

# Range Flow in Varying Illumination: Algorithms and Comparisons

Tobias Schuchert, Til Aach, *Senior Member, IEEE*, and Hanno Scharf

**Abstract**—We extend estimation of range flow to handle brightness changes in image data caused by inhomogeneous illumination. Standard range flow computes 3D velocity fields using both range and intensity image sequences. Toward this end, range flow estimation combines a depth change model with a brightness constancy model. However, local brightness is generally not preserved when object surfaces rotate relative to the camera or the light sources, or when surfaces move in inhomogeneous illumination. We describe and investigate different approaches to handle such brightness changes. A straightforward approach is to prefilter the intensity data such that brightness changes are suppressed, for instance, by a highpass or a homomorphic filter. Such prefiltering may, though, reduce the signal-to-noise ratio. An alternative novel approach is to replace the brightness constancy model by 1) a gradient constancy model, or 2) by a combination of gradient and brightness constancy constraints used earlier successfully for optical flow, or 3) by a physics-based brightness change model. In performance tests, the standard version and the novel versions of range flow estimation are investigated using prefiltered or nonprefiltered synthetic data with available ground truth. Furthermore, the influences of additive Gaussian noise and simulated shot noise are investigated. Finally, we compare all range flow estimators on real data.

**Index Terms**—Range flow, illumination changes, brightness constancy constraint, prefiltering, homomorphic filter, gradient constancy, structure tensor, 3D motion estimation.

## 1 INTRODUCTION

IN this paper, the importance of adequate brightness models in 3D velocity field estimation by range flow is investigated. Brightness modeling is an essential component in standard motion estimation approaches. It may provide explicit equations describing brightness constraints, and lead to methods such as prefiltering, discretizations of the constraint equations by convolution filters, and finally, a parameter estimation scheme calculating velocities from input data.

The motivation of this work is the estimation of plant growth. Growth is one of the most important processes in plant life and therefore of high botanical interest. However, this paper does not focus on a best estimation *system* or *method* for this application; rather, it discusses the influence of brightness changes on 3D motion estimation, which frequently occur in such botanical as well as in many other types of data, such as outdoor image sequences.

Estimation of 3D motion fields was (and still is) the subject of considerable research efforts. Range flow estimation [2], [3] uses solely data from range sensors, whereas

Spies et al. [4], [1] incorporate information from both range and image sensors. Reconstruction of scene flow and 3D structure from the optical flow observed in several cameras has been proposed by [5], [6], [7]. These scene flow approaches and most optical flow-based approaches [8], [9] imply brightness constancy and are therefore not suitable if substantial brightness changes are present in a sequence (see Fig. 1, upper half of the shown leaf). For optical flow estimation, less brightness sensitive models, e.g., constancy of the intensity gradient vector [10], [11] have been proposed, as well as physics-based brightness change models [12], [13]. The physics-based models presented in [13] have recently been adapted and extended to moving surfaces under inhomogeneous illumination [14]. Also, suppressing brightness changes by appropriate prefiltering of the intensity data has been shown to be very efficient. One of the simplest approaches is to apply a spatial high-pass filter to minimize the effect of global brightness inhomogeneities. Toth et al. [15] show that using homomorphic prefilters [16] can highly improve motion detection in image sequences with inhomogeneous illumination. More recent approaches for scene flow estimation use statistical similarity measures [17], a gradient constancy constraint [18], or probability distributions for optical flow and disparity [19] to make scene flow estimation more robust against brightness changes.

**Our contribution.** The contribution of this paper is twofold: On the one hand, range flow estimation, as presented by Spies et al. [1], is extended to cope with inhomogeneous illumination. To this end, different techniques known from optical flow estimation are introduced into the range flow constraints, namely, 1) gradient constancy [10], 2) combining of brightness and gradient constancy [11], and 3) physics-based brightness modeling [14]. On the other hand, the performance of the standard and the novel range flow models is investigated. The

• T. Schuchert is with the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Department of Autonomous Systems and Machine Vision (ASM), Fraunhoferstr. 1, 76131 Karlsruhe, Germany. E-mail: tobias.schuchert@iosb.fraunhofer.de.

• T. Aach is with the Institute of Imaging and Computer Vision, RWTH Aachen University, 52056 Aachen, Germany. E-mail: Til.Aach@ifv.rwth-aachen.de.

• H. Scharf is with the Institute for Chemistry and Dynamics of the Geosphere, Forschungszentrum Jülich, 52428 Jülich, Germany. E-mail: h.scharf@fz-juelich.de.

Manuscript received 20 Nov. 2008; revised 13 May 2009; accepted 8 Aug. 2009; published online 19 Aug. 2009.

Recommended for acceptance by D. Hogg.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2008-11-0799.

Digital Object Identifier no. 10.1109/TPAMI.2009.162.

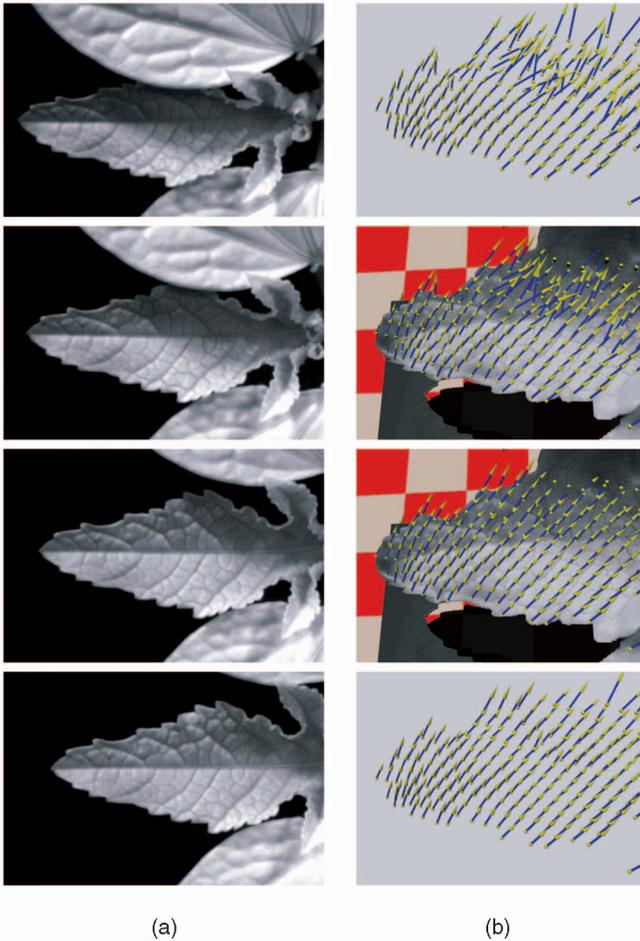


Fig. 1. Castor bean plant leaf. (a) Images of the input sequence, with strong brightness changes occurring especially in the upper half of the central leaf. (b) Estimated motion vector fields of two range flow models, with and without 3D structure shown. Top two images: standard range flow [1]. Bottom two images: proposed TAYLOR model.

models are tested on two kinds of synthetic data sets with known ground truth. These sequences are either used with or without suppression of illumination inhomogeneities by high-pass or homomorphic filtering. A 3D motion estimation result for a “real” image sequence from a botanical experiment on leaf growth is shown in Fig. 1.

Our experiments focus on the effects of brightness models on 3D motion estimation and are therefore especially designed to reveal their influence. For quantitative experiments we selected image areas without occlusions, inner borders, holes, or aperture problems. This affects the choice of the parameter estimation scheme used to finally calculate the 3D motion. For optical flow estimation, estimators such as local least squares [20], local total least squares [21], or variational approaches [22] are often applied. Occlusions and other model violations are typically handled using robust error norms instead of plain least squares [23], [24]. Variational estimators allow us to incorporate additional prior knowledge by, e.g., regularization terms, which tend to close holes and compensate for the aperture problems. Our study, however, seeks to shed light on model violations; thus *no robust estimators* handling occlusion and *no variational estimators* regularizing data are employed. Instead, we apply local total least squares estimation on a neighborhood to

avoid compensation of model violations by, say, regularization. In a final system, robust and/or variational estimators may, of course, be used if needed.

With respect to our earlier work in [25], this paper provides a more detailed derivation of the physics-based brightness model, and deeper analyses of 1) prefilters, 2) prefilters and models for additive Gaussian noise and intensity dependent noise, and 3) the influence of the neighborhood size. Moreover, more experiments on synthetic data and one additional experiment on real data are supplied.

**Paper organization.** In Section 2, we derive the differential range flow model. Then, different prefilters and intensity constraints for range flow are presented in Section 3. In Section 4, we briefly review parameter estimation methods, followed by experiments on synthetic and real data in Section 5.

## 2 RANGE FLOW

Range flow is based on two motion constraints: One pertains to the range data and the other one to the intensity data. Following [1], we briefly review these two constraints.

