Optical Flow and Motion Fields

Optical Flow Estimation: Advances and Comparisons

M. Otte[1] and H.-H. Nagel[1,2]

[1] Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Postfach 6980, D-76128 Karlsruhe, Germany
[2] Fraunhofer - Institut für Informations- und Datenverarbeitung (IITB), Karlsruhe

Abstract. This contribution investigates local differential techniques for estimating optical flow and its derivatives based on the brightness change constraint. By using the tensor calculus representation we build the Taylor expansion of the gray-value derivatives as well as of the optical flow in a spatiotemporal neighborhood. Such a formulation simplifies a unifying framework for all existing local differential approaches and allows to derive new systems of equations to estimate the optical flow and its derivatives. We also tested various optical flow estimation approaches on real image sequences recorded by a calibrated camera fixed on the arm of a robot. By moving the arm of the robot along a precisely defined trajectory we can determine the true displacement rate of scene surface elements projected into the image plane and compare it quantitatively with the results of different optical flow estimators.

1 Introduction

Estimation of optical flow and its derivatives is an important task in the area of computer vision. [Koenderink & van Doorn 76] studied the role of differential invariants of optical flow with respect to 3D-interpretation of image sequences. Specific 3D-tasks like obstacle detection ([Subbarao 90]) and computation of bounds for time to collision ([Nelson & Aloimonos 88; Cipolla & Blake 92]) may be solved based only on 0^{th} and 1^{st} order properties of optical flow. Furthermore, first-order properties, [Baraldi *et al.* 89; Girosi *et al.* 89; Negahdaripour & Lee 92], can be used as features for the classification of image patches into regions corresponding to independently moving objects.

[Nagel 92] proposed an approach to estimate spatiotemporal derivatives of the optical flow, whereas [Werkhoven & Koenderink 90] limited their approach to compute only spatial ones. All of them use at least second order derivatives of the gray-value pattern in order to capture the variation of optical flow in the neighborhood of the point under consideration. These differential approaches are to be distinguished from 'neighborhood-sampling' approaches which use the actual values of the gray-value gradient at every point of the observed neighborhood like [Nagel 85; Kearney *et al.* 87; Campani & Verri 90]. Regarding the above mentioned approaches to estimate optical flow and its derivatives, we have been able to build a common framework to derive all local differential methods based on the brightness change constraint and to present a method which combines differential with neighborhood-sampling techniques. Furthermore, within this framework we show that if one refers strictly to the assumptions of [Werkhoven & Koenderink 90], it will turn out that their approach is equal to the optical

Lecture Notes in Computer Science, Vol. 800
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

flow estimation technique presented by [Nagel 87].

Most publications presenting a new optical flow estimator discuss their results only qualitatively. A remarkably broad comparison has been presented by [Barron *et al.* 92], who implemented various optical flow estimation techniques and tested them quantitatively on several synthetic and quasi-synthetic (i.e. one real image with simulated camera motion) image sequences. Their comparison with real image sequences as input data has been limited to a qualitative judgement, since the true displacement rate fields of their image sequences are unknown.

In this contribution we use an image sequence recorded with a calibrated camera fixed on the arm of a robot which moves along a precisely defined 3D-trajectory. The calibration data as well as the known trajectory allow us to compute the true displacement rate field which is compared with results obtained by the new optical flow estimation approaches presented in this paper and with some of the estimators mentioned above.

2 Estimation of optical flow and of its derivatives

Optical flow is defined as the apparent velocity of gray-value structures. Assuming temporal constancy of a moving gray-value structure $g(x, y, t)$ results in the well known Optical Flow Constraint Equation (OFCE) postulated by [Horn & Schunck 81]:

$$\frac{\mathrm{d}}{\mathrm{d}t} g(x, y, t) \;=\; \nabla g^{\mathrm{T}} \boldsymbol{u} \;= g_{\mathrm{x}} u_1 + g_{\mathrm{y}} u_2 + g_{\mathrm{t}} \;=\; 0 \qquad (1)$$

with $\boldsymbol{u} = (u_1, u_2, 1)^{\mathrm{T}}$. This equation does only allow to estimate a linear combination of the components u_1 and u_2 of the optical flow. It has to be supplemented, therefore, by additional assumptions.

[Srinivasan 90; Chen *et al.* 93; Weber & Malik 93] estimate the gray-value gradient with a set of spatiotemporal filters to obtain two or more constraint equations. Unfortunately, this kind of estimating optical flow must fail the more, the better the estimated partial derivatives approximate the real derivatives, because in this case, the equations tends to become linearly dependent.

[Schunck 84; Aisbett 89; Negahdaripour & Yu 93] use a generalized form of the OFCE by assuming intensity changes due to shading or due to changes of the surface orientation with respect to light sources. [Verri & Poggio 89] argue that different biological visual systems do compute different optical flows. In biological visual systems it suffices to comply with the qualitative properties of the motion field as good candidates for subsequent analyzing cells. In this connection it must be allowed to define different "optical flows", since they have to be considered as an approximation of the true displacement rate field. In this contribution we use the OFCE as a basic constraint, since it allows to estimate shifts of gray-value patterns without any specific assumptions about surface properties or about the direction of light sources.

2.1 Common basis for local gradient based estimators

Apart from partially occluded objects or some artificial image sequences, the image of projected scene surfaces does not change abruptly with time in general, if the relative movement between camera and scenery is not too large or, more precisely, if the temporal sampling rate is high enough.

Thus, the OFCE can be considered valid not only as an approximation for a pixel position (x, y, t) but also for some local environment $(x + \delta x, y + \delta y, t + \delta t)$ of the actual position. This assumption allows us to write the OFCE in form of:

$$0 = \nabla g^{\mathrm{T}}(\boldsymbol{x} + \delta\boldsymbol{x})\boldsymbol{u}(\boldsymbol{x} + \delta\boldsymbol{x}) \ . \tag{2}$$

If we take the optical flow to vary at most linearly, we can substitute the term $\boldsymbol{u}(\boldsymbol{x} + \delta\boldsymbol{x})$ by a first order Taylor expansion

$$0 = \nabla g^{\mathrm{T}}(\boldsymbol{x} + \delta\boldsymbol{x})[\boldsymbol{u}(\boldsymbol{x}) + \nabla \boldsymbol{u}^{\mathrm{T}}(\boldsymbol{x})\delta\boldsymbol{x}] \ . \tag{3}$$

We distinguish optical flow estimation approaches based on equation (3) into two groups, namely neighborhood-sampling and gray-value gradient Taylor expansion estimation approaches, depending on a description for $\nabla g^{\mathrm{T}}(\boldsymbol{x} + \delta\boldsymbol{x})$.

2.2 Neighborhood-sampling approaches

If we consider a (spatial) region of $n \times n$ pixels around the actual point $\boldsymbol{x}$, we can sample the gray-value gradient at n^2 positions which yields an overconstrained system of n^2 equations. This method is used by [Nagel 85; Kearney *et al.* 87] to estimate the optical flow itself and by [Campani & Verri 90] to estimate the optical flow and in addition its linear spatial variation:

$$\begin{aligned} 0 = \quad & g_{\mathrm{x}}(x_0, y_0)u_1 + g_{\mathrm{y}}(x_0, y_0)u_2 + g_{\mathrm{t}}(x_0, y_0) \\ 0 = \quad & g_{\mathrm{x}}(x_1, y_1)(u_1 + u_{1\mathrm{x}}(x_1 - x_0) + u_{1\mathrm{y}}(y_1 - y_0)) \\ & + g_{\mathrm{y}}(x_1, y_1)(u_2 + u_{2\mathrm{x}}(x_1 - x_0) + u_{2\mathrm{y}}(y_1 - y_0)) + g_{\mathrm{t}}(x_1, y_1) \ . \\ & \vdots \end{aligned} \tag{4}$$

[Campani & Verri 90] used a region of between 10×10 and 70×70 pixels to achieve acceptable results. Obviously, this method can be extended to estimate not only the spatial variation of the optical flow but also its linear variation with time. In this case, the gray-value gradient has to be sampled in the temporal as well as in the spatial domain. Theoretically one can choose a similar region of up to 70×70×70 pixels which yields a system of 343,000(!) equations for eight unknowns, which is not practical in general. For our implementation we restricted the region to a 5×5×5 neighborhood.

2.3 Gray-value gradient Taylor expansion approaches

Instead of sampling the gray-value gradient in a small neighborhood, the gray-value gradient can be described as a Taylor series. In order to obtain a compact presentation, we write g instead of $g(\boldsymbol{x})$ and we use the Einstein summation convention for a three dimensional space:

$$\begin{array}{llll} x^n := \boldsymbol{x} = (x, y, t)^{\mathrm{T}}, & u^n := \boldsymbol{u} = (u_1, u_2, 1)^{\mathrm{T}}, & r^n := \boldsymbol{r} = (\delta x, \delta y, \delta t)^{\mathrm{T}}, \\ g_n := \nabla g = (\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial t})^{\mathrm{T}}, & g_{nm} = \mathrm{H} = \frac{\partial}{\partial \boldsymbol{x}} \nabla g^{\mathrm{T}}, & u_n^m = \mathrm{A} = \frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{u}^{\mathrm{T}}. \end{array} \tag{5}$$

Using this notation, equ. (3) can be replaced by Taylor series for both $\nabla g^{\mathrm{T}}(\boldsymbol{x} + \delta\boldsymbol{x})$ and $\boldsymbol{u}(\boldsymbol{x} + \delta\boldsymbol{x})$:

$$0 = (g_n + g_{nm}r^m + \frac{1}{2} g_{nmk} r^m r^k + \mathcal{O}((r^i)^3))(u^n + u_s^n r^s + \mathcal{O}((r^i)^2)) \tag{6}$$

Supposing an at most linearly varying optical flow and neglecting gray-value derivatives of order four and higher, we obtain the Basic Optical Flow Equation

$(\mathrm{BOFE}_{2,1})$ of third order[3]:

$$0 = g_n u^n + (g_{nm} u^n + g_n u_m^n) r^m + (\frac{1}{2} g_{nmk} u^n + g_{nm} u_k^n) r^m r^k + \frac{1}{2} g_{nmk} u_l^n r^m r^k r^l \quad (7)$$

We can now postulate three conditions for which the polynomial vanishes:

1. **Rigorous condition (RC) :** The polynomial vanishes identically, i.e. that all coefficients of the polynomial must be zero. This method was introduced by [Nagel 87].
2. **Integrated condition (IC):** The integral of the polynomial over a small region must vanish. It is used to rederive the approach of [Werkhoven & Koenderink 90].
3. **Sampling condition (SC):** The SC merges the proposed neighborhood-sampling method with the gray-value gradient Taylor expansion method by choosing appropriate values $\delta x, \delta y, \delta t$ to express the neighboring points at which the BOFE is postulated to be valid. This method is new.

Solution under the Rigorous Condition (RC) The RC demands that the polynomial should vanish identically in the neighborhood. As a consequence, all coefficients of the polynomial in equ. (7) must be zero. This leads to the following system of 20 equations for 8 unknowns:

$$\begin{pmatrix}
g_x & 0 & 0 & 0 & g_y & 0 & 0 & 0 \\
g_{xx} & g_x & 0 & 0 & g_{xy} & g_y & 0 & 0 \\
g_{xy} & 0 & g_x & 0 & g_{yy} & 0 & g_y & 0 \\
g_{xt} & 0 & 0 & g_x & g_{yt} & 0 & 0 & g_y \\
\frac{1}{2} g_{xxx} & g_{xx} & 0 & 0 & \frac{1}{2} g_{xxy} & g_{xy} & 0 & 0 \\
g_{xxy} & g_{xy} & g_{xx} & 0 & g_{xyy} & g_{yy} & g_{xy} & 0 \\
g_{xxt} & g_{xt} & 0 & g_{xx} & g_{xyt} & g_{yt} & 0 & g_{xy} \\
\frac{1}{2} g_{xyy} & 0 & g_{xy} & 0 & \frac{1}{2} g_{yyy} & 0 & g_{yy} & 0 \\
g_{xyt} & 0 & g_{xt} & g_{xy} & g_{yyt} & 0 & g_{yt} & g_{yy} \\
\frac{1}{2} g_{xtt} & 0 & 0 & g_{xt} & \frac{1}{2} g_{ytt} & 0 & 0 & g_{yt} \\
0 & \frac{1}{2} g_{xxx} & 0 & 0 & 0 & \frac{1}{2} g_{xxy} & 0 & 0 \\
0 & g_{xxy} & \frac{1}{2} g_{xxx} & 0 & 0 & g_{xyy} & \frac{1}{2} g_{xxy} & 0 \\
0 & g_{xxt} & 0 & \frac{1}{2} g_{xxx} & 0 & g_{xyt} & 0 & \frac{1}{2} g_{xxy} \\
0 & \frac{1}{2} g_{xyy} & g_{xxy} & 0 & 0 & \frac{1}{2} g_{yyy} & g_{xyy} & 0 \\
0 & g_{xyt} & g_{xxt} & g_{xxy} & 0 & g_{yyt} & g_{xyt} & g_{xyy} \\
0 & \frac{1}{2} g_{xtt} & 0 & g_{xxt} & 0 & \frac{1}{2} g_{ytt} & 0 & g_{xyt} \\
0 & 0 & \frac{1}{2} g_{xyy} & 0 & 0 & 0 & \frac{1}{2} g_{yyy} & 0 \\
0 & 0 & g_{xyt} & \frac{1}{2} g_{xyy} & 0 & 0 & g_{yyt} & \frac{1}{2} g_{yyy} \\
0 & 0 & \frac{1}{2} g_{xtt} & g_{xyt} & 0 & 0 & \frac{1}{2} g_{ytt} & g_{yyt} \\
0 & 0 & 0 & \frac{1}{2} g_{xtt} & 0 & 0 & 0 & \frac{1}{2} g_{ytt}
\end{pmatrix}
\begin{pmatrix} u_1 \\ u_{1x} \\ u_{1y} \\ u_{1t} \\ u_2 \\ u_{2x} \\ u_{2y} \\ u_{2t} \end{pmatrix}
= -
\begin{pmatrix}
g_t \\ g_{xt} \\ g_{yt} \\ g_{tt} \\ \frac{1}{2} g_{xxt} \\ g_{xyt} \\ g_{xtt} \\ \frac{1}{2} g_{yyt} \\ g_{ytt} \\ \frac{1}{2} g_{ttt} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{pmatrix} \quad (8)$$

In general the solution under the RC of the $\mathrm{BOFE}_{n,m}$ (denoted as $\mathrm{RC}_{n,m}$) results into a system of $\binom{n+m+3}{3}$ equations. The system of 35 equations derived from the $\mathrm{BOFE}_{3,1}$ can be found in [Otte 94].

If one neglects all third order partial gray-value derivatives corresponding to the $\mathrm{RC}_{1,1}$, one obtains exactly the system of equations presented by [Nagel 92]. Assuming in addition constancy of the optical flow with time, one can eliminate the fourth and last column as well as the fourth, seventh and ninth to last row of

[3] The two indices of the BOFE depends on the order of the Taylor expansion of the gray-value gradient and of the optical flow. The order of the BOFE itself is defined as the order of the resulting polynomial with respect to r^n (= sum of the two indices).

the resulting system of equations which is equal to those presented by [Nagel 87]. [Nagel 87; Nagel 92] pointed out that the respective systems of equations do not have full rank. In contrast to the $RC_{1,1}$, the $RC_{2,1}$ has full rank in general which indicates that third order derivatives of the gray-value pattern are important to estimate first order derivatives of the optical flow.

Solution under the Integrated Condition (IC) In case of the IC, we assume temporal constancy of the average gray-value over a small region, i.e. the integral of the polynomial (7) over a spatiotemporal region $\mathcal{R}$ must vanish. Before we determine the $IC_{2,1}$, we want to rederive the approach of [Werkhoven & Koenderink 90]. Since they describe variations of gray-value derivatives by first order Taylor expansions, we have to derive a solution for the $BOFE_{1,1}$:

$$0 = \int_{\mathcal{R}} (g_n u^n + (g_{nm} u^n + g_n u^n_m) r^m + g_{nm} u^n_k r^m r^k) \mathrm{d}r^i \tag{9}$$

Let the region $\mathcal{R}$ be a spatiotemporal sphere symmetrically centered around the origin with radius $\sigma\sqrt{5}$. Dividing equ. (9) through $\int_{\mathcal{R}} \mathrm{d}r^i$ yields:

$$0 \;=\; g_n u^n + \sigma^2 \; g_{nm} u^n_k \delta^{mk} \;=\; g_n u^n + \sigma^2 \; g_{nm} u^n_m \;, \tag{10}$$

where δ^{mk} represents the Kronecker delta. Since the OFCE (1) is considered as a constraint at all pixel positions, the partial derivatives of equ. (1) must vanish. Executing the same steps for the partial derivatives of equ. (1) as for equ. (9) and (10) yields

$$0_l = \frac{\partial}{\partial x_l}(g_n u^n + \sigma^2 g_{nm} u^n_m) = g_{nl} u^n + g_n u^n_l + \sigma^2 g_{nml} u^n_m \tag{11a}$$

$$0_{ls} = \frac{\partial^2}{\partial x_l \partial x_s}(g_n u^n + \sigma^2 g_{nm} u^n_m) = g_{nls} u^n + g_{nl} u^n_s + g_{ns} u^n_l + \sigma^2 g_{nmls} u^n_m. \tag{11b}$$

If we consider variations only in the spatial domain as [Werkhoven & Koenderink 90] did, equ. (11a) and (11b) have to be differentiated only with respect to x and y which leads to five equations. Together with equ. (10) we obtain a system of six equations for six unknowns, which represents the same system for optical flow estimation as that derived by [Werkhoven & Koenderink 90][4].

Considering the steps we presented in order to derive the system of equations of [Werkhoven & Koenderink 90], one can establish an inconsistency in the assumptions underlying their approach. First of all we started with the $BOFE_{1,1}$ which implies that all gray-value derivatives higher than second order can be neglected. In the final result (10) – (11b), however, we keep third and fourth order spatiotemporal derivatives of the gray-value pattern. An equivalent inconsistency can be found in [Werkhoven & Koenderink 90]. They introduce coefficients $l_{n,m}$ as the correlation of gray-values $g(x, y)$ with a receptive field $\Psi_{n,m}$ which is the $(n+m)^{\text{th}}$ order derivative of a two-dimensional Gaussian, scaled with the $(n+m)^{\text{th}}$ power of its standard deviation. In their equ. (18) they express a receptive field in terms of a linear Taylor expansion regardless of its order. Although they use receptive fields of up to fourth order, they neglect third and higher order terms when approximating a first order receptive field through a linear Taylor expansion.

[4] Apart from the fact that the C matrix of [Werkhoven & Koenderink 90] is multiplied by -1 due to a mistake in the sign in their equation (18).

We derive now the solution under the IC starting with the $\mathrm{BOFE}_{2,1}$. Integration of the polynomial (7) over a region $\mathcal{R}$ yields under the same assumptions as for equ. (10):

$$0 = g_n u^n + \sigma^2 \, g_{nm} u_m^n + \frac{\sigma^2}{2} g_{nmm} u^n. \tag{12}$$

Since we considered gray-value derivatives of up to third order in this derivation, we can formulate appropriate conditions analogous to equ. (11a) to (11b). If we neglect all gray-value derivatives of higher than third order and all second order derivatives of the optical flow in accordance with the assumptions underlying the $\mathrm{BOFE}_{2,1}$, we obtain:

$$0_l = g_{nl} u^n + g_n u_l^n + \sigma^2 \, g_{nml} u_m^n + \frac{\sigma^2}{2} g_{nmm} u_l^n \tag{13a}$$

$$0_{ls} = g_{nls} u^n + g_{nl} u_s^n + g_{ns} u_l^n \tag{13b}$$

$$0_{lsr} = g_{nls} u_r^n + g_{nlr} u_s^n + g_{nsr} u_l^n \; . \tag{13c}$$

The equations with $l = s$ in (13b) and (13c) allow to eliminate the gray-value variations scaled with σ^2 in equ. (12) and (13a). The resulting system of equations is exactly the same result as the $\mathrm{RC}_{2,1}$ (8). Neglecting all third order gray-value derivatives corresponding to the $\mathrm{BOFE}_{1,1}$, and assuming constancy of the optical flow with time as [Werkhoven & Koenderink 90] did, one obtains the same approach as [Nagel 87].

Solution under the Sampling Condition (SC) The solution under the SC combines the neighborhood-sampling method with the gray-value gradient Taylor expansion method. Disadvantage of the neighborhood-sampling method is the large number of equations usually used ([Campani & Verri 90] consider up to 4900 points, i.e. up to 4900 equations). As we showed, the $\mathrm{RC}_{2,1}$ and $\mathrm{RC}_{3,1}$ need only 20 or 35 equations but one has to estimate at least third order derivatives of the gray-value distribution.

The RC requires that all coefficients of the polynomial (7) must be zero. Instead of this rigorous condition we can choose appropriate values $\delta x, \delta y, \delta t$ to select a number of neighboring points. Since the $\mathrm{BOFE}_{2,1}$ (7) represents a polynomial with respect to $r^n = (\delta x, \delta y, \delta t)$, it can be written in the form[5]

$$\begin{aligned} 0 = \quad & c + c_1 r^1 + c_2 r^2 + c_3 r^3 \\ & + c_{11}(r^1)^2 + c_{12} r^1 r^2 + c_{13} r^1 r^3 + c_{22}(r^2)^2 + c_{23} r^2 r^3 + c_{33}(r^3)^2 \\ & + c_{111}(r^1)^3 + c_{112}(r^1)^2 r^2 + c_{113}(r^1)^2 r^3 + c_{122} r^1 (r^2)^2 + c_{123} r^1 r^2 r^3 \\ & + c_{133} r^1 (r^3)^2 + c_{222}(r^2)^3 + c_{223}(r^2)^2 r^3 + c_{233} r^2 (r^3)^2 + c_{333}(r^3)^3 . \end{aligned} \tag{14}$$

Selecting all possible values for r^1, r^2, r^3 from the set $\{-1, 0, 1\}$ which is related to the 3×3×3 neighborhood around the actual position leads to a system of 27 equations:

$$\begin{aligned} & r^1 \; r^2 \; r^3 \\ & 0 \quad 0 \quad 0\text{:} \;\; 0 = c \\ & 1 \quad 0 \quad 0\text{:} \;\; 0 = c + c_1 + c_{11} + c_{111} \\ & \qquad \vdots \\ & -1 -1 -1\text{:} \;\; 0 = c - c_1 - c_2 - c_3 + c_{11} + c_{12} + c_{13} + c_{22} + c_{23} + c_{33} - c_{111} \\ & \qquad\qquad - c_{112} - c_{113} - c_{122} - c_{123} - c_{133} - c_{222} - c_{223} - c_{233} - c_{333}, \end{aligned} \tag{15}$$

[5] Note: the upper index denotes the component; if a component has to be raised to second or third power, it is enclosed by parentheses.

which can be reduced to a system of 17 equations. This system of equations is similar to the $RC_{2,1}$ (8). The 11^{th}, 17^{th}, and 20^{th} equation of the $RC_{2,1}$ is a summand in the 2^{nd} to 4^{th} row of the SC result. Exploiting this observation reduces the number of equations by 3. The $SC_{3,1}$ reduces the number of 35 equations in case of the $RC_{3,1}$ to only 23 [Otte 94].

To sum up the three approaches – RC, IC and SC –, one has to consider at least third order gray-value derivatives to be able to estimate first order derivatives of the optical flow. Another important result is the fact that the IC leads to the same result as the RC does, if one applies the initial assumptions consistently. Last but not least, with the proposed SC we presented a new method which reduces the number of equations from 20 and 35 of the $RC_{2,1}$, and $RC_{3,1}$ to 17 and 23, respectively.

3 Comparison between optical flow estimators

[Barron *et al.* 92] compared quantitatively the deviation of estimated optical flow vectors of various approaches with respect to the true displacement rate field. However, the image sequences used in their comparison were generated synthetically. In contrast, we record real image sequences prepared with a camera mounted on the moving arm of a robot. Fig. 1 shows two frames of such an image sequence. The camera moves with pure 3D-translation towards the depicted scene which is stationary with the exception of the marbled light block, which translates to the left.

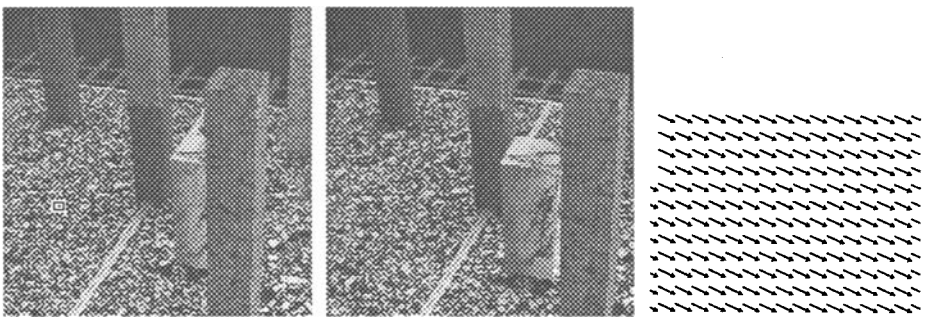

Fig. 1. 7^{th}, and 50^{th} frame of an image sequence recorded with a camera mounted on a robot arm moving with pure translation toward the scene. In addition to the camera displacement, the marbled block translates to the left. Right: True displacement rates of the little ground surface section marked in the left image.

If one considers a region with a homogeneous displacement – for example a part of the ground surface – one expects a robust estimation of optical flow vectors since there are no discontinuities. Fig. 1 (right) shows the true displacement rates correspond to the clipping of the ground surface marked in the left image of Fig. 1. In Fig. 2 one can see in the upper row the optical flow vectors estimated by the approach of [Werkhoven & Koenderink 90], by the $SC_{2,1}$ and by the neighborhood-sampling method. In the lower row, the resulting difference vectors between estimated and true displacement rate are shown.

The gray-value derivatives were computed by a convolution of the images with trivariate spatiotemporal derivatives of a Gaussian with a standard deviation of 1.5 in the spatial and 1.0 in the temporal domain. Optical flow vectors are only shown if the smallest singular value of the corresponding coefficient matrix exceeds a chosen threshold.

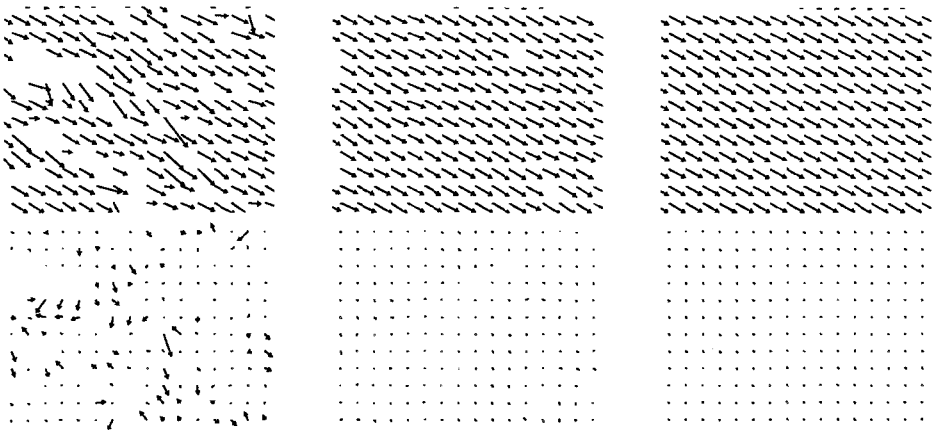

Fig. 2. Difference vectors (lower row) between the true displacement rates and estimated optical flow vectors (upper row) estimated by the approach of [Werkhoven & Koenderink 90] (left), by the $SC_{2,1}$ (center), and by the neighborhood-sampling method including a spatiotemporal neighborhood of 5×5×5 pixels (right).

3.1 Quantitative comparison

[Barron *et al.* 92] use an angular measure of error for their quantitative comparisons. The optical flow as well as the true displacement vectors are extended to 3D vectors with an arbitrarily selected value of 1 as third component in order not to overrate relative differences of short vectors. Let $\boldsymbol{u} = (u_1, u_2, 1)^{\mathrm{T}}$ be the true displacement rate and $\hat{\boldsymbol{u}} = (\hat{u}_1, \hat{u}_2, 1)^{\mathrm{T}}$ be the estimated optical flow. The angular error is then defined as $\varepsilon_\alpha = \arccos(\boldsymbol{u}^{\mathrm{T}}\hat{\boldsymbol{u}}/\|\boldsymbol{u}\|\,\|\hat{\boldsymbol{u}}\|)$. The problem of this error measure is that differences of large vectors correspond to relatively small angular errors. In addition, symmetrical deviations of estimated vectors from the true value result in different angular errors: let $\boldsymbol{u} = (1.5, 0, 1)^{\mathrm{T}}$ be the true displacement rate, $\hat{\boldsymbol{u}}_1 = (2.0, 0, 1)^{\mathrm{T}}$, and $\hat{\boldsymbol{u}}_2 = (1.0, 0, 1)^{\mathrm{T}}$ two estimated optical flow vectors. The two angular errors in this example are $\varepsilon_{\alpha 1} = 7.12°$ and $\varepsilon_{\alpha 2} = 11.3°$. To avoid this effect, we use the absolute magnitude of difference vectors as an error measure

$$\varepsilon_\Delta = \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|, \tag{16}$$

which prevents the above described effects and expresses the individually illustrated difference vectors of Fig. 2 as average values.

Table 1 shows a quantitative comparison between our implemented local optical flow estimation approaches. The left hand side shows the results with selected singular value thresholds used in the preceding Fig. 2. In the right hand side of Table 1, the thresholds were chosen to obtain exactly 50.000 optical flow vectors. The magnitude of the true displacement rate vary between 0.471 and 2.571 pixel per frame with an average magnitude of 1.371 pixel per frame.

Approach	SV Thresh.	# comp. vectors	$\bar{\varepsilon}_\Delta$ [pixel]	$\sigma_{\varepsilon_\Delta}$ [pixel]	SV Thresh.	# comp. vectors	$\bar{\varepsilon}_\Delta$ [pixel]	$\sigma_{\varepsilon_\Delta}$ [pixel]
Werk.'90	0.15	97742	0.369	0.46	0.424	50000	0.228	0.24
$RC_{2,1}$	0.8	113362	0.139	0.18	2.332	50000	0.104	0.09
$SC_{2,1}$	0.8	103675	0.134	0.17	1.985	50000	0.106	0.09
$RC_{3,1}$	0.8	113195	0.127	0.16	2.342	50000	0.097	0.08
$SC_{3,1}$	0.8	115029	0.128	0.17	2.475	50000	0.100	0.07
NA (5×5×5)	100	115208	0.107	0.15	902.1	50000	0.087	0.05

Table 1. Quantitative comparison between local optical flow estimation approaches. The compared estimation approaches are [Werkhoven & Koenderink 90] (Werk'90), the solution under the RC and SC of the $BOFE_{2,1}$, and $BOFE_{3,1}$, and the neighborhood-sampling (NS) method including a spatiotemporal neighborhood of 5×5×5 pixels related to the approach of [Campani & Verri 90]. The next four columns contain the selected threshold of the smallest singular value and the remaining number of thresholded vectors, the average difference vector magnitude and its standard deviation. For the first four columns, the singular value threshold has been selected as for Fig. 2. In contrast to this, the four right columns show the analogous results where a fixed number of optical flow vectors has been compared.

Although the approach of [Werkhoven & Koenderink 90] includes up to 4^{th} order partial derivatives of the gray-value structure, the obtained estimates differ strongly from the true displacement rate. The solutions under the RC and SC of the $BOFE_{2,1}$ and $BOFE_{3,1}$ as well as the neighborhood-sampling method allow comparable optical flow estimation, whereas the last one gives slightly better results. But one has to keep in mind, that the performance of the NS approach is based on 125 equations for 8 unknowns, whereas the approach of [Werkhoven & Koenderink 90] uses only 6 equations for 6 unknowns. It thus does not exploit the advantages of an overdetermined system of equations for the estimation of the optical flow and its derivatives as the other approaches.

4 Acknowledgements

This work was supported in part by the Basic Research Action project INSIGHT II of the European Community. We thank V. Gengenbach for the improvement of the mechanics and calibration of the robot used in these experiments and H. Kollnig for his comments on a draft version of this contribution. We also thank K. Daniilidis for the helpful discussions which stimulated the derivation of the approach of [Werkhoven & Koenderink 90] with the IC.

References

[Aisbett 89] J. Aisbett, Optical Flow with an Intensity-Weighted Smoothing, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-11** (1989) 512–522.

[Baraldi *et al.* 89] P. Baraldi, E. De Micheli, S. Uras, Motion and Depth from Optical Flow, *Proc. Fifth Alvey Vision Conference*, University of Reading, Reading/UK, September 25-28, 1989, pp. 205–208.

[Barron *et al.* 92] J.L. Barron, D. J. Fleet, S. S. Beauchemin, Performance of Optical Flow Techniques, Technical Report RPL-TR-9107, Dept. of Computing Science, Queen's University, Kingston, Ontario, July 1992. (Revised July 1993).

[Campani & Verri 90] M. Campani, A. Verri, Computing Optical Flow from an Overconstrained System of Linear Algebraic Equations, *Proc. Third Int. Conf. on Computer Vision ICCV'90*, Osaka, Japan, Dec. 4-7, 1990, pp. 22–26.

[Chen *et al.* 93] H.-J. Chen, Y. Shirai, M. Asada, Obtaining Optical Flow with Multi-Orientation Filters, *Proc. IEEE Conf. Computer Vision and Pattern Recognition CVPR'93*, New York City, NY, June 15-17, 1993, pp. 736–737.

[Cipolla & Blake 92] R. Cipolla, A. Blake, Surface Orientation and Time to Contact from Image Divergence and Deformation, *Proc. Second European Conference on Computer Vision ECCV'92*, S. Margherita, Italy, May 19-22, 1992, G. Sandini (ed.), Lecture Notes in Computer Science 588, Springer-Verlag, Berlin etc., pp. 187–202.

[Girosi *et al.* 89] F. Girosi, A. Verri, V. Torre, Constraints for the Computation of Optical Flow, *Proc. IEEE Workshop on Visual Motion*, Irvine, CA, March 20-22, 1989, pp. 116–124.

[Horn & Schunck 81] B.K.P. Horn, B.G. Schunck, Determining Optical Flow, *Artificial Intelligence* **17** (1981) 185–203.

[Kearney *et al.* 87] J.K. Kearney, W.B. Thompson, D.L. Boley, Optical Flow Estimation: An Error Analysis of Gradient-Based Methods with Local Optimization, *IEEE Trans. Pattern Analysis and Machine Intelligence* **PAMI-9** (1987) 229–244.

[Koenderink & van Doorn 76] J.J. Koenderink, A.J. van Doorn, Local Structure of Movement Parallax of the Plane, *Journal of the Optical Society of America* **66** (1976) 717–723.

[Nagel 85] H.-H. Nagel, Analyse und Interpretation von Bildfolgen, *Informatik-Spektrum* **8** (1985) 178–200 und 312–327.

[Nagel 87] H.-H. Nagel, On the Estimation of Optical Flow: Relations between Different Approaches and Some New Results, *Artificial Intelligence* **33** (1987) 299–324.

[Nagel 92] H.H. Nagel, Direct Estimation of Optical Flow and its Derivatives, in *Artificial and Biological Vision Systems*, G.A Orban and H.-H. Nagel (eds.), Springer-Verlag Berlin etc., 1992, pp. 193–224.

[Negahdaripour & Lee 92] S. Negahdaripour, S. Lee, Motion Recovery from Image Sequences Using Only First Order Optical Flow Information, *International Journal of Computer Vision* **9** (1992) 163–184.

[Negahdaripour & Yu 93] S. Negahdaripour, C.-H. Yu, A Generalized Brightness Change Model for Computing Optical Flow, *Proc. Fourth Int. Conf. on Computer Vision ICCV'93*, Berlin, Germany, May 11-14, 1993, pp. 2–11.

[Nelson & Aloimonos 88] R.C. Nelson, J. Aloimonos, Using Flow Field Divergence for Obstacle Avoidance: Towards Qualitative Vision, *Proc. Second Int. Conf. on Computer Vision ICCV'88*, Tampa, FL, Dec. 5-8, 1988, pp. 188–196.

[Otte 94] M. Otte, Extraktion von linienförmigen Merkmalen und Ermittlung des optischen Flusses mit seinen Ableitungen aus Voll- und Halbbildfolgen, eingereichte Dissertation, Institut für Algorithmen und Kognitive Systeme, Fakultät für Informatik der Universität Karlsruhe (TH), Karlsruhe, Deutschland, Januar 1994.

[Schunck 84] B.G. Schunck, The Motion Constraint Equation for Optical Flow, *Proc. Int. Conf. on Pattern Recognition ICPR'84*, Montréal, Canada, July 30 - Aug. 2, 1984, pp. 20–22.

[Srinivasan 90] M.V. Srinivasan, Generalized Gradient Schemes for the Measurement of Two-Dimensional Image Motion, *Biological Cybernetics* **63** (1990) 421–431.

[Subbarao 90] M. Subbarao, Bounds on Time-to-Collision and Rotational Component from First-Order Derivatives of Image Flow, *Computer Vision, Graphics, and Image Processing* **50** (1990) 329–341.

[Verri & Poggio 89] A. Verri, T. Poggio, Motion Field and Optical Flow: Qualitative Properties, *IEEE Trans. Pattern Analysis and Machine Intelligence* **PAMI-11** (1989) 490–498.

[Weber & Malik 93] J. Weber, J. Malik, Robust Computation of Optical Flow in a Multi-Scale Differential Framework, *Proc. Fourth Int. Conf. on Computer Vision ICCV'93*, Berlin, Germany, May 11-14, 1993, pp. 12–20.

[Werkhoven & Koenderink 90] P. Werkhoven, J.J. Koenderink, Extraction of Motion Parallax Structure in the Visual System, *Biological Cybernetics* **63** (1990) 185–199.