using this convention, the matrix $(A'WA)$ can be written in the form

$$A'WA = \frac{1}{\sigma^2}\begin{pmatrix} S_{000} & 0 & 0 & 1/2S_{200} & 0 & 1/2S_{020} \\ 0 & S_{200} & 0 & 0 & 0 & 0 \\ 0 & 0 & S_{020} & 0 & 0 & 0 \\ 1/2S_{200} & 0 & 0 & 1/4S_{400} & 0 & 1/4S_{220} \\ 0 & 0 & 0 & 0 & S_{220} & 0 \\ 1/2S_{020} & 0 & 0 & 1/4S_{220} & 0 & 1/4S_{040} \end{pmatrix} \tag{A13}$$

with

$$S_{000} = N = (2k + 1)^2 \tag{A14a}$$

$$S_{200} = S_{020} = 1/3k(k + 1)(2k + 1)^2 \tag{A14b}$$

$$S_{220} = 1/9k^2(k + 1)^2(2k + 1)^2 \tag{A14c}$$

$$S_{400} = S_{040} = 1/15k(k + 1)(2k + 1)^2(3k^2 + 3k - 1) \tag{A14d}$$

from which it follows that

$$S_{000}S_{220} = S_{200}^2 \tag{A15}$$

$$\det(A'WA) = \tfrac{1}{16}(1/\sigma^2)^6 S_{000}^2 S_{220}^2 (S_{400} - S_{220})^2. \tag{A16}$$

The inverse of $A'WA$ represents the effects of error propagation to the parameters, i.e., it is the covariance matrix for the parameters. It is given by

$$W_P^{-1} = (A'WA)^{-1} \tag{A17}$$

$$= e^2 \begin{pmatrix} S_{220}(S_{400}+S_{220}) & 0 & 0 & -2S_{200}S_{220} & 0 & -2S_{020}S_{220} \\ 0 & S_{200}(S_{400}-S_{220}) & 0 & 0 & 0 & 0 \\ 0 & 0 & S_{020}(S_{400}-S_{220}) & 0 & 0 & 0 \\ -2S_{200}S_{220} & 0 & 0 & 4S_{000}S_{220} & 0 & 0 \\ 0 & 0 & 0 & 0 & S_{000}(S_{400}-S_{220}) & 0 \\ -2S_{020}S_{220} & 0 & 0 & 0 & 0 & 4S_{000}S_{220} \end{pmatrix}$$

with

$$e^2 = \sigma^2/(S_{000}S_{220}(S_{400} - S_{220})). \tag{A17a}$$

For $A'WG$ we obtain

$$A'WG = \frac{1}{\sigma^2}\begin{pmatrix} S_{001} \\ S_{101} \\ S_{011} \\ 1/2S_{201} \\ S_{111} \\ 1/2S_{021} \end{pmatrix}. \tag{A18}$$

Combining (A17) and (A18) yields

$$P_1 = f_0 = \frac{(S_{400}+S_{220})S_{001} - S_{200}S_{201} - S_{020}S_{021}}{S_{000}(S_{400}-S_{220})}, \tag{A19a}$$

$$P_2 = f_x = S_{200}S_{101}/(S_{000}S_{220}) = S_{101}/S_{200}, \tag{A19b}$$

$$P_3 = f_y = S_{020}S_{011}/(S_{000}S_{220}) = S_{011}/S_{020}, \tag{A19c}$$

$$P_4 = f_{xx} = \frac{2(S_{000}S_{201} - S_{200}S_{001})}{S_{000}(S_{400}-S_{220})}, \tag{A19d}$$

$$P_5 = f_{xy} = S_{111}/S_{220}, \tag{A19e}$$

$$P_6 = f_{yy} = \frac{2(S_{000}S_{021} - S_{020}S_{001})}{S_{000}(S_{400}-S_{220})}. \tag{A19f}$$

Using the expressions given in Eqs. (A14) will yield the operators presented by Beaudet [38]—with the exception of Eq. (A19e), where Beaudet obtained the opposite sign. Kitchen and Rosenfeld [32] pointed out that Beaudet used row–column coordinates rather than the xy system used here.