### 2.1 The Range Constraint

Let a surface be described by its depth  $Z = Z(X, Y, t)$  as a function of space and time, where  $X$ ,  $Y$ , and  $Z$  are spatial coordinates and  $t$  denotes time. Without loss of generality, we define  $X$  and  $Y$  to be aligned with the camera sensor coordinates  $x$  and  $y$ , respectively. The  $Z$ -axis is the optical axis of the camera, which is assumed as projective. The total derivative of  $Z$  with respect to time then yields the so-called range flow motion constraint equation

$$\frac{dZ}{dt} = \partial_X Z \frac{dX}{dt} + \partial_Y Z \frac{dY}{dt} + \partial_t Z, \quad (1)$$

where partial derivatives are denoted by  $\partial_X Z := \frac{\partial Z}{\partial X}$  and so on. The range flow is now defined as  $f = [U, V, W]^T := [\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt}]^T$ . Some range sensors, like those used in [4], and 3D structure from motion algorithms (cf. [14]) produce range data as data sets  $X = X(x, y, t)$ ,  $Y = Y(x, y, t)$ , and  $Z = Z(x, y, t)$  over the sensor coordinates  $x$ ,  $y$ , and time  $t$ . Rewriting the range flow constraint (1) as in [4] allows us to compute partial derivatives directly on the sensor grid rather than in world coordinates, thus avoiding interpolation artifacts and expensive preprocessing steps. Range flow, i.e., the total derivatives of the world coordinates with respect to time, may then be calculated as

$$U = \frac{dX}{dt} = \partial_x X \dot{x} + \partial_y X \dot{y} + \partial_t X, \quad (2)$$

$$V = \frac{dY}{dt} = \partial_x Y \dot{x} + \partial_y Y \dot{y} + \partial_t Y, \quad (3)$$

$$W = \frac{dZ}{dt} = \partial_x Z \dot{x} + \partial_y Z \dot{y} + \partial_t Z, \quad (4)$$

where total derivatives with respect to time are indicated by a dot. Not being interested in the changes on the sensor grid, i.e., the optical flow,  $\dot{x}$  and  $\dot{y}$  in (2)-(4) can be eliminated. This yields

$$\frac{\partial(Z, Y)}{\partial(x, y)}U + \frac{\partial(X, Z)}{\partial(x, y)}V + \frac{\partial(Y, X)}{\partial(x, y)}W + \frac{\partial(X, Y, Z)}{\partial(x, y, t)} = 0, \quad (5)$$

where

$$\frac{\partial(Z, Y)}{\partial(x, y)} = \begin{vmatrix} \partial_x Z & \partial_x Y \\ \partial_y Z & \partial_y Y \end{vmatrix} = \partial_x Z \partial_y Y - \partial_y Z \partial_x X \quad (6)$$

is the Jacobian of  $Z, Y$  with respect to  $x, y$ , and so on. Equation (5) depends only on derivatives in sensor coordinates, and can be calculated easily using derivative kernels. Assuming aligned world and sensor coordinate systems ( $\partial_y X = \partial_x Y = 0$ ), (5) reduces to

$$(\partial_y Y \partial_x Z)U + (\partial_x X \partial_y Z)V - (\partial_x X \partial_y Y)W + (\partial_x X \partial_y Y \partial_t Z - \partial_x X \partial_t Y \partial_y Z - \partial_t X \partial_y Y \partial_x Z) = 0. \quad (7)$$

## 2.2 The Intensity Constraint

The range flow constraint pertains solely to range data, and the full flow can only be estimated where three or more distinct depth planes intersect. Plant surfaces are often nearly planar, smooth surfaces resulting in aperture problems almost everywhere when using the range flow constraint only. As proposed in [4], intensity data should therefore be additionally incorporated. Let the intensity of a point remain constant over the observation time interval. Then, the so-called brightness constancy constraint equation often used for optical flow estimation (see, e.g., [8]) holds. Linearization of this constraint yields

$$\frac{dI}{dt} = \partial_x I \dot{x} + \partial_y I \dot{y} + \partial_t I = 0. \quad (8)$$

Eliminating optical flow  $(\dot{x}, \dot{y})$  using (2) and (3) yields

$$\frac{\partial(I, Y)}{\partial(x, y)}U + \frac{\partial(X, I)}{\partial(x, y)}V + \frac{\partial(X, Y, I)}{\partial(x, y, t)} = 0. \quad (9)$$

The estimated range flow  $f = [U, V, W]^T$  has to fulfill both the range flow constraint (5) and the intensity constraint (9). The intensity constraint is more reliable to provide point-to-point correspondences and therefore often solves the aperture problem. Together with the range constraint, it allows us to solve for  $f$  in places where the range constraint alone is insufficient. The combination of range and intensity constraints and the estimation of  $f$  via total least squares is described in Section 4.

## 3 HANDLING BRIGHTNESS CHANGES

Range flow estimation, as presented in the previous section, yields good results for objects under homogeneous, diffuse illumination. Problems occur for directed, inhomogeneous illumination because the intensity constraint (9) is not sufficiently well satisfied anymore. For an illumination-independent range sensor, such as a laser range sensor, the range constraint is still fulfilled. However, range data estimated from a structure from motion approach may also be corrupted. In the next section, different approaches to handle illumination changes are presented. Three of these lead to constraints novel in range flow estimation.

### 3.1 Prefiltering

Prefiltering is a well-known technique for illumination change suppression, making the image data more or less illumination invariant.

Temporal and/or spatial high-pass filtering approximately eliminates slow brightness changes in the data. However, faster illumination changes in both the spatial and temporal domains still remain present in the data. In our experiments, we evaluate high-pass filtering according to

$$\tilde{I} = (\mathbb{1} - G_{\sigma_{pre}}) * I \quad (10)$$

where  $G_{\sigma_{pre}}$  denotes a spatial Gaussian filter (see, e.g., [26]) with standard deviation  $\sigma_{pre}$ , and where  $\mathbb{1}$  is an identity filter whose impulse response is the unit impulse, and where  $*$  is a convolution.

A more sophisticated approach uses homomorphic filtering [16], [15]. Following [15], we briefly derive a straightforward implementation of homomorphic filtering, which proved to be very successful in suppressing illumination changes. Homomorphic filtering is based on modeling the image intensity  $I(x, y)$  as being determined by illumination  $L(x, y)$  reflected toward the camera by the surfaces of the objects present in the scene. The reflectance of the surfaces is denoted by  $R(x, y)$ . Note that in this model, both the illumination and the object reflectances are already projected onto the image plane, i.e., they depend on the sensor coordinates  $x, y$ . For Lambertian surfaces, the image intensity can thus be modeled by

$$I(x, y) \propto L(x, y) \cdot R(x, y). \quad (11)$$

The structure of the scene is represented by the reflectance  $R(x, y)$ , from which motion therefore needs to be estimated. A logarithmic point transform converts the multiplicative relation<sup>1</sup> between illumination  $L$  and reflectance  $R$  into an additive one according to

$$\log(I(x, y)) \propto \log(L(x, y)) + \log(R(x, y)). \quad (12)$$

In a rough approximation,  $\log(L)$  and  $\log(R)$  may be considered separated in frequency domain as  $L(x, y)$  is assumed to be low-frequent (such that even its harmonics generated by the nonlinear point transform can be neglected), while  $R$ , and hence  $\log(R)$ , is regarded as predominantly high-frequent such that it can be extracted by a high-pass filter. In practice, though, the two components overlap. The design of the high-pass filter thus implies a trade-off between suppressing brightness changes and loss of relevant signal. After high-pass filtering, exponentiation returns an approximation of the sought reflectance component. Note that the nonlinear log-operation turns the camera noise, which, neglecting quantum noise, is often modeled as being signal-independent, into signal-dependent noise, which may affect motion parameter estimation. Fig. 2 illustrates the effects of the proposed prefilterers on a 2D signal  $I(x, t)$  with exponential temporal brightness changes. Results for linear high-pass and homomorphic prefiltering of the signal for noise-free data (top), for added Gaussian noise (middle), and for intensity-dependent noise (bottom) are shown (see Section 5 for details on intensity-dependent

1. Note that this multiplicative relation is also preserved for cameras with a nonlinear conversion from image irradiance to intensity, if the nonlinearity follows a  $\gamma$ -curve.

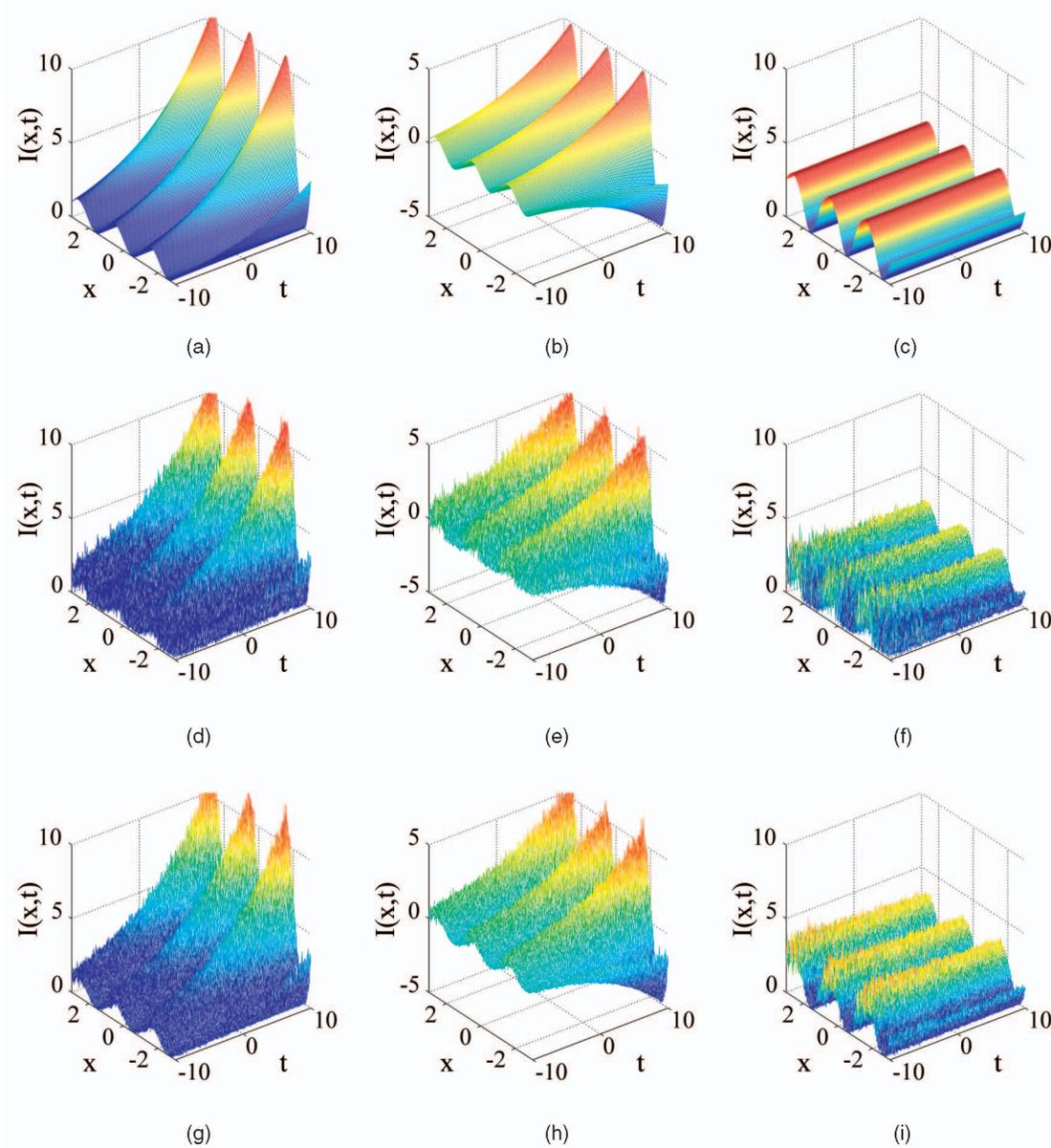


Fig. 2. Sine wave with increasing illumination, parameter  $a_1 = 0.1$ ,  $a_{1,x} = 0$ , (left) without prefilter, (center) after high-pass, and (right) homomorphic prefiltering (low-pass with  $\sigma_{pre} = 35$ ). Noise added: (top) no noise, (middle) additive Gaussian noise ( $\sigma_n = 0.5$ ), and (bottom) shot noise ( $k = 0.5$ , see (33)).

noise). Evidently, homomorphic prefiltering removes all brightness changes for noise-free data, whereas high-pass prefiltering only removes signal offset. Fig. 2f shows how the nonlinear log-operation leads to increased noise in regions with small intensity values, especially for  $t < 0$ . If the noise is intensity-dependent like shot noise, homomorphic prefiltering may reduce this effect (Fig. 2i). A general possibility to reduce the signal dependence of shot or quantum noise is to subject the data first to a square-root transform [27]. We will, though, not consider this any further here.

### 3.2 The Gradient Constancy Constraint

The intensity constraint (9) can be replaced by an illumination-invariant or illumination-insensitive constraint, which can be applied instead of or together with prefiltering the data.

For the estimation of optical flow, a successful such constraint is to assume that the 2D image intensity *gradient* remains constant along the motion trajectory [11]. As gradient computation corresponds to derivative filtering, which is a high-pass operation, this constraint therefore attenuates illumination changes, as discussed in Section 3.1. This leads to two linearized gradient constancy constraints

$$\frac{dI_x}{dt} = \partial_x I_x \dot{x} + \partial_y I_x \dot{y} + \partial_t I_x = 0, \quad (13)$$

$$\frac{dI_y}{dt} = \partial_x I_y \dot{x} + \partial_y I_y \dot{y} + \partial_t I_y = 0, \quad (14)$$

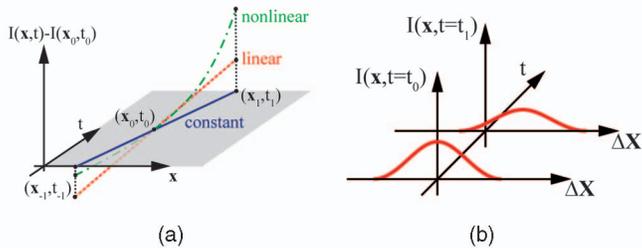


Fig. 3. (a) Change of intensity values caused by brightness changes (cf. [13]) and (b) temporal intensity changes caused by illumination changes within a spatial neighborhood with local 3D coordinates  $\Delta X$ .

where lower indices indicate partial derivatives, e.g.,  $I_x := \partial_x I$ . As already done for the derivation of (9), the optical flow  $\dot{x}, \dot{y}$  is eliminated using (2) and (3), yielding

$$\frac{\partial(I_x, Y)}{\partial(x, y)} U + \frac{\partial(X, I_x)}{\partial(x, y)} V + \frac{\partial(X, Y, I_x)}{\partial(x, y, t)} = 0, \quad (15)$$

$$\frac{\partial(I_y, Y)}{\partial(x, y)} U + \frac{\partial(X, I_y)}{\partial(x, y)} V + \frac{\partial(X, Y, I_y)}{\partial(x, y, t)} = 0. \quad (16)$$

### 3.3 Combined Intensity and Gradient Constancy Constraint

A known drawback of the gradient constancy constraint is that it noticeably reduces structure in the images and leads to aperture problems. Using both the intensity constraint and the gradient constraint simultaneously reduces this effect in optical flow estimation [11]. Doing so leads to three constraint equations, i.e., (9), (15), and (16), which should be satisfied simultaneously by the horizontal and vertical range flow components  $U$  and  $V$ .

### 3.4 A Physics-Based Brightness Change Model

A different approach to handling brightness changes is to model these explicitly and to then estimate both optical flow and brightness change parameters based on this model. Haußecker and Fleet [13] proposed a generalized formulation of optical flow estimation based on models of brightness variation that are caused by time-dependent physical processes. Brightness changes along a temporal trajectory  $\mathbf{x}(t) = (x(t), y(t))^T$ . This is described by a parameterized function  $h_I$

$$I(\mathbf{x}(t), t) = h_I(I_0, t, \mathbf{a}), \quad (17)$$

where  $I_0 = I(\mathbf{x}(0), 0)$  denotes image intensity at time  $t = 0$  and  $\mathbf{a} = [a_1, \dots, a_n]^T$  contains  $n$  brightness change parameters. Taking the total derivative on both sides yields

$$\partial_x I \dot{x} + \partial_y I \dot{y} + \partial_t I = \dot{h}_I(I_0, t, \mathbf{a}). \quad (18)$$

Assuming brightness constancy, i.e.,  $h_I(I_0, t, \mathbf{a}) = c$ , (18) reduces to (8). Fig. 3a shows constant, linear, and nonlinear changes of intensity due to brightness changes. Given a physical model  $h$  for brightness changes, both the optical flow  $(\dot{x}, \dot{y})$  and the parameter vector  $\mathbf{a}$  need to be estimated. Several time-dependent brightness change models are proposed in [13], i.e., changing surface orientation, motion of the illuminant, and physical models of heat transport in

infrared images. We use the brightness change model presented in [14], which describes spatially-varying time-dependent illumination changes caused by directed, inhomogeneous illumination and changing surface orientation. Fig. 3b illustrates spatially-varying intensity changes in a neighborhood  $\Omega$ .  $\Delta X$  and  $\Delta Y$  are local coordinates of  $\Omega$ . This model is also able to describe illumination by spotlights, whereas the models presented in [13] assume that the intensity change of pixels is constant over a local neighborhood. Following [14], the brightness change function is set to

$$h_I(I_0, t, \mathbf{a}) = I_0 \exp(h(\Delta X, \Delta Y, t, \mathbf{a})). \quad (19)$$

The incident irradiance caused by the moving illuminant is assumed to be spatially inhomogeneous. It therefore changes not only according to a time-dependent parameter, but varies also smoothly in space. Approximating these brightness changes by a second order Taylor series, and with the notation

$$\mathbf{a} = [a_1, a_{1,x}, a_{1,y}, a_2, a_{2,x}, a_{2,y}]^T, \quad (20)$$

this yields

$$h(\Delta X, \Delta Y, t, \mathbf{a}) := \sum_{i=1}^2 (a_i + a_{i,x} \Delta X + a_{i,y} \Delta Y) t^i \quad (21)$$

and its temporal derivative is

$$\dot{h}(\Delta X, \Delta Y, t, \mathbf{a}) = \sum_{i=1}^2 i (a_i + a_{i,x} \Delta X + a_{i,y} \Delta Y) t^{i-1}. \quad (22)$$

As shown in [14], we may drop  $a_{2,x}$  and  $a_{2,y}$  as these have only a negligible effect on the estimation. Similarly to Section 2.2, we obtain the total differential

$$\frac{dI}{dt} = \partial_x I \dot{x} + \partial_y I \dot{y} + \partial_t I = I \dot{h}(\Delta X, \Delta Y, t, \mathbf{a}) \quad (23)$$

where we, as before, eliminate the optical flow  $(\dot{x}, \dot{y})$  using (2) and (3), leading to

$$\begin{aligned} \frac{\partial(I, Y)}{\partial(x, y)} U + \frac{\partial(X, I)}{\partial(x, y)} V + \frac{\partial(X, Y, I)}{\partial(x, y, t)} \\ - I a_1 - I a_{1,x} \Delta X - I a_{1,y} \Delta Y - 2I a_2 t = 0. \end{aligned} \quad (24)$$

## 4 ESTIMATION OF MOTION AND BRIGHTNESS PARAMETERS

So far, we have described modeling and prefiltering of the data. Estimating the 3D motion from the input data then corresponds to fitting the selected model to the original or, if applicable, prefiltered data. The parameters to be estimated comprise the motion parameters  $U, V, W$  for all models and, for the physics-based brightness change model from Section 3.4, the brightness change parameters  $a_1, a_{1,x}, a_{1,y}, a_2$  in the Taylor series (21).

To estimate parameters within a total least-squares framework, we closely follow [1]. The range constraint (7) yields for every pixel an equation of the form  $\mathbf{d}_{rc}^T \mathbf{p} = 0$  with

$$\mathbf{d}_{rc} = \left[ \frac{\partial(Z, Y)}{\partial(x, y)}, \frac{\partial(X, Z)}{\partial(x, y)}, \frac{\partial(Y, X)}{\partial(x, y)}, \frac{\partial(X, Y, Z)}{\partial(x, y, t)} \right]^T \quad (25)$$

and parameter vector

$$\mathbf{p} = [U, V, W, 1]^T. \quad (26)$$

To solve this equation which contains three unknowns, we assume that, within a local neighborhood  $\Omega$ , one parameter vector  $\mathbf{p}$  solves all equations up to an error  $\mathbf{e}$ . Minimizing the weighted  $L_2$ -norm of the error  $\mathbf{e}$  yields

$$\|\mathbf{e}\|_2 = \mathbf{p}^T \mathbf{J}_{rc} \mathbf{p} \stackrel{!}{=} \min, \quad (27)$$

with structure tensor  $\mathbf{J}_{rc} = \mathbf{W} * (\mathbf{d}_{rc} \mathbf{d}_{rc}^T)$  and an averaging filter  $\mathbf{W}$  which determines the neighborhood  $\Omega$ . As described in Sections 2.2, 3.2, 3.3, and 3.4, we have more than one constraint for the  $U$  and  $V$  components of the range flow. In the same way as expressing the range constraint by  $\mathbf{d}_{rc}^T \mathbf{p} = 0$ , the intensity constraint (9) and the gradient constancy constraints (15) and (16) may be expressed by  $\mathbf{d}_Q^T \mathbf{p} = 0$ , where  $\mathbf{d}_Q$  is computed from the observed data according to

$$\mathbf{d}_Q = \left[ \frac{\partial(Q, Y)}{\partial(x, y)}, \frac{\partial(X, Q)}{\partial(x, y)}, 0, \frac{\partial(X, Y, Q)}{\partial(x, y, t)} \right]^T, \quad (28)$$

with  $Q \in \{I, I_x, I_y\}$  selected as appropriate. We adapt all constraints to identical dimensions by inserting zeros at those positions into the data vector  $\mathbf{d}$ , which correspond to positions of parameters in the parameter vector  $\mathbf{p}$  which are not part of the respective constraint. In particular, the physics-based brightness change model in Section 3.4 contains motion *and* brightness change parameters, thus leading to the enlarged parameter vector

$$\mathbf{p} = [U, V, W, 1, \mathbf{a}^T]^T. \quad (29)$$

The data vectors for both the range constraint and the brightness change constraint hence have to be enlarged correspondingly by appropriate insertion of zeros.

As shown in [1], combining the different constraints yields a new structure tensor which is simply the weighted sum of the tensor  $\mathbf{J}_{rc}$  from the range constraint (27), and the tensors  $\mathbf{J}_i$ ,  $i = 1, \dots, 4$ , from the intensity-dependent constraints (9), (15), (16), and (24). With the weights  $\beta_i$ , the overall tensor thus is

$$\mathbf{J} = \mathbf{J}_{rc} + \sum_{i=1}^4 \beta_i \mathbf{J}_i. \quad (30)$$

The weights  $\beta_i$  may even be used to switch between the models and to account for different signal-to-noise-ratios in the structure tensors. Furthermore, the data channels should be scaled to the same mean and variance before they are combined.

The optimal estimate of the sought parameter vector is then given by the eigenvector  $\mathbf{b}$  corresponding to the lowest eigenvalue of  $\mathbf{J}$ . The lowest eigenvalue can be regarded as an error measure, which vanishes for an error-free estimate. As the eigenvector is only defined up to a scaling factor, the range flow is finally computed by the normalization (see (26))

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \frac{1}{b_4} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}. \quad (31)$$

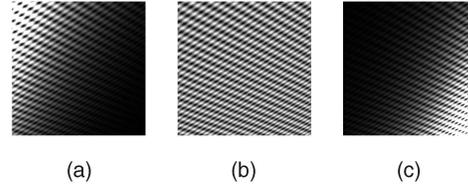


Fig. 4. Scaled (a) first, (b) central, (c) and last frame of sinusoidal sequence with illumination parameters  $a_1 = 0$  and  $a_{1,x} = 0.06$ .

## 5 EXPERIMENTS

The models compared are combinations of the range constraint with

1. the intensity constancy constraint (9), i.e.,  $\beta_2 = \beta_3 = \beta_4 = 0$  (INT),
2. the gradient constancy constraint (15) and (16), i.e.,  $\beta_1 = \beta_4 = 0$  (GRAD),
3. the combined intensity and gradient constancy constraint (9), (15), and (16), i.e.,  $\beta_4 = 0$  (INTGRAD), and
4. the intensity constraint with modeling of brightness changes by Taylor series (24), i.e.,  $\beta_1 = \beta_2 = \beta_3 = 0$  (TAYLOR).

Model INT without prefiltering is the original range flow as introduced by Spies et al. [1]. We investigate the influence of temporal and spatial-varying temporal brightness changes for noise-free input sequences as well as for data corrupted by additive Gaussian noise or by intensity-dependent noise. Furthermore, accuracy of the 3D motion estimates is evaluated on a rendered cube illuminated by a directed light source, and finally, on real data. For the synthetic sequences, the ground truth of the range flow is available, while for all other experiments, we generate reference range data by multicamera stereo reconstruction, as described in [28]. Using multicamera data allows for comparison with [18] and has no further effect on range flow.

To compute the intensity derivatives, we used the optimized  $5 \times 5 \times 5$  filter sets described in [29] for all experiments. Motion and brightness change parameters are determined as the minimizer of the model error given in (27), as described in Section 4. Spies et al. [1] use  $\beta_1 = 1$  for their experiments (cmp. (30)). We scale all structure tensors from the different models and prefilters to have the same variance as model INT without prefiltering. In our earlier work [25], structure tensors were not scaled, therefore leading to slightly different results.

### 5.1 Sinusoidal Patterns

For a systematic error analysis of the different models, we use patches with synthetically generated sinusoidal patterns under varying illumination. The varying illumination is generated according to (17), (19), and (21). Three frames of a typical test sequence are shown in Fig. 4. Its parameters were set as follows: Rendered surfaces translate with  $U = 0.0073$  mm/frame,  $V = 0$  mm/frame, and  $W = 0.5$  mm/frame and rotate around the  $Y$ -axis with an angular velocity of  $\omega = 0.002$  radians/frame. For  $t = 0$ , the surface normal of the patch is  $\mathbf{n} = (1, 2, -1)^T$ , and the distance of the patch center to the camera is  $Z_0 = 100$  mm. The synthetic sensor contains  $301 \times 301$  pixels of size  $(0.0044 \text{ mm})^2$ . The focal

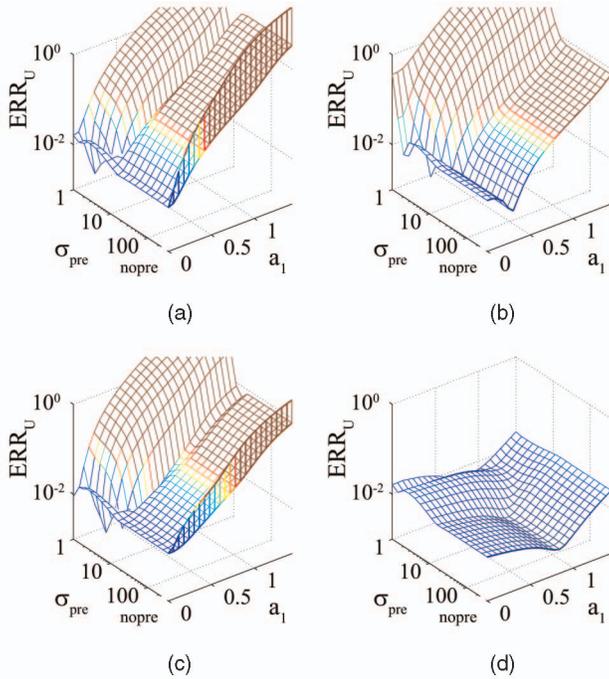


Fig. 5. Error of motion estimates  $U$  versus increasing *temporal* brightness changes and *high-pass* prefiltering or without prefiltering (“nopro”). *Temporal* brightness changes are modeled by  $a_1 \neq 0$ . All other parameters of  $\mathbf{a}$  in (20) are zero. Standard deviation  $\sigma_{pre}$  according to (10). Models are (a) INT, (b) GRAD, (c) INTGRAD, and (d) TAYLOR.

length of the synthetic projective camera is  $f = 12$  mm. For each experiment, we evaluate the mean absolute value of the relative error of  $U$

$$ERR_U = \frac{1}{N} \sum_i \frac{|U_{estimated}(i) - U_{reference}(i)|}{|U_{reference}(i)|} \quad (32)$$

over all pixels  $i$  at a minimum distance of 60 pixels from the nearest image border. The structure tensor weighting matrix  $\mathbf{W}$  is realized by a large, 65-tab Gaussian with standard deviation  $\sigma_W = 19$  in order to reduce systematic errors stemming from the phase of the sinusoidal pattern. Prefilters are designed as described in Section 3.1, i.e.,  $\sigma_{pre}$  denotes the standard deviation of the low-pass  $G_{\sigma_{pre}}$  in (10). As in [14], we compare estimation errors of  $U$  for increasing illumination parameters  $a_1|_{a_{1,x}=0}$  and  $a_{1,x}|_{a_1=0}$  to simulate brightness changes. We provide here the errors for  $U$  only, as errors

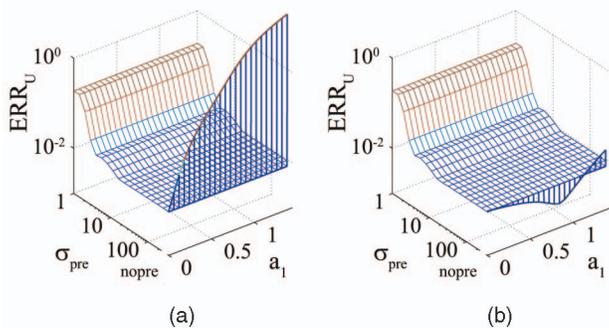


Fig. 6. Same as Fig. 5 but with *homomorphic* prefiltering instead of *high-pass*. Models are (a) INT and (b) TAYLOR.

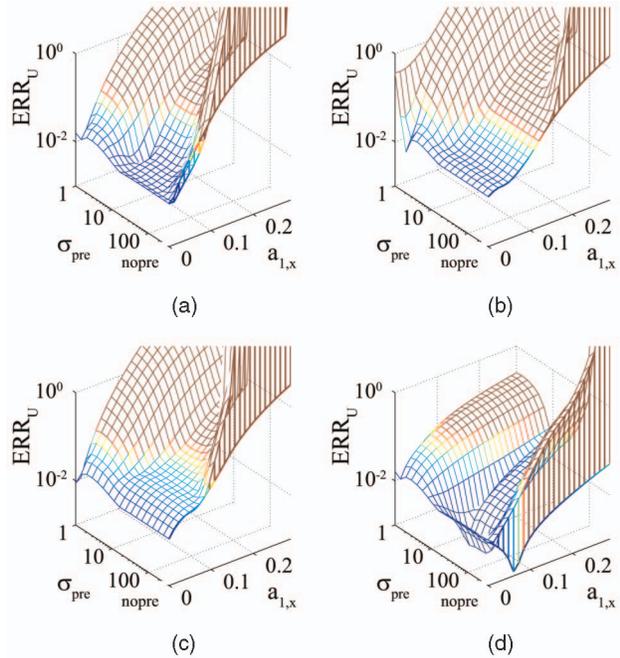


Fig. 7. Error of motion estimates  $U$  versus increasing *spatially-varying temporal* brightness changes and *high-pass* prefiltering or without prefiltering (“nopro”). *Spatially-varying temporal* brightness changes are modeled by  $a_{1,x} \neq 0$ . All other parameters of  $\mathbf{a}$  in (20) are zero. Standard deviation  $\sigma_{pre}$  according to (10). Models are (a) INT, (b) GRAD, (c) INTGRAD, and (d) TAYLOR.

of  $V$  and  $W$  exhibited very similar behavior. Furthermore, we investigate the influence of adding Gaussian noise with  $\sigma_n = 0.025$  or shot noise to the intensity data. The Poisson-distributed shot noise is approximated by adding Gaussian noise with intensity-dependent standard deviation

$$\sigma_{sn} = k \sqrt{\frac{I(x,t)}{\alpha}} \quad (33)$$

with  $k = 0.025$  and  $\alpha = 1$ . The sinusoidal pattern for  $t = 0$  has intensity values in the range of  $0 \leq I \leq 2$ . To facilitate the evaluation of the effects of prefiltering, we show the errors for all tested models *with and without* prefiltering in Figs. 5, 6, 7, 8, 9, and 10.

In Fig. 5, the errors of  $U$  for the described models with and without high-pass prefiltering are shown over the parameter  $a_1$ . High-pass prefiltering reduces estimation errors for models INT and INTGRAD. For models GRAD

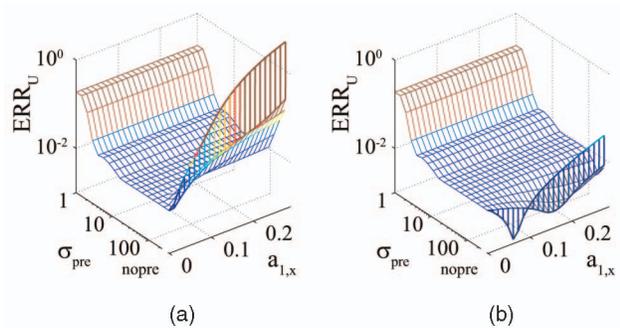


Fig. 8. Same as Fig. 7 but with *homomorphic* prefiltering instead of *high-pass*. Models are (a) INT and (b) TAYLOR.

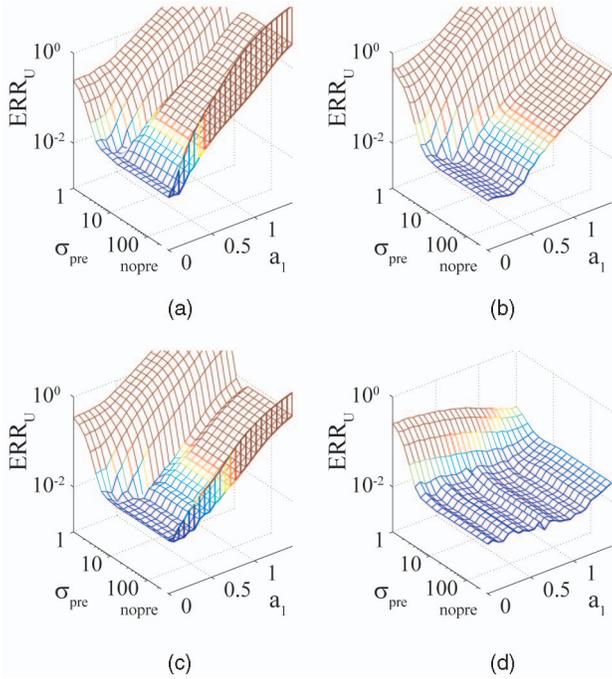


Fig. 9. Same as Fig. 5 but for data with additive Gaussian noise of standard deviation  $\sigma_n = 0.025$ . Error of motion estimates  $U$  versus increasing *temporal* brightness changes and *high-pass* prefiltering or without prefiltering (“nopre”). Models are (a) INT, (b) GRAD, (c) INTGRAD, and (d) TAYLOR.

and TAYLOR, using no prefilter yields results as good as using a high-pass with standard deviation  $\sigma_{pre} > 15$  (see (10)). Strong filtering, i.e., small  $\sigma_{pre}$ , degrades estimation results for all models. Overall, the model TAYLOR yields the best results for most cases; only for  $a_1 \approx 0.3$ , the model GRAD performs slightly better. Grooves in the error surfaces where errors become low, e.g., for model INT at  $a_1 \approx 0.1$  and  $\sigma_{pre} < 10$ , may be the result of improved gradients due to brightness changes and are highly pattern-dependent. All models perform comparable when applying homomorphic prefiltering. Therefore, in Fig. 6, we show errors for the models INT and TAYLOR only. Homomorphic prefiltering with  $\sigma_{pre} > 10$  reduces the brightness changes efficiently for all values of  $a_1$ , while prefiltering with  $\sigma_{pre} < 10$  degrades the signal. For small brightness changes, model TAYLOR performs even better without prefiltering.

Figs. 7 and 8 repeat the experiments shown in Figs. 5 and 6, but with brightness changes now being governed by the parameter  $a_{1,x}$  instead of  $a_1$ . This means that brightness changes vary now not only temporally, but also spatially.

Evidently, the errors shown in Fig. 7 depend more strongly on  $\sigma_{pre}$  than those in Fig. 5.

Prefiltering with standard deviation of  $12 \leq \sigma_{pre} \leq 50$  yields the best results. As before, models GRAD and TAYLOR perform well without prefilter. Fig. 8 shows similar behavior for homomorphic prefiltering. Prefiltering with  $\sigma_{pre} > 50$  degrades the signal for all models but model TAYLOR. For model TAYLOR, using no prefilter yields best results if  $a_{1,x} \lesssim 0.07$ .

Figs. 9 and 10 illustrate the effects of the proposed prefilterers on noisy data. In Section 3.1, high-pass filtering was shown to have no effect on noise distributions, in contrast to homomorphic prefiltering. We present errors of

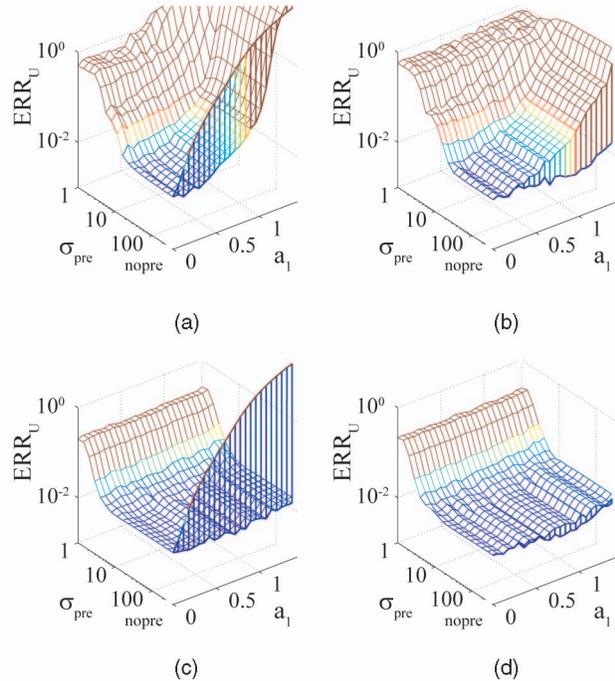


Fig. 10. Same as Fig. 6 but for input data with added Gaussian noise with  $\sigma_n = 0.025$  (top) or shot noise with  $k = 0.025$  (bottom). Error of motion estimates  $U$  versus increasing *temporal* brightness changes and *homomorphic* prefiltering or without prefiltering (“nopre”). Models are (left) INT and (right) TAYLOR.

$U$  for high-pass prefiltering for added Gaussian noise with standard deviation  $\sigma_n = 0.025$  only (Fig. 9) because results for shot noise are almost the same. In comparison to the noise-free case, the errors increased and the grooves in the error surfaces vanished. Overall, the model TAYLOR yields best results for  $a_1 > 0$  and similar results as the other models for  $a_1 \approx 0$ .

Fig. 10 shows how homomorphic prefiltering affects estimation results for added Gaussian noise and simulated shot noise. As stated above, the nonlinear log-operation makes originally signal-independent noise signal-dependent. Therefore, homomorphic prefiltering of noisy intensity data with brightness changes  $a_1 > 0.5$  causes heavily degraded signals. If the noise itself is signal-dependent, such as shot noise, homomorphic prefiltering performs much better. In both noise experiments, model TAYLOR without prefilter yields the best and most reliable results.

## 5.2 Synthetic Cube

The synthetic cube sequence allows us to test the models on more realistic data, but still with an available ground truth. The cube moves with  $U = -0.2$  mm/frame,  $V = 0$  mm/frame, and  $W = -2$  mm/frame. In addition to ambient light, the cube is illuminated by a fixed spotlight from the right. The input data consists of sequences with nine frames each, acquired by five synthetic cameras positioned on the  $x$ -axis. This setup allows us to use the optimized filter sets proposed in [29]. We compute range data with the stereo estimation algorithm presented in [14]. Fig. 11 shows two frames of the cube sequence, a frame after high-pass prefiltering and one after homomorphic prefiltering. Also shown are the regions investigated on the left and right side of the cube, as well as a rendering of the ground truth. Both prefilterers have a standard deviation of  $\sigma_{pre} = 2$ , and the weighting matrix  $\mathbf{W}$  is

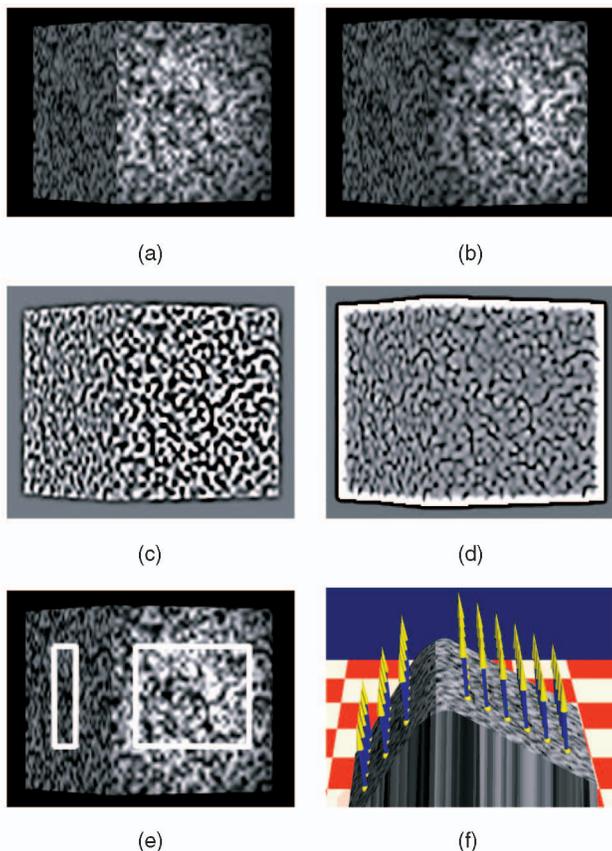


Fig. 11. Cube experiment data. (a) First and (b) last frame of cube sequence, (c) central frame of cube sequence after high-pass and (d) homomorphic prefiltering, (e) regions of cube used for error analysis, and (f) rendering of ground truth.

implemented using a 31-tab Gaussian with standard deviation  $\sigma_w = 11$ . In Fig. 12, motion estimates of the original range flow [1] (i.e., model INT) are compared to the following three models: GRAD with high-pass prefilter, INTGRAD with homomorphic prefilter, and TAYLOR without prefilter. For all models, the errors are too small to be visible without amplification. Thus, *errors are amplified* for  $U$  and  $V$  by 100 and for  $W$  by 50.

Model INT without prefilter (Fig. 12a) yields highly corrupted estimation results on the right side of the cube, where illumination changes because of the fixed spotlight. The other side does not suffer from illumination changes. Range flow estimates there are thus much more accurate. As expected, estimates are substantially improved by the other models where brightness changes are present.

All combinations of models and prefilters yield more or less the same results as original range flow on the left side of the cube.

On its right side, where the brightness changes dominate, applying a high-pass prefilter, for instance, together with model GRAD (Fig. 12b), improves the motion estimates where spatial brightness changes are small, i.e., in the middle of the light spot. Motion estimates are visibly worse at the borders of the light spot. Model TAYLOR (Fig. 12d) and model INTGRAD yield more uniform and more accurate motion vectors for the right side of the cube.

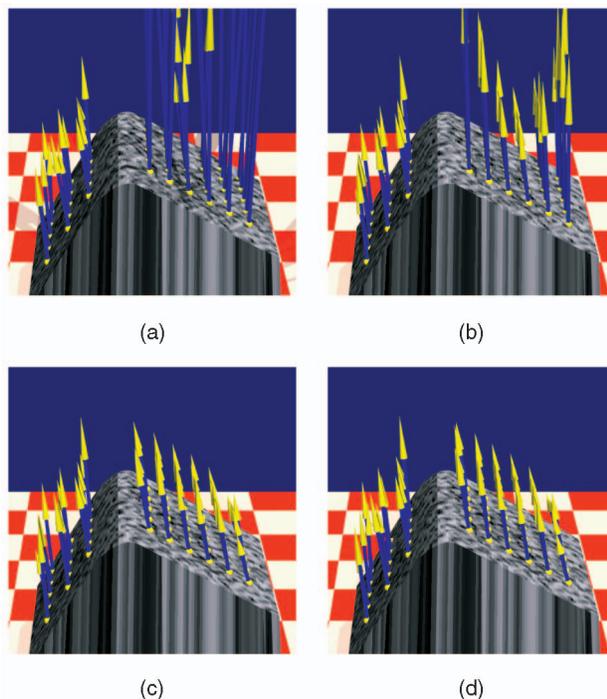


Fig. 12. Scaled motion estimates with amplified errors for different models: (a) INT without prefiltering, (b) GRAD with high-pass prefiltering, (c) INTGRAD with homomorphic prefiltering, (d) TAYLOR without prefiltering. The ground truth is shown in Fig. 11f.

Table 1 shows numerical errors of the different models for the regions on the left and right side of the cube (see Fig. 11). We show the average angular error [8]

$$AAE = \arccos\left(\frac{f_c f_e}{|f_c| |f_e|}\right) [^\circ] \quad (34)$$

and its standard deviation where  $f_c$  is the true and  $f_e$  the estimated flow.

Furthermore, since our target application is plant growth estimation, we compare average relative growth rates. In this experiment, the cube does not grow, thus estimated relative growth rates should be zero. According to [30], the relative area change  $dA$  of a local surface  $s$  may be calculated by

$$dA = \frac{[|s(x+1, y) + f(x+1, y) - (s(x, y) + f(x, y))| \times |s(x, y+1) + f(x, y+1) - (s(x, y) + f(x, y))|]}{[|s(x+1, y) - s(x, y)| \times |s(x, y+1) - s(x, y)|]} \quad (35)$$

when  $s(x, y) = [X(x, y), Y(x, y), Z(x, y)]^T$  is parameterized in sensor coordinates  $x$  and  $y$ , and when the 3D displacement vector field  $f(x, y) = [U(x, y), V(x, y), W(x, y)]^T$  is given. The relative growth rate is determined by  $RGR = (dA - 1) \cdot 100\%$ .

Table 1 confirms the observations from Fig. 12 that all models perform comparably on the left side of the cube. On the right side of the cube, where illumination changes are present, homomorphic prefiltering results in excellent estimates for all models. Model TAYLOR is most robust with respect to prefiltering, whereas results for the other models are much more prefilter-dependent.

TABLE 1  
Average Angular Error (AAE) in Degrees, Average Relative Growth Rate (RGR) in Percent per Frame and Their Standard Deviations of Regions on the Left and Right Sides of the Cube (see Fig. 11e)

model / method	prefilterer	left region		right region	
		AAE	RGR	AAE	RGR
INT	NO	0.148 ± 0.074	0.093 ± 0.092	4.311 ± 4.024	4.372 ± 9.211
	HP	0.136 ± 0.071	0.108 ± 0.096	0.255 ± 0.278	0.137 ± 0.195
	HOM	0.132 ± 0.069	0.111 ± 0.099	0.061 ± 0.015	-0.007 ± 0.054
GRAD	NO	0.146 ± 0.075	0.096 ± 0.095	0.681 ± 0.719	0.352 ± 0.627
	HP	0.139 ± 0.070	0.111 ± 0.098	0.425 ± 0.455	0.236 ± 0.321
	HOM	0.131 ± 0.067	0.115 ± 0.100	0.062 ± 0.015	-0.007 ± 0.055
INTGRAD	NO	0.150 ± 0.075	0.095 ± 0.094	4.764 ± 4.309	5.055 ± 10.029
	HP	0.143 ± 0.071	0.110 ± 0.097	0.671 ± 0.726	0.395 ± 0.552
	HOM	0.132 ± 0.068	0.113 ± 0.099	0.062 ± 0.015	-0.006 ± 0.053
TAYLOR	NO	0.148 ± 0.074	0.090 ± 0.094	0.056 ± 0.014	-0.009 ± 0.052
	HP	0.136 ± 0.071	0.108 ± 0.096	0.046 ± 0.015	-0.012 ± 0.051
	HOM	0.132 ± 0.069	0.111 ± 0.099	0.061 ± 0.015	-0.007 ± 0.054
[18]	NO	4.941 ± 1.812	-0.739 ± 11.59	2.216 ± 1.223	-3.36 ± 8.694
	HP	8.977 ± 3.714	-3.20 ± 15.51	4.041 ± 4.078	-4.19 ± 1.557
	HOM	8.907 ± 2.995	-1.83 ± 13.02	2.093 ± 0.784	0.347 ± 7.358

Errors or standard deviations above 1 degree (AAE) or 1 percent/frame (RGR) are indicated in red, below 0.1 degree (AAE) or 0.01 percent/frame (RGR) in green. Please note that the method from [18] uses only four input frames in total.

Additionally, we show results obtained by the recent scene flow approach of Huguet and Devernay [18]. This is essentially a warping technique for stereo camera sequences analogous to the optical flow approach of Papenberg et al. [11]. We apply the algorithm as provided by Huguet and Devernay using parameter settings of the *rotating sphere* experiment. Only the weighting parameter  $\gamma$  of the gradient constraint was increased to  $\gamma = 30$  in order to reduce the effect of the severe brightness variations in the data. The algorithm handles four input images, i.e., two for each camera in a stereo setup, instead of nine frames from five cameras. We chose the first and the last frame of the two outer cameras from the cube data set.

Compared to the original range flow of [1], and on the right side of the cube, errors are reduced by approximately a factor of 2. However, on the stronger tilted left side of the cube, results are drastically worse than for all other models. In contrast to the other models, prefiltering does not or not significantly improve results. Probably incorporating more input data, i.e., more than four images, may significantly

improve results. This would confirm that using an elaborate estimator alone does not necessarily help; instead, the whole estimation framework, including appropriate constraints, discretizations, and the estimator, needs to be optimized.

In Fig. 13, we compare the performance of the different models for different sizes of the neighborhood  $\Omega$  used in the estimation process. As stated in Section 4, the neighborhood  $\Omega$  is defined by a normalized Gaussian filter  $\mathbf{W}$  with standard deviation  $\sigma_W$ . On the left side of the cube, all models yield similar results for most sizes of  $\mathbf{W}$ . Only for higher values of  $\sigma_W$  does model INTGRAD with homomorphic prefiltering yield slightly higher errors. For neighborhoods with  $\sigma_W > 25$ , edges in the data degrade estimation results. On the right side of the cube, model INT without prefiltering performs worst, as expected. Also, model GRAD with high-pass prefiltering yields unsatisfying results. Model INTGRAD with homomorphic prefiltering and model TAYLOR without prefilter perform similarly and significantly better than the other models. The optimal size of filter  $\mathbf{W}$  is  $\sigma_W \lesssim 15$ .

We conclude that, if the brightness constancy assumption is violated, homomorphic prefiltering rather than linear high-pass prefiltering should be used. Furthermore, the size of the prefilter has a great impact on estimation results and has to be chosen carefully. Nevertheless, even for small neighborhoods, modeling brightness changes yields slightly more accurate results than using a model which merely attenuates the effects of brightness changes.

### 5.3 Real Data

The previous experiments showed that model TAYLOR yields the most reliable estimates. Haußecker and Fleet [13] tested their brightness change model on a sequence of an arm under changing illumination. We replicate this experiment for range flow and estimate 3D structure and 3D motion of the arm. We use a stereo camera setup and calculate 3D structure using the algorithms described in [14]. The prefilters have the same size as in the cube experiment, and

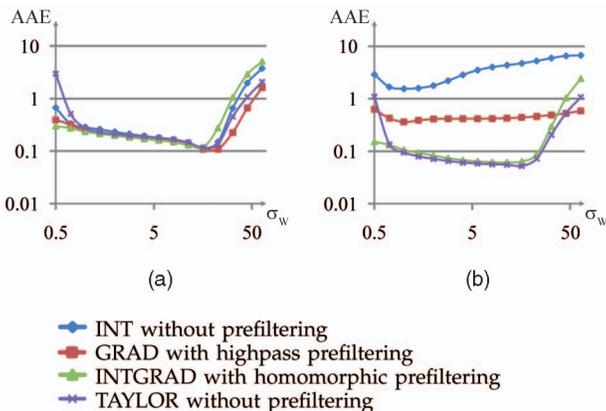


Fig. 13. Average angular error of the motion estimates in the (a) left and (b) right region of the cube for increasing  $\sigma_W$  of weighting matrix  $\mathbf{W}$  for different models.

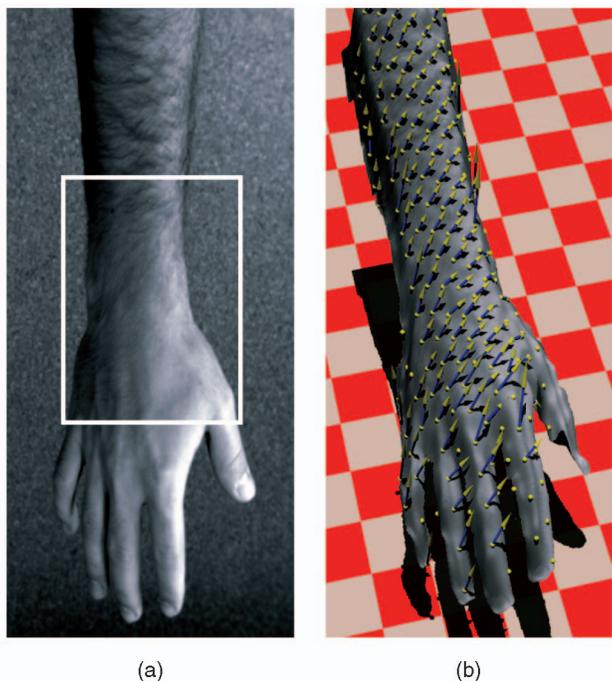


Fig. 14. (a) One image of the rotating arm sequence indicating the area shown in Fig. 15, (b) 3D structure and 3D motion estimated with model TAYLOR without prefilter.

the neighborhood matrix  $\mathbf{W}$  is implemented using a Gaussian filter with  $\sigma_{\mathbf{W}} = 5$ . We do not show motion estimates if  $|U_X|, |U_Y| > 10$  or  $|U_Z| > 20$ .

Fig. 14 shows one frame of the sequence (rotation under directional illumination from the right) and the estimated 3D structure and 3D motion of the arm using model TAYLOR without prefilter. Hausßcker and Fleet [13] calculated error ellipses to indicate confidence in optical flow estimates. They use an approximation of the error covariance matrix, as derived in [31]. We adapt this concept to show uncertainties of range flow solutions. Error covariance matrices  $\Sigma$  are approximated by  $\Sigma = \mathbf{H}^{-1}$ , where the Hessian  $\mathbf{H}$  is given by

$$\mathbf{H} = \frac{\gamma}{\sigma_n^2 \|\mathbf{p}\|^2} \left( \mathbf{M} - \frac{1}{\|\mathbf{p}\|^2} (\mathbf{p}^T \mathbf{J} \mathbf{p}) \mathbf{I} \right) \quad (36)$$

and  $\mathbf{M}$  contains all but the last column and last row of structure tensor  $\mathbf{J}$  in (30).  $\mathbf{I}$  is an identity matrix of the same size as  $\mathbf{M}$ , and  $\mathbf{p}$  is the estimate (27) of the parameter vector. The parameter  $\gamma$  is given by  $\gamma = \sigma^2 / (\sigma_n^2 + \sigma^2)$ , where  $\sigma^2$  denotes the signal power in the measurements. Note that this equation differs from the one presented in [31] as the additional terms in the latter collapse to zero at the solution  $\mathbf{p}$ . For the arm sequence, we show error ellipsoids, i.e., ellipsoids satisfying  $\mathbf{e}^T \Sigma_f^{-1} \mathbf{e} = 6.25$ , where  $\Sigma_f$  is the 3D error covariance submatrix for range flow  $f$ , and  $\mathbf{e}$  is a point on the ellipsoid. These ellipsoids cover approximately 90 percent of the expected errors.

Fig. 15 shows a closeup view of two frames of the input sequence in Fig. 14 as well as the estimated motion and uncertainty ellipsoids for different models. Estimates of model INT without prefilter are strongly degraded in regions where brightness changes occur. Motion estimates

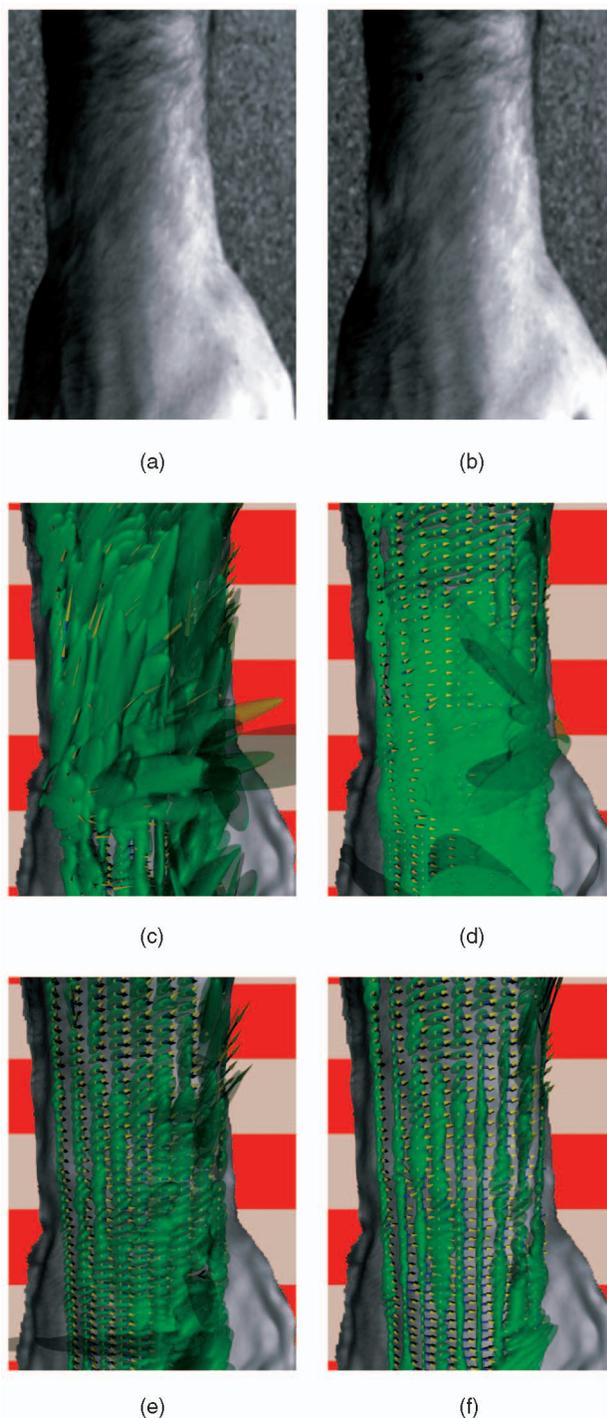


Fig. 15. (a) and (b) The area of interest of frames 1 and 9 of the rotating arm sequence. Motion estimates and uncertainty ellipsoids are shown for (c) model INT without prefilter, (d) model GRAD with high-pass prefilter, (e) model INTGRAD with homomorphic prefilter, and (f) model TAYLOR without prefilter.

are significantly better for model GRAD with high-pass prefilter, but uncertainty ellipsoids are still quite large. Model INTGRAD increases reliability of the estimates, but at the border of the illuminated region (right side of the arm) estimates are still noticeably degraded. Model TAYLOR without prefilter performs best. Uncertainty ellipsoids are significantly smaller compared to all other models, and the motion estimates capture the motion of the arm well.

In Fig. 1, in Section 1, we show estimated 3D velocity fields for a freely moving castor bean leaf. The scene is illuminated by directed infrared light emitting diodes from the top right causing shadows on the leaf of interest. Depth reconstruction was obtained according to [28] using five images from different camera positions at each time step. For motion estimation, a sequence of nine frames from the center camera with a sampling rate of one frame per 2 minutes was taken, i.e., acquisition time for all images was 16 minutes.

The leaf rotates around the node where it is attached to the stem. This results in a visible motion toward the camera and to the right, while the shadow area caused by the top leaf decreases. Model TAYLOR visibly improves estimation results in regions with illumination changes compared to the standard range flow model [1]. As expected, for an almost rigid motion, which can be assumed because of the relatively high temporal resolution, we obtain a smoothly varying vector field.

## 6 CONCLUSIONS

In this paper, we extended range flow estimation described in [1] by different approaches to handle inhomogeneous illumination. We presented a detailed error analysis for four different brightness constraints in combination with high-pass or homomorphic prefiltering on synthetic image sequences. Experiments on two real sequences showed the applicability of the models in more realistic scenarios.

Prefiltering improved estimation results on data when illumination changes were present; however, the standard deviation  $\sigma_{pre}$  of the low pass used in the filters should be large enough. If not, too much signal is lost, resulting in less accurate motion estimates. High-pass filtering performs well if Gaussian noise or shot noise is present. Homomorphic filtering works excellently for shot noise, but errors increase when Gaussian noise and strong brightness changes are present. We conclude that, except for this case, well-tuned homomorphic prefiltering allows for accurate range flow estimation in conjunction with all models. However, a wrongly tuned filter severely corrupts motion estimates. Furthermore, prefiltering increased uncertainty of motion estimates as the relevant signal is then also affected by the filtering.

Modeling brightness changes by the models GRAD or TAYLOR instead of high-pass prefiltering results in motion estimates of the same or higher accuracy in almost all scenarios investigated. The model INTGRAD mostly showed results lying between the results of standard range flow (model INT) and the model GRAD. Generally, the best results were achieved by model TAYLOR and, especially when noise was present, without prefiltering. Only for brightness changes with strong spatial variations, i.e.,  $a_{1,x} \gg 0$ , and noise-free data is homomorphic prefiltering recommended. We conclude that suitably modeling brightness changes almost always outperforms other approaches and increases overall accuracy without tuning of prefilters.

Typical growth rates of fast growing leaves are approximately 3 percent per hour or 0.1 percent/frame when a typical frame rate of 0.5 frames per minute is assumed. The cube experiment showed that growth estimates are accurate when homomorphic prefiltering or model TAYLOR are applied. On the right side of the cube, best growth estimates

are well below 0.1 percent/frame and accurate enough for plant growth studies where a relative measurement error of approximately 10 percent is acceptable. However, variances are high, such that spatial resolution is not yet in the desired range. In addition, results for the accuracy on the left, stronger tilted side of the cube are not yet sufficient. In future work, an estimation scheme needs to be designed which is suitable to estimate growth also for more inclined surfaces. This may allow to build a *method* accurate enough for leaf growth estimation by divergence of range flow.

## ACKNOWLEDