Motion Boundary Detection in Image Sequences by Local Stochastic Tests

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

[1] Fraunhofer-Institut für Informations- und Datenverarbeitung IITB
Fraunhoferstr. 1, D-76131 Karlsruhe
Telephone +49 (721) 6091-210 (Fax -413)
[2] Institut für Algorithmen und Kognitive Systeme
Fakultät für Informatik der Universität Karlsruhe

Abstract. While estimating both components of optical flow based on the postulated validity of the Optical Flow Constraint Equation (OFCE), it has been tacitly assumed so far that the partial derivatives of the gray value distribution - which are required for this approach at the pixel positions involved - are independent from each other. [Nagel 94] has shown in a theoretical investigation how dropping this assumption affects the estimation procedure. The advantage of such a more rigorous approach consists in the possibility to replace heuristic tests for the local detection of discontinuities in optical flow fields by well known stochastic tests. First results from various experiments with this new approach are presented and discussed.

1 Introduction

Most attempts to estimate optical flow so far took only partial account of the stochastic nature of the observed digitized image sequences. As a consequence, heuristic procedures to cope with signal variations as well as 'intuitive' choices of threshold parameters have been used to select results upon which further inferences had to be based. A careful discussion about the assumptions underlying various approaches towards the estimation of optical flow can be found, too, in [Fleet 92] and in [Nagel 92]. [Barron et al. 92] offer an in-depth comparison between various approaches. In the following, therefore, we concentrate on recent approaches towards the segmentation of estimated optical flow fields.

In cases of relative motion between camera and scene background, one can attempt to segment the field of view based on significant discontinuities in the optical flow field. [Spoerri & Ullman 87] discuss local computations to detect motion boundaries. [Bouthemy & Francois 93] and [Irani et al. 92] have exploited the evaluation of more than two or three consecutive frames from an image sequence in order to acquire clues to regions corresponding to the images of moving objects. [Thompson & Pong 90] compare - in qualitative terms - various approaches towards the segmentation of images of moving objects from the image of moving background. A valuable update on detecting images of moving objects can be found in [Letang et al. 93]. More recent approaches have been reported by [Wang & Adelson 93], [Gu et al. 93], [Torr & Murray 92], and [Kollnig et al.

Lecture Notes in Computer Science, Vol. 801
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

94]. None of these approaches investigates the influence of uncertainties in the raw digitized gray values on the final segmentation results in a stringent manner.

[Negahdaripour & Lee 92] tessellate frames from an image sequence and estimate optical flow as well as its first spatial derivatives within each image region by a fit to the Optical Flow Constraint Equation (OFCE), see [Horn 86]. They do not, however, take into account that the estimates for the spatio-temporal gray value gradient at neighboring pixel positions are not independent from each other.

In most cases, the assumption is not even made explicit that the partial derivatives $\partial g(x,y,t)/\partial x$, $\partial g(x,y,t)/\partial y$, and $\partial g(x,y,t)/\partial t$ of the gray value distribution $g(x,y,t)$ at neighboring pixel positions are treated as independent estimates. Each such estimate is tacitly associated with an error taken to be an independently identically distributed (i.i.d.) sample from a Gaussian and then used in a Least Squared Error approach - see, for example, [Burt et al. 91], [Campani & Verri 92], [Torr & Murray 92], and [Otte & Nagel 94]. [Gu et al. 93], however, at least mention this assumption explicitly in their study to segment images of moving objects from those of moving background, see also [Etoh & Shirai 93]. Similarly, [Weber & Malik 93] emphasize the fact that all partial derivatives of the gray value distribution used for estimating optical flow are in principle corrupted by noise, not only the temporal derivative. They treat, however, each estimated spatio-temporal partial derivative of the gray value distribution as being an independent measurement.

Some authors explicitly posit that the deviation from an exact validity of the OFCE at each pixel position is i.i.d. Gaussian, for example [Bouthemy & Francois 93]. An analogous remark is made by [Simoncelli et al. 91]. Similarly, [Chou & Chen 93] point out that noise could be taken into account by associating their version of the OFCE with a suitably chosen noise term. [Bouthemy & Santillana Rivero 87] postulated that the difference between a constant optical flow for a region and a local estimate, projected onto the local gradient direction, is i.i.d. Gaussian. [De Micheli et al. 93] pointed out that in case the Hessian is well conditioned, the technique to estimate optical flow from second order spatial derivatives of the gray value distribution 'can be computed almost independently at any pixel position' - if one neglects the correlation between adjacent estimates introduced by a smoothing step in their approach.

In his study of uncertainty in Low-Level Vision, [Szeliski 89], [Szeliski 90] does not treat the influence of uncertainties of the individual gray values on the uncertainty of the resulting optical flow estimates. His approach clearly shows that i.i.d. Gaussian noise is assumed for each estimate of the gradient at a location $\boldsymbol{x}$, but the interrelationship between the various gray values contributing to gradient estimates at adjacent pixel positions is not treated in detail.

Based on well established approaches of estimation theory, [Nagel 94] describes an estimation of optical flow vector fields which takes the dependencies between estimates of partial derivatives of gray values into account.

2 Notation

A vector in the spatio-temporal domain of image plane location (x,y) at time t

is denoted by a bold face character, i.e. $\boldsymbol{x} = (x, y, t)^T = (x_1, x_2, x_3)^T$ where the latter notation is used if a coordinate should be referenced by a running index.

It is assumed that the irradiance distribution $g(x, y, t)$ impinging at the image plane location (x, y) at time t is transduced into a gray value $g_{observed}(x, y, t)$. The continuum $\boldsymbol{x} = (x, y, t)^T$ is sampled at grid position $\boldsymbol{x}_{k_{sample}} = (x_{k_{sample}}, y_{k_{sample}}, t_{k_{sample}})^T$ with $n_{sample} = n_{x_{sample}} \times n_{y_{sample}} \times n_{t_{sample}}$ and $k_{sample} \in \{1, ..., n_{sample}\}$. Let $g(\boldsymbol{x}_{k_{sample}})$ denote the undistorted irradiance value at the raster position $\boldsymbol{x}_{k_{sample}}$. The sensing process is assumed to distort the incoming signal by noise $\delta g(\boldsymbol{x}_{k_{sample}})$ which is taken to be independently identically distributed with zero mean and variance σ_g^2, thus

$$g_{observed}(\boldsymbol{x}_{k_{sample}}) = g(\boldsymbol{x}_{k_{sample}}) + \delta g(\boldsymbol{x}_{k_{sample}}) \quad . \tag{1}$$

The pixel position at the center of the sampling area is indicated by the position vector $\boldsymbol{x}_{0_{test}} = (x_{0_{test}}, y_{0_{test}}, t_{0_{test}})^T$ with $t_{0_{test}} = t_0$. We select a subset from the digitized image around location $\boldsymbol{x}_{0_{test}}$ in order to estimate the optical flow vector. This subset is denoted as the test area around $\boldsymbol{x}_{0_{test}}$ and comprises - for example - a cube of pixels around the central pixel $\boldsymbol{x}_{0_{test}}$ in the image plane at time t_0. The other pixel positions are serialized in raster scan order starting at the northwest position of $\boldsymbol{x}_{0_{test}}$ - but in the preceding frame - with $j_{test} = 1, ..., n_{test} - 1$, skipping $\boldsymbol{x}_{0_{test}}$ since this center pixel of the test area is taken to be the first in this serialization.

It is postulated that the structure of the impinging irradiance distribution is such that the Optical Flow Constraint Equation (OFCE) [Horn 86] is satisfied at location $\boldsymbol{x}_{test}$. Although other constraint equations - see, e.g., [Negahdaripour & Yu 93] - could be introduced, the emphasis of this study can be demonstrated based on the well-known OFCE. We thus have

$$u_1(\boldsymbol{x}_{test}) \left.\frac{\partial g(\boldsymbol{x})}{\partial x}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}} + u_2(\boldsymbol{x}_{test}) \left.\frac{\partial g(\boldsymbol{x})}{\partial y}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}} + \left.\frac{\partial g(\boldsymbol{x})}{\partial t}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}} = 0 \quad . \tag{2}$$

Moreover, it is assumed that the optical flow vector $\mathbf{u}(\boldsymbol{x})$ varies only linearly within the test area, i.e.

$$\begin{pmatrix} u_1(\boldsymbol{x}_{test}) \\ u_2(\boldsymbol{x}_{test}) \end{pmatrix} = \begin{pmatrix} u_1(\boldsymbol{x}_{0_{test}}) \\ u_2(\boldsymbol{x}_{0_{test}}) \end{pmatrix} + \begin{pmatrix} \left.\frac{\partial u_1(\boldsymbol{x})}{\partial x}\right|_{\boldsymbol{x}=\boldsymbol{x}_{0_{test}}} & \left.\frac{\partial u_1(\boldsymbol{x})}{\partial y}\right|_{\boldsymbol{x}=\boldsymbol{x}_{0_{test}}} & \left.\frac{\partial u_1(\boldsymbol{x})}{\partial t}\right|_{\boldsymbol{x}=\boldsymbol{x}_{0_{test}}} \\ \left.\frac{\partial u_2(\boldsymbol{x})}{\partial x}\right|_{\boldsymbol{x}=\boldsymbol{x}_{0_{test}}} & \left.\frac{\partial u_2(\boldsymbol{x})}{\partial y}\right|_{\boldsymbol{x}=\boldsymbol{x}_{0_{test}}} & \left.\frac{\partial u_2(\boldsymbol{x})}{\partial t}\right|_{\boldsymbol{x}=\boldsymbol{x}_{0_{test}}} \end{pmatrix} \cdot (\boldsymbol{x}_{test} - \boldsymbol{x}_{0_{test}})$$

or – with the abbreviated notation $(u_1(\boldsymbol{x}_{0_{test}}), u_2(\boldsymbol{x}_{0_{test}}))^T = (u_1, u_2)^T$ –

$$\begin{pmatrix} u_1(\boldsymbol{x}_{test}) \\ u_2(\boldsymbol{x}_{test}) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} u_{1x} & u_{1y} & u_{1t} \\ u_{2x} & u_{2y} & u_{2t} \end{pmatrix} \cdot (\boldsymbol{x}_{test} - \boldsymbol{x}_{0_{test}}) \quad . \tag{3}$$

The partial derivatives of $g(\boldsymbol{x})$ with respect to x, y, and t are computed by discrete convolution of $g(\boldsymbol{x}_{k_{sample}})$ with discretized versions of the corresponding partial derivatives of a trivariate spatiotemporal Gaussian distribution

$$w(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{|\Sigma_w|}} \exp^{-\frac{1}{2}\boldsymbol{x}^T \Sigma_w^{-1} \boldsymbol{x}} \quad , \tag{4}$$

$$\left.\frac{\partial g(\boldsymbol{x})}{\partial x_i}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}} = \iiint d\xi \cdot \frac{\partial w(\xi)}{\partial \xi_i} g(\boldsymbol{x}_{test} - \xi) = \iiint d\xi \cdot \left.\frac{\partial w(\boldsymbol{x}-\xi)}{\partial x_i}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}} g(\xi) \; . \tag{5}$$

The samples of $\left.\frac{\partial w(\boldsymbol{x}-\xi)}{\partial x_i}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}}$ are arranged into column vectors

$$\mathbf{w}_i(\boldsymbol{x}_{test}) = \left(\left.\frac{\partial w(\boldsymbol{x} - \boldsymbol{x}_{k_{sample}})}{\partial x_i}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}} \right) = \left(w_{ik_{sample}}(\boldsymbol{x}_{test}) \right) \tag{6}$$

where $i = 1, 2, 3$ corresponds to x, y, t, respectively, and $k_{sample} = 1, ..., n_{sample}$. The convolution can then be written as the scalar product

$$\left.\frac{\partial g(\boldsymbol{x})}{\partial x_i}\right|_{\boldsymbol{x}=\boldsymbol{x}_{test}} = \mathbf{w}_i(\boldsymbol{x}_{test})^T \mathbf{g} = \sum_{k_{sample}=1}^{n_{sample}} w_{ik_{sample}}(\boldsymbol{x}_{test}) \cdot g_{k_{sample}}, \; i = 1, 2, 3 \; . \tag{7}$$

Since we shall have to study the optical flow vector $(u_1(\boldsymbol{x}_{test}),\ u_2(\boldsymbol{x}_{test}))^T$ as some function of all n_{test} pixel positions $\boldsymbol{x}_{test}$ in the test area, it will be convenient to index these positions and to write

$$\mathbf{w}_i(\boldsymbol{x}_{j_{test}}) = \mathbf{w}_{ij_{test}} = \left(w_{ik_{sample}}(\boldsymbol{x}_{j_{test}})\right) = \left(w_{ij_{test}k_{sample}}\right)$$

$$i = 1, 2, 3 \; ; \; j_{test} = 0, 1, ..., n_{test} - 1; \; k_{sample} = 1, 2, ..., n_{sample}. \tag{8}$$

3 Estimation Problem and Solution Approach

We may now write the OFCEs as a set of n_{test} equations

$$u_{1j_{test}}(\mathbf{w}_{1j_{test}}^T \mathbf{g}) + u_{2j_{test}}(\mathbf{w}_{2j_{test}}^T \mathbf{g}) + (\mathbf{w}_{3j_{test}}^T \mathbf{g}) = 0 \qquad \text{or}$$

$$(u_{1j_{test}}\mathbf{w}_{1j_{test}} + u_{2j_{test}}\mathbf{w}_{2j_{test}} + \mathbf{w}_{3j_{test}})^T \mathbf{g} = 0, \; j_{test} = 0, ..., n_{test} - 1. \tag{9}$$

This formulation exhibits the linear dependence of the n_{test} OFCEs on the set of all samples from the irradiance field within the sampling area. It will be advantageous to consider this set of n_{test} OFCEs as a column vector with n_{test} components, where each component depends on the gray values $\mathbf{g}$ and on the set of $n_{parameter} = 8$ parameters written as components of

$$\mathbf{u} = (u_1, u_{1x}, u_{1y}, u_{1t}, u_2, u_{2x}, u_{2y}, u_{2t})^T \tag{10}$$

We then have for each component $c_{j_{test}}(\mathbf{g}, \mathbf{u})$ of this column vector

$$c_{j_{test}}(\mathbf{g}, \mathbf{u}) = \left\{ \left[u_1 + \begin{pmatrix} u_{1x} \\ u_{1y} \\ u_{1t} \end{pmatrix}^T (\boldsymbol{x}_{j_{test}} - \boldsymbol{x}_{0_{test}}) \right] \mathbf{w}_{1j_{test}} + \right. \tag{11}$$

$$\left. \left[u_2 + \begin{pmatrix} u_{2x} \\ u_{2y} \\ u_{2t} \end{pmatrix}^T (\boldsymbol{x}_{j_{test}} - \boldsymbol{x}_{0_{test}}) \right] \mathbf{w}_{2j_{test}} + \mathbf{w}_{3j_{test}} \right\}^T \mathbf{g} = 0 ,$$

$$j_{test} = 0, 1, 2, ..., n_{test} - 1 \quad .$$

Since by assumption - see equ. (1) - we can write $\mathbf{g}_{observed} = \mathbf{g} + \delta\mathbf{g}$ where the error vector $\delta\mathbf{g}$ is assumed to be distributed according to

$$p(\delta\mathbf{g}) = \frac{1}{(2\pi\sigma_g^2)^{n_{sample}/2}} \exp^{-\frac{1}{2}\delta\mathbf{g}^T \Sigma_g^{-1} \delta\mathbf{g}} \tag{12}$$

with - $I_{n\times n}$ denoting the $n \times n$ unity matrix -

$$\Sigma_g = \sigma_g^2 \cdot I_{n_{sample}\times n_{sample}} . \tag{13}$$

We are now in a position to precisely formulate the estimation problem: given the assumptions introduced above, determine the estimate $\hat{\mathbf{u}}$ for the parameter vector $\mathbf{u}$ which maximizes the joint probability to observe $\delta\mathbf{g}$, subject to the constraints $c_{j_{test}}(\mathbf{g}, \mathbf{u}) = 0$ for $j_{test} = 0, 1, ...n_{test-1}$. This is equivalent to minimizing

$$\delta\mathbf{g}^T \Sigma_g^{-1} \delta\mathbf{g} + 2\boldsymbol{\lambda}^T \left(c_{j_{test}}(\mathbf{g}, \mathbf{u}) \right) \implies \text{minimum} \tag{14}$$

by suitable choices of $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, ..., \lambda_{n_{test}-1})^T$ and of the components of $\mathbf{u}$. This is a non-linear constraint problem which is solved by an iterative approach - see [Nagel 94]. As usual, it is assumed that all higher than first order terms in the constraint equations can be neglected, i.e.

$$c_{j_{test}}(\mathbf{g}, \mathbf{u}) = c_{j_{test}} \left(\mathbf{g}_{(k)} + \Delta_{(k)}\mathbf{g} , \mathbf{u}_{(k)} + \Delta_{(k)}\mathbf{u} \right) \tag{15}$$

$$= c_{j_{test}}(\mathbf{g}_{(k)}, \mathbf{u}_{(k)}) + \left(\left. \frac{\partial c_{j_{test}}}{\partial \mathbf{g}} \right|_{\mathbf{g}_{(k)}, \mathbf{u}_{(k)}} \cdot \Delta_{(k)}\mathbf{g} \right)$$

$$+ \left(\left. \frac{\partial c_{j_{test}}}{\partial \mathbf{u}} \right|_{\mathbf{g}_{(k)}, \mathbf{u}_{(k)}} \cdot \Delta_{(k)}\mathbf{u} \right) = 0$$

where (k) denotes an iteration index, $k = 0, 1, ...$, and $\Delta_{(k)}\mathbf{g}$ as well as $\Delta_{(k)}\mathbf{u}$ are assumed to be small corrections so that higher than first powers of the components of $\Delta_{(k)}\mathbf{g}$ and $\Delta_{(k)}\mathbf{u}$ can be neglected.

It is further assumed that a start estimate $\hat{\mathbf{u}}_{(0)}$ for $\mathbf{u}$ is available and that the observed gray values $\mathbf{g}_{observed}$ may serve as start values for the undistorted gray values $\mathbf{g}$ to be determined during the iterative estimation procedure, i.e. $\mathbf{g}_{(0)} = \mathbf{g}_{observed}$. In order to simplify the notation, the iteration index (k) will be suppressed in subsequent expressions unless it is explicitly manipulated. The following abbreviations are introduced :

$$\mathbf{c} = \left(c_{j_{test}}(\mathbf{g}_{(k)}, \mathbf{u}_{(k)}) \right) \text{ , a } n_{test} \times 1 \qquad \text{vector} \quad , \tag{16a}$$

$$C_g = \left(\left. \frac{\partial c_{j_{test}}}{\partial \mathbf{g}} \right|_{\mathbf{g}_{(k)},\, \mathbf{u}_{(k)}} \right) \text{ , a } n_{test} \times n_{sample} \qquad \text{matrix} \quad , \tag{16b}$$

$$C_u = \left(\left. \frac{\partial c_{j_{test}}}{\partial \mathbf{u}} \right|_{\mathbf{g}_{(k)},\, \mathbf{u}_{(k)}} \right) \text{ , a } n_{test} \times n_{parameter} \text{ matrix} \quad . \tag{16c}$$

The minimization problem can now be rewritten as:

$$\Delta\mathbf{g}^T \Sigma_g^{-1} \Delta\mathbf{g} + 2\boldsymbol{\lambda}^T (\mathbf{c} + C_g \Delta\mathbf{g} + C_u \Delta\mathbf{u}) \implies \text{minimum} \quad . \tag{17}$$

If we define

$$\Sigma_u = \left(C_u^T \left(C_g \Sigma_g C_g^T \right)^{-1} C_u \right)^{-1} \quad , \tag{18}$$

the solution to this constrained minimization problem is given in [Nagel 94] :

$$\Delta\hat{\mathbf{g}} = -\Sigma_g C_g^T \left(C_g \Sigma_g C_g^T \right)^{-1} \left[I - C_u \Sigma_u C_u^T \left(C_g \Sigma_g C_g^T \right)^{-1} \right] \mathbf{c} \quad , \tag{19}$$

$$\Delta\hat{\mathbf{u}} = -\Sigma_u C_u^T \left(C_g \Sigma_g C_g^T \right)^{-1} \mathbf{c} \quad , \tag{20}$$

$$\hat{\boldsymbol{\lambda}} = \left(C_g \Sigma_g C_g^T \right)^{-1} \left[I - C_u \Sigma_u C_u^T \left(C_g \Sigma_g C_g^T \right)^{-1} \right] \mathbf{c} \quad . \tag{21}$$

We thus have

$$\hat{\mathbf{g}}_{(k+1)} = \hat{\mathbf{g}}_{(k)} + \Delta\hat{\mathbf{g}}_{(k)} \qquad \text{for } k = 0, 1, \ldots \tag{22}$$

with $\mathbf{g}_{(0)} = \mathbf{g}_{observed}$, $\Delta\hat{\mathbf{g}}_{(0)} = 0$,

$$\hat{\mathbf{u}}_{(k+1)} = \hat{\mathbf{u}}_{(k)} + \Delta\hat{\mathbf{u}}_{(k)} \qquad \text{for } k = 0, 1, \ldots \tag{23}$$

with $\Delta\hat{\mathbf{u}}_{(0)} = 0$, and $\hat{\mathbf{u}}_{(0)} = -(G^T G)^{-1} G^T \mathbf{g}_t$. The matrix G is built from partial spatial derivatives of the grayvalues - see [Nagel 94]. The iteration will be stopped if

$$\| \Delta\hat{\mathbf{u}}_{(k+1)} - \Delta\hat{\mathbf{u}}_{(k)} \| < \text{ iteration threshold} \quad . \tag{24}$$

As is well known, $\Delta\hat{\mathbf{g}}^T \Sigma_g^{-1} \Delta\hat{\mathbf{g}}$ follows a χ^2-distribution with $(n_{test} - n_{parameter})$ degrees of freedom.

If we possess an estimate of σ_g^2, for example from measurements, we can compute

$$\Delta\hat{\mathbf{g}}^T \Sigma_g^{-1} \Delta\hat{\mathbf{g}} = \sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}} \tag{25}$$

and test whether this result is compatible with the assumptions: if

$$\sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}} > \chi^2(n_{test} - n_{parameter}; 1-\alpha) \ , \tag{26}$$

then the estimate $\hat{\mathbf{u}}$ is rejected, the estimated error being incompatible at a confidence level of $1-\alpha$ with the assumptions made.

As can be seen from equ. 20, the inverse of Σ_u - i.e. $C_u^T \left(C_g \Sigma_g C_g^T\right)^{-1} C_u$ - has to have sufficiently large eigenvalues in order to ensure that it can be inverted without numerical problems. We thus may require that the smallest eigenvalue of Σ_u^{-1} does exceed a threshold.

We are now able to detect algorithmically not only instability due to insufficient gray value variations. In addition, we may use equ. (26) in order to perform a test on inappropriately large variations of the gray value distribution which will result in an excessively high value for $\sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}}$ and thus can be detected with a given confidence $1-\alpha$.

As a result of this analysis, we thus have tests for both situations, inappropriately small as well as inappropriately large variations of the gray value distribution within the test area ! These tests can, of course, be easily extended for larger and less regular areas than discussed so far.

4 Results

Quantitative experiments with the approach outlined in the preceding section have been performed with image frames from a sequence recorded by a camera on the moving arm of a calibrated robot - see Figure 1 [Otte & Nagel 94].

Since the χ^2-distribution requires the knowledge of the variance of the original measurements, a subseries of 29 image frames has been recorded at a fixed position of the robot hand with the illumination kept stationary, yielding 3.6 as the average gray value variance. Optical flow vectors estimated according to the

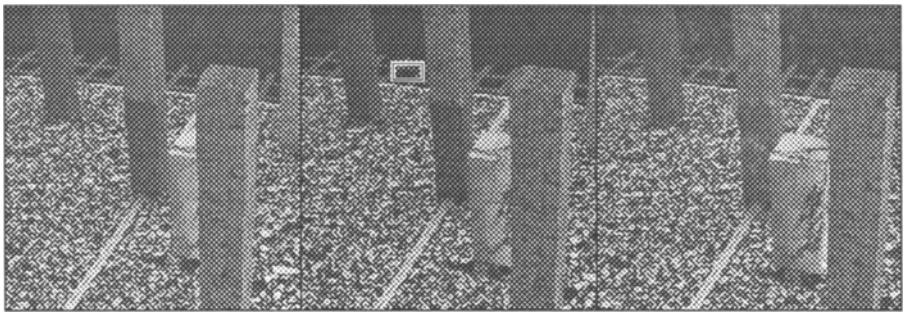

Fig. 1. Three frames from a sequence of non-interlaced images. The video camera translates into the scene. In addition, the white marble parallelepiped is moving to the left with respect to the rest of the scene.

approach of [Nagel 94] outlined in the preceding section in an image area where we do not expect discontinuities of the optical flow field do not differ significantly from those obtained according to a version of the approach of [Campani & Verri 92] extended by the inclusion of temporal derivatives of the optical flow. Differences are significant only at those locations where the estimation has been suppressed in the approach of [Nagel 94] due to insufficient structure in the local gray value variation.

It turned out, however, that in image areas with sufficient structure to determine all unknown components of the vector $\mathbf{u}$, the estimated gray value variances are greater than expected. These image areas are selected by requiring that the smallest singular value exceeds 0.3 as a threshold. If we plot the histogram of $\Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}} / (4.2)^2$ instead of $\sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}}$ with $\sigma_g^2 = (1.902)^2$, we observe acceptable compatibility with a χ^2-distribution for 3 degrees of freedom. The standard deviation of the gray values had to be increased, however, by a factor of roughly 2.3 compared to the measured standard deviation in order to shift the maximum of the histogram to the expected position. This observation could be taken as a hint that the OFCE imposes constraints onto the recorded spatiotemporal data which are not quite compatible with the true variations. It should not come as a surprise that these constraints make themselves felt more strongly in image areas with sufficient spatiotemporal variation to facilitate the estimation of all components of the unknown vector $\mathbf{u}$ than in image areas with only small variations.

Significant differences in the histogram of $\sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}}$ can be observed if we compare one obtained from image areas where we do not expect discontinuities of the optical flow and a histogram computed for an image area in which discontinuities of the optical flow field occur. The tail of $\sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}}$ extends to much larger values in the latter case. A threshold applied to such χ^2-values appears as a suitable local means to discriminate discontinuities in the optical flow field from image areas with a smooth variation of the optical flow field. This hypothesis is well supported by the results shown in Figure 2. Image locations with large values of $\sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}}$ clearly cluster along the expected lines of discontinuity - in this case caused by both depth discontinuities and, in addition, by the movement of the white marble block relative to the scene background which itself moves relative to the camera due to the camera motion on the robot arm.

Figure 3 illustrates how moving object contours can be extracted in real world image sequences without depending on a-priori knowledge about a stationary camera. Significant contour segments of the moving bus can be detected solely by the requirement that χ^2 exceeds a threshold, in this case of 48. For a gray value standard deviation of $\sigma_g = 1.902$, the χ^2-threshold at the 99% confidence level is 11.37. The estimate $\sigma_g = 1.902$ for the standard deviation of the observed gray values has been obtained by recording with a CCD-camera whereas the image sequence 'Ettlinger-Tor-Platz' has been recorded with an older tube TV-camera. We thus expect that analogous noise measurements for this latter camera would yield larger values than those obtained with the CCD-camera under controlled

laboratory conditions. This would allow to lower the threshold for χ^2 in the 'Ettlinger-Tor-Platz' image sequence from 48 to some more reasonable value without serious deterioration of the detection ability.

5 Conclusion

Experimental evidence has been presented for the hypothesis that taking the stochastic aspects of gray value variation into account for the estimation of an optical flow field facilitates the segmentation of such a field based solely on a χ^2-test - even if the background should change due to camera motion. The pixel positions at which such a test will fail - and thus generate a cue towards the presence of a discontinuity in the estimated optical flow field - cluster around discontinuities. The width of such clusters depends on the size of the masks which are used to estimate the spatiotemporal derivatives of the gray value distribution. In the cases presented here, these masks had a width of seven pixels. It can be seen in the last Figure that the χ^2-test tends to emphasize the contour lines of the bus. Obviously, further experiments are required in order to consolidate the preliminary conclusions presented in this contribution.

6 Acknowledgements

Partial support of these investigations by the ESPRIT Basic Research Action project INSIGHT II as well as by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the Sonderforschungsbereich 'Künstliche Intelligenz - Wissensbasierte Systeme' are gratefully acknowledged.

References

[Barron et al. 92] Barron, J.L., Fleet, D.J., Beauchemin, S.S., and Burkitt, T.A.: Performance of Optical Flow Techniques. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR '92*, 15-18 June 1992, Champaign, IL, pp. 236-242. See, too, Int. Journal of Computer Vision 12:1 (1994) 43-77.

[Bouthemy & Santillana Rivero 87] Bouthemy, P., and Santillana Rivero, J.: A Hierarchical Likelihood Approach for Region Segmentation According to Motion-Based Criteria. In *Proc. First Intern. Conference on Computer Vision ICCV '87*, London, UK, 8-11 June 1987, pp. 463-467.

[Bouthemy & Francois 93] Bouthemy, P., and Francois, E.: Motion Segmentation and Qualitative Scene Analysis from an Image Sequence. Int. Journal of Computer Vision 10:2 (1993) 157-182

[Burt et al. 91] Burt, P.J., Hingorani, R., and Kolczynski, R.: Mechanisms for Isolation Component Patterns in the Sequential Analysis of Multiple Motion. In *Proc. IEEE Workshop on Visual Motion*, Princeton, NJ, 7-9 October 1991, pp. 187-193.

[Campani & Verri 92] Campani, M., and Verri, A.: Motion Analysis from First Order Properties of Optical Flow. CVGIP: Image Understanding 56:1 (1992) 90-107.

[Chou & Chen 93] Chou, W.-S., and Chen, Y.-C.: Estimation of the Velocity Field of Two-Dimensional Deformable Motion. Pattern Recognition 26:2 (1993) 351-364.

[De Micheli et al. 93] De Micheli, E., Torre, V., and Uras, S.: The Accuracy of the Computation of Optical Flow and of the Recovery of Motion Parameters. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-15:5 (1993) 434-447.

[Etoh & Shirai 93] Etoh, M., and Shirai, Y.: Segmentation and 2D Motion Estimation by Region Fragments. In *Proc. Fourth Intern. Conference on Computer Vision ICCV '93*, 11-14 May 1993, Berlin, Germany, pp. 192-199.

[Fleet 92] Fleet, D.J.: Measurement of Image Velocity. Kluwer Academic Publishers: Boston, MA; London, UK; Dordrecht, NL, 1992

[Gu et al. 93] H. Gu, M. Asada, and Y. Shirai: The Optimal Partition of Moving Edge Segments. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR '93*, 15-17 June 1993, New York City, NY, pp. 367-372.

[Horn 86] Horn, B.K.P.: Robot Vision. The MIT Press: Cambridge, MA, 1986

[Irani et al. 92] Irani, M., Rousso, B., and Peleg, S.: Detecting and Tracking Multiple Moving Objects Using Temporal Integration. In *Proc. Second European Conference on Computer Vision ECCV '92*, Santa Margherita Ligure, Italy, 18-23 May 1992, Lecture Notes in Computer Science 588, G. Sandini (ed.), Springer-Verlag: Berlin Heidelberg New York and others, pp. 282-287.

[Kollnig et al. 94] H. Kollnig, H.-H. Nagel, and M. Otte: Association of Motion Verbs with Vehicle Movements Extracted from Dense Optical Flow Fields. Proc. ECCV '94, Stockholm / Sweden, 2-6 May 1994.

[Letang et al. 93] Letang, J.M., Rebuffel, V., and Bouthemy, P.: Motion Detection Robust to Perturbations: a Statistical Regularization and Temporal Integration Framework. In *Proc. Fourth Intern. Conference on Computer Vision ICCV '93*, 11-14 May 1993, Berlin, Germany, pp. 21-30.

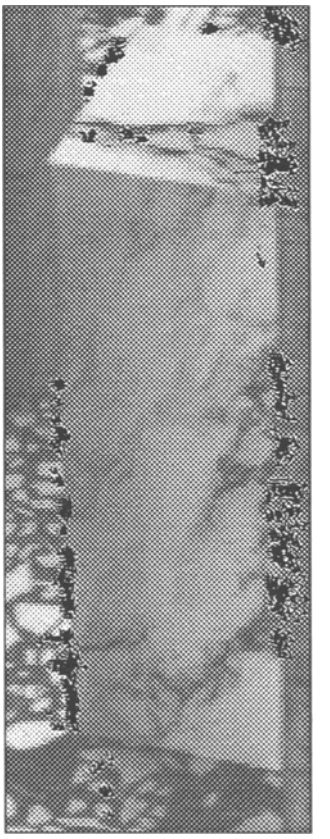

Fig. 2. A black dot is superimposed on the original image at all locations where $\sigma_g^{-2} \Delta\hat{\mathbf{g}}^T \Delta\hat{\mathbf{g}}$ exceeds a threshold of 100 for $\sigma_g^2 = 3.6$. This demonstrates the possibility to detect discontinuities in optical flow fields by local methods, here the boundaries of the marbled epiped which moves with respect to the rest of the field of view recorded by a camera moving itself. Obviously, the local gray value variation must be significant enough to facilitate estimation of an optical flow vector by a local approach.

[Nagel 92] Nagel, H.-H.: Direct Estimation of Optical Flow and of Its Derivatives. In *Artificial and Biological Vision Systems*, G. Orban and H.-H. Nagel (eds.). Springer-Verlag: Berlin Heidelberg New York and others, 1992, pp. 193-224.

[Nagel 94] Nagel, H.-H.: Optical Flow Estimation and the Interaction Between Measurement Errors at Adjacent Pixel Positions. Int. Journal of Computer Vision, to appear 1994.

[Negahdaripour & Lee 92] Negahdaripour, S., and Lee, S.: Motion Recovery from Image Sequences Using Only First Order Optical Flow Information. Int. Journal of Computer Vision 9:3 (1992) 163-184.

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

[Otte & Nagel 94] M. Otte and H.-H. Nagel: Optical Flow Estimation: Advances and Comparisons. Proc. ECCV '94, Stockholm / Sweden, 2-6 May 1994.

[Simoncelli et al. 91] Simoncelli, E.P., Adelson, E.H., and Heeger, D.J.: Probability Distributions of Optical Flow. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Lahaina, Maui, Hawaii, 3-6 June 1991, pp. 310-315.

[Spoerri & Ullman 87] Spoerri, A., and Ullman, S.: The Early Detection of Motion Boundaries. In *Proc. First Intern. Conference on Computer Vision ICCV '87*, London, UK, 8-11 June 1987, pp. 209-218.

[Szeliski 89] Szeliski, R.: Bayesian Modeling of Uncertainty in Low-level Vision. Kluwer Academic Publishers: Boston, MA; Dordrecht, NL; London, UK, 1989

[Szeliski 90] Szeliski, R.: Bayesian Modeling of Uncertainty in Low-Level Vision. Int. Journal of Computer Vision 5:3 (1990) 271-301.

[Thompson & Pong 90] Thompson, W.B., and Pong, T.-C.: Detecting Moving Objects. Intern. Journal of Computer Vision 4:1 (1990) 39-57.

[Torr & Murray 92] Torr, P.H.S., and Murray, D.W.: Statistical Detection of Independent Movement from a Moving Camera. In *Proc. British Machine Vision Conference*, Leeds, UK, 22-24 Sept. 1992, D. Hogg and R. Boyle (eds.), Springer-Verlag: London Berlin Heidelberg and others, pp. 79-88. See, too, Image and Vision Computing 11:4 (1993) 180-187.

[Wang & Adelson 93] Wang, J.Y.A., and Adelson, E.H.: Layered Representation for Motion Analysis. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR '93*, 15-17 June 1993, New York City, NY, pp. 361-366.

[Weber & Malik 93] Weber, J., and Malik, J.: Robust Computation of Optical Flow in a Multiscale Differential Framework. In *Proc. Fourth Intern. Conference on Computer Vision ICCV '93*, 11-14 May 1993, Berlin, Germany, pp. 12-20.

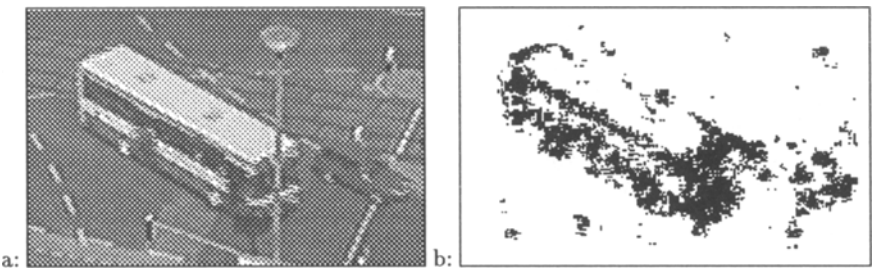

Fig. 3. (a) Enlarged subregion from a sequence recorded at an intersection in downtown Karlsruhe, the 'Ettlinger-Tor-Platz'. (b) Moving contour segments are separated from the background solely by requiring that χ^2 exceeds a threshold of 48; no additional thresholds for the magnitude of the optical flow or for the smallest singular value have been applied.