The uncertainties propagated from the measurements $\mathbf{G}$ to the parameters $\mathbf{P}$ introduce uncertainties in the values of $\mathbf{F}(\mathbf{P})$. Employing error propagation again yields the $N * N$ covariance matrix W_F^{-1} for the values of $\mathbf{F}(\mathbf{P})$:

$$W_F^{-1} = A W_P^{-1} A'. \tag{A20a}$$

The element in row i and column j of this covariance matrix is given by the expression

$$\begin{aligned}
\left(W_F^{-1}\right)_{ij} = e^2 * \{ & (S_{400} + S_{220}) S_{220} - \left(x_i^2 + x_j^2\right) S_{200} S_{220} \\
& - \left(y_i^2 + y_j^2\right) S_{020} S_{220} + x_i x_j (S_{400} - S_{220}) S_{200} \\
& + y_i y_j (S_{400} - S_{220}) S_{020} + \left(x_i^2 x_j^2 + y_i^2 y_j^2\right) S_{000} S_{220} \\
& + x_i x_j y_i y_j S_{000} (S_{400} - S_{220}) \}
\end{aligned} \tag{A20b}$$

$$\begin{aligned}
\left(W_F^{-1}\right)_{ij} = {} & \sigma^2 / k^2 (k+1)^2 (2k+1)^2 (2k+3)(2k-1) \\
& * \{ k^2 (k+1)^2 (14k^2 - 14k - 3) - 15 k^2 (k+1)^2 \left(x_i^2 + x_j^2 + y_i^2 + y_j^2\right) \\
& + 3k(k+1)(2k+3)(2k-1)(x_i x_j + y_i y_j) \\
& + 45k(k+1)\left(x_i^2 x_j^2 + y_i^2 y_j^2\right) + 9(2k+3)(2k-1) x_i x_j y_i y_j \}
\end{aligned} \tag{A20c}$$

for $i, j = 1, 2, \ldots, N = (2k+1)^2$. To illustrate the dependencies, this matrix is given for $k = 1$, i.e., a 3×3 domain.

$$W_F^{-1} = \frac{\sigma^2}{36} \begin{pmatrix}
29 & 8 & -1 & 8 & -4 & -4 & -1 & -4 & 5 \\
8 & 20 & 8 & -4 & 8 & -4 & -4 & 8 & -4 \\
-1 & 8 & 29 & -4 & -4 & 8 & 5 & -4 & -1 \\
8 & -4 & -4 & 20 & 8 & 8 & 8 & -4 & -4 \\
-4 & 8 & -4 & 8 & 20 & 8 & -4 & 8 & -4 \\
-4 & -4 & 8 & 8 & 8 & 20 & -4 & -4 & 8 \\
-1 & -4 & 5 & 8 & -4 & -4 & 29 & 8 & -1 \\
-4 & 8 & -4 & -4 & 8 & -4 & 8 & 20 & 8 \\
5 & -4 & -1 & -4 & -4 & 8 & -1 & 8 & 29
\end{pmatrix}. \tag{A21}$$

It can be seen that W_F^{-1} is symmetric about both diagonals and that the sum of each row or column is equal to σ^2.

Since these properties will be used later, their general validity will be shown here. The symmetry about the main diagonal follows from the symmetry of Eq. (A20c) for the interchange of row and column indices i and j. The symmetry property of W_F^{-1} about the other diagonal from the lower left to the upper right corner can be formally expressed as

$$\left(W_F^{-1}\right)_{ij} = \left(W_F^{-1}\right)_{N+1-j,\, N+1-i} \quad \text{for} \quad i, j = 1, \ldots, N = (2k+1)^2. \tag{A22}$$

If we increase i or j from 1 to N, the index value $N + 1 - j$ or $N + 1 - i$, respectively, decreases from N to 1, i.e., we proceed through the raster points just in the opposite order than that given by Eqs. (A2). This is equivalent, however, to a reflection of the coordinate system at the origin. Inspection of Eq. (A20c) will show that it remains invariant if we replace x by $-x$ and y by $-y$. As a consequence, W_F^{-1} is symmetric about the other diagonal, too.

When Eq. (A20c) is summed over the row index i, all terms containing odd powers of x_i or y_i will vanish due to the symmetry of the environment around the origin. The remaining terms can be evaluated using Eqs. (A12), (A14), and (A15). Straightforward algebraic manipulations will then yield equality of this sum with σ^2. The equivalent statement for a column follows from the symmetry of W_F^{-1} about the main diagonal.

APPENDIX 2: SYMMETRIES AND INVERSION OF A COVARIANCE MATRIX FOR INTERFRAME DIFFERENCES

If the errors in the gray value measurements at frame times $t1$ and $t2$ are uncorrelated, the covariance matrix for the interframe difference defined by Eq. (28) can be written in the form

$$\left(W_G^{-1}\right)_{ij} = \sigma^2 \delta_{ij} + \left(W_F^{-1}\right)_{ij}. \tag{30}$$

This matrix is symmetric about the main diagonal as well as about the other diagonal from the lower left to the upper right corner. Since the rows and columns of W_F^{-1} sum to σ^2, the rows and columns of W_G^{-1} sum to $2\sigma^2$. It will now be shown that corresponding statements can be proven with respect to the inverse of W_G^{-1}, namely, W_G.

Symmetry of W_G about the main diagonal, given this property for W_G^{-1}, follows from general matrix theory. Let $''$ indicate transposition about the other diagonal. In analogy to Eq. (A22), we have

$$\left(W_G^{-1}\right)''_{ij} = \left(W_G^{-1}\right)_{N+1-j,\,N+1-i} = \left(W_G^{-1}\right)_{ij}. \tag{A22a}$$

By definition of W_G as the inverse of W_G^{-1} we can write

$$\delta_{ik} = \sum_j \left(W_G^{-1}\right)_{ij} (W_G)_{jk} \tag{A23a}$$

$$= \sum_j \left(W_G^{-1}\right)_{ji} (W_G)_{kj} = \sum_j \left(W_G^{-1}\right)_{N+1-j,\,i} (W_G)_{k,\,N+1-j} \tag{A23b}$$

$$= \sum_j \left(W_G^{-1}\right)_{N+1-j,\,N+1-i} (W_G)_{N+1-k,\,N+1-j}$$

$$= \sum_j \left(W_G^{-1}\right)''_{ij} (W_G)''_{jk}, \tag{A23c}$$

where we used the symmetry of W_G^{-1} and W_G about the main diagonal and the fact that the name of the summation index as well as the order in which the summation is performed do not change the expressions in Eq. (A23b). The last expression of (A23c) shows that W_G'' is the inverse of $(W_G^{-1})''$.

It remains to be shown that the sum of a row of W_G equals $1/(2\sigma^2)$. The (i, j)th element of W_G can be expressed by the cofactor for the corresponding element of W_G^{-1}:

$$(W_G)_{ij} = \text{cofactor}\left(W_G^{-1}\right)_{ij} / \det\left(W_G^{-1}\right). \tag{A24}$$

The determinant of W_G^{-1} remains invariant if one row is added to another. Therefore, $\det(W_G^{-1})$ remains invariant if we add all other rows ($k \neq i$) to row i. Since adding all elements of a column yields $2\sigma^2$, row i will then contain only elements equal to $2\sigma^2$. If we now develop $\det(W_G^{-1})$ for row i, we obtain

$$\det\left(W_G^{-1}\right) = 2\sigma^2 \sum_j \text{cofactor}\left(W_G^{-1}\right)_{ij}. \tag{A25}$$

If all elements in row i of W_G are added, Eq. (A24) together with (A25) will yield the result

$$\sum_j (W_G)_{ij} = 1/(2\sigma^2). \tag{A26}$$

The properties of W_G allow us to drastically reduce the number of elements which actually have to be computed. All elements below the main and the other diagonal are given by the symmetry properties of W_G. Moreover, the fact that the sum of elements in a row or column equals $1/(2\sigma^2)$ can be exploited to further reduce the number of elements which have to be computed in the case of an $N * N$ matrix (with $N = (2k + 1)^2$) to $(N + 1)(N - 1)/4$. For a square environment with $2k + 1 = 3$ pixels on one side such as that depicted in Fig. A1, only 20 instead of 81 elements need to be computed. For this case, W_F^{-1} is given explicitly in (A21). The inverse W_G has been determined based on this W_F^{-1} according to Eq. (30),

$$W_G = \frac{1}{72\sigma^2} \begin{pmatrix} 43 & -8 & 1 & -8 & 4 & 4 & 1 & 4 & -5 \\ -8 & 52 & -8 & 4 & -8 & 4 & 4 & -8 & 4 \\ 1 & -8 & 43 & 4 & 4 & -8 & -5 & 4 & 1 \\ -8 & 4 & 4 & 52 & -8 & -8 & -8 & 4 & 4 \\ 4 & -8 & 4 & -8 & 52 & -8 & 4 & -8 & 4 \\ 4 & 4 & -8 & -8 & -8 & 52 & 4 & 4 & -8 \\ 1 & 4 & -5 & -8 & 4 & 4 & 43 & -8 & 1 \\ 4 & -8 & 4 & 4 & -8 & 4 & -8 & 52 & -8 \\ -5 & 4 & 1 & 4 & 4 & -8 & 1 & -8 & 43 \end{pmatrix}. \tag{A27}$$

The covariance matrix W_F^{-1} does not have a full rank of N. Because of the symmetries of the environment around **X**, the errors in the function values computed for $f1(x, y; \mathbf{P1})$ are not all independent from each other. The addition of the diagonal covariance matrix for the measurements $g(x, y, t2)$ restores W_G^{-1} to full rank of N, however. To avoid a misunderstanding, the reader should note that the covariance matrix W_F^{-1} for the computed values of $f1(x, y; \mathbf{P1})$ is different from the (diagonal) covariance matrix of the measurements $g(x, y, t1)$ based on which the parameter vector **P1** has been determined.

ACKNOWLEDGMENTS

Discussions with L. Dreschler, W. Enkelmann, and Y. S. Hsu as well as the help of L. Dreschler and W. Enkelmann in providing Figs. 1, 2, 4, and 5 are gratefully acknowledged. I thank B. Radig and H. Westphal for valuable comments on a draft version of this paper. Mrs. R. Jancke helped me much by her patient editing support.

REFERENCES

1. S. Ullman, *The Interpretation of Visual Motion*, MIT Press, Cambridge, Mass., 1979.
2. J. K. Aggarwal and N. I. Badler (Eds.), Special issue on motion and time-varying imagery, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**, 1980.
3. M. Asada and S. Tsuji, Reconstruction of Three-Dimensional Motions from Image Sequences, IEEE Conf. on Pattern Recognition and Image Processing, 1981, pp. 88–90.
4. W. F. Clocksin, Perception of surface slant and edge labels of optical flow: A computational approach, *Perception* **9**, 1980, 253–269.
5. L. Dreschler and H.-H. Nagel, Volumetric Model and 3D-Trajectory of a Moving Car Derived from Monocular TV-Frame Sequences of a Street Scene, International Joint Conference on Artificial Intelligence, 1981, pp. 692–697; extended version to appear in *Computer Graphics and Image Processing* **20** (3) (1982), pp. 199–228.
6. H.-H. Nagel, On the Derivation of 3D Rigid Point Configurations from Image Sequences, IEEE Conf. on Pattern Recognition and Image Processing, 1981, pp. 103–108.
7. H.-H. Nagel and B. Neumann, On 3D Reconstruction from Two Perspective Views, International Joint Conf. on Artificial Intelligence, 1981, pp. 661–663.
8. B. Neumann, 3D-Information aus mehrfachen Ansichten, in *Modelle und Strukturen 4th DAGM Symposium Hamburg, 6–8 Oktober 1981, Informatik-Fachberichte 49* (B. Radig, Ed.), pp. 93–111, Springer-Verlag, Berlin/Heidelberg/New York, 1981.
9. K. Prazdny, Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer, *Computer Graphics and Image Processing* **17**, 1981, 238–248.
10. R. Y. Tsai and T. S. Huang, Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces, Report R-921, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1981.
11. J. A. Webb and J. K. Aggarwal, Structure from Motion of Rigid and Jointed Objects, International Joint Conf. on Artificial Intelligence, 1981, pp. 686–691.
12. J. A. Webb, Shape and Structure from Motion of Objects, Ph.D. Dissertation, University of Texas, Austin, December 1981.
13. D. Marr and T. Poggio, A computational theory of human stereo vision, *Proc. Roy. Soc. London, Ser. B* **204**, 1979, 301–308.
14. W. E. L. Grimson, A computer implementation of a theory of human stereo vision, *Philos. Trans. Roy. Soc. London Ser. B* **292**, 1981, 217–253.
15. J. E. W. Mayhew and J. P. Frisby, Psychophysical and computational studies towards a theory of human stereopsis, *Artif. Intell.* **17**, 1981, 349–385.
16. A. N. Netravali and J. O. Limb, Picture coding: A review, *Proc. IEEE* **68**, 1980, 366–406.
17. A. K. Jain, Image data compression: A review, *Proc. IEEE* **69**, 1981, 349–389.
18. E. Dubois, B. Prasada, and M. S. Sabri, Image sequence coding, in *Image Sequence Analysis* (T. S. Huang, Ed.), pp. 229–287, Springer-Verlag, Berlin/Heidelberg/New York, 1981.
19. J. Roach and J. K. Aggarwal, Computer tracking of objects moving in space, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**, 1979, 127–135.
20. R. Kraasch, B. Radig, and W. Zach, Automatische Dreidimensionale Beschreibung bewegter Gegenstaende, in *Angewandte Szenenanalyse* (J. P. Foith, Ed.), pp. 208–215, Informatik Fachberichte 20, Springer-Verlag, Berlin/Heidelberg/New York, 1979. Preliminary results have been presented at the Workshop on Computer Analysis of Time-Varying Imagery, Philadelphia, PA., April 1979, pp. 42–43; see also International Conference on Pattern Recognition, 1980, pp. 1081–1084.
21. A. L. Gilbert, M. K. Giles, G. M. Flachs, R. B. Rogers, and Y. Hsun U, A real-time video tracking system, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**, 1980, 47–56.
22. B. Radig, Auswertung von digitisierten Fernsehbildern zur Beschreibung bewegter Objekte, Dissertation, Fachbereich Informatik (Maerz 1978), Universitaet Hamburg; see also Image region

extraction of moving objects, in *Image Sequence Analysis* (T. S. Huang, Ed.), pp. 311–354, Springer-Verlag, Berlin/Heidelberg/New York, 1981.
23. B. Radig, Inferential Region Extraction in TV Sequences, International Joint Conference on Artificial Intelligence, 1981, pp. 719–721.
24. H.-H. Nagel, Image sequence analysis: What can we learn from applications? in *Image Sequence Analysis* (T. S. Huang, Ed.), pp. 19–228, Springer-Verlag, Berlin/Heidelberg/New York, 1981.
25. J. K. Aggarwal, L. S. Davis, and W. N. Martin, Correspondence processes in dynamic scene analysis, *Proc. IEEE* **69**, 1981, 562–572.
26. H. P. Moravec, Towards Automatic Visual Obstacle Avoidance, International Joint Conference on Artificial Intelligence, 1977, p. 584; ibid., 1979, pp. 598–600.
27. M. J. Hannah, Bootstrap Stereo, Proc. Image Understanding Workshop, April 1980, pp. 201–208.
28. H. P. Moravec, Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover, Ph.D. Thesis, Department of Computer Science, STAN-CS-80-813, Stanford University. Available also as CMU-RI-TR-3, September 1980, Robotics Institute, Carnegie-Mellon University, Pittsburgh, Pa., see also International Joint Conference on Artificial Intelligence, 1981, pp. 785–790.
29. L. Dreschler, Ermittlung markanter Punkte auf den Bildern bewegter Objekte und Berechnung einer 3D-Beschreibung auf dieser Grundlage, Dissertation, June 1981, Fachbereich Informatik der Universitaet Hamburg.
30. H. -H. Nagel, From Digital Picture Processing to Image Analysis, Proc. International Conference on Image Analysis and Processing (V. Cantoni, Ed.), pp. 27–40, Pavia, Italy, October 22–24, 1980.
31. L. Dreschler and H. -H. Nagel, On the frame-to-frame correspondence between greyvalue characteristics in the images of moving objects, in *5th GI-Fachtagung GWAI, 1981, Bad Honnef, January 26–30, 1981 Informatik Fachberichte 47* (J. Siekmann, Ed.), pp. 18–29, Springer-Verlag, Berlin/Heidelberg/New York, 1981.
32. L. Kitchen and A. Rosenfeld, Gray-Level Corner Detection, TR-887, Computer Science Center, University of Maryland, College Park, 1980.
33. D. Marr and S. Ullman, Directional selectivity and its use in early visual processing, *Proc. Roy. Soc. London, Ser. B* **211**, 1981, 151–180.
34. D. Marr and E. Hildreth, Theory of edge detection, *Proc. Roy. Soc. London, Ser. B* **207**, 1980, 187–217.
35. H. H. Baker and T. O. Binford, Depth from Edge and Intensity Based Stereo, International Joint Conference on Artificial Intelligence, 1981, pp. 631–636.
36. S. Yam and L. S. Davis, Image Registration Using Generalized Hough Transforms, *Pattern Recognition and Image Processing*, 1981, pp. 526–533.
37. L. Dreschler and H.-H. Nagel, On the Selection of Critical Points and Local Curvature Extrema of Region Boundaries for Interframe Matching, International Conference on Pattern Recognition, 1982, pp. 542–544.
38. P. R. Beaudet, Rotationally Invariant Image Operators, International Joint Conference on Pattern Recognition, 1978, pp. 579–583.
39. J. O. Limb and J. A. Murphy, Measuring the speed of moving objects from television signals, *IEEE Trans. Commun.* **COM-23**, 1975, 474–478.
40. J. O. Limb and J. A. Murphy, Estimating the velocity of moving images in television signals, *Computer Graphics and Image Processing* **4**, 1975, 311–327.
41. C. Cafforio and F. Rocca, Methods for measuring small displacements of television images, *IEEE Trans. Inform. Theory* **IT-22**, 1976, 573–579.
42. C. Cafforio and F. Rocca, Tracking moving objects in television images, *Signal Processing* **1**, 1979, 133–140.
43. T. Fukinuki, H. Yoshigi, and K. Fukushima, Improvement of inter-frame predictive coding of TV-signals by utilizing visual properties for moving objects, *Trans. IECE Japan* **E 59**, 1976, 28–29; **59-A** 1976, 764–771. [Japanese]
44. C. L. Fennema and W. B. Thompson, Velocity determination in scenes containing several moving objects, *Computer Graphics and Image Processing* **9**, 1979, 301–315.
45. R. J. Schalkoff and E. S. McVey, Algorithm Development for Real-Time Automatic Video Tracking Systems, Proc. 3rd International Computer Software and Applications Conference, Chicago, November 1979, pp. 504–511; *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-4**, 1982, 2–10.
46. A. N. Netravali and J. D. Robbins, Motion-compensated coding: Some new results, *Bell Sys. Tech. J.* **59**, 1980, 1735–1745.

47. W. B. Thompson and S. T. Barnard, Lower-level estimation and interpretation of visual motion, *IEEE Comput.* **14**, No. 8, August 1981, 20–28.
48. B. K. P. Horn and B. G. Schunck, Determining optical flow, *Artif. Intell.* **17**, 1981, 185–203.
49. M. Yachida, Determining Velocity Map by 3-D Iterative Estimation, International Joint Conference on Artificial Intelligence, 1981, pp. 716–718.
50. T. Dinse, W. Enkelmann, and H.-H. Nagel, Untersuchung von Verschiebungsvektorfeldern in Bildfolgen, in *Modelle und Strukturen, 4th DAGM Symposium, Hamburg, Oktober 6–8, 1981, Informatik Fachberichte 49* (B. Radig, Ed.), pp. 69–75, Springer-Verlag, Berlin/Heidelberg/New York, 1981.
51. B. G. Schunck and B. K. P. Horn, Constraints on Optical Flow, IEEE Pattern Recognition and Image Processing, 1981, pp. 205–210.
52. W. E. Snyder, S. A. Rajala, and G. Hirzinger, Image Modeling, the Continuity Assumption and Tracking, International Conference on Pattern Recognition, 1980, pp. 1111–1114.
53. K. Prazdny, A Simple Method for Recovering Relative Depth Map in the Case of a Translating Sensor, International Joint Conference on Artificial Intelligence, 1981, pp. 698–699.
54. A. Korn and R. Kories, Motion Analysis in Natural Scenes Picked up by a Moving Optical Sensor, International Conference on Pattern Recognition, 1980, 1251–1254.
55. R. Kories, Determination of Displacement Vector Fields for General Camera Motions, *IEEE* Pattern Recognition and Image Processing, 1981, pp. 115–117.
56. W. B. Thompson, Combining motion and contrast for segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**, 1980, 543–549.
57. S. M. Haynes and R. Jain, Detection of Moving Edges, CSC-82-004, Department of Computer Science, Wayne State University, Detroit, Mich., 1981.
58. H.-H. Nagel and W. Enkelmann, Investigation of Second Order Grey Value Variations to Estimate Corner Point Displacements, International Conference on Pattern Recognition, 1982, pp. 768–773.
59. J. Batali and S. Ullman, Motion Detection and Analysis, Proc. Image Understanding Workshop, pp. 69–75, November 1978. (L. S. Bauman, Ed.), Science Applications, Arlington, Virginia 22209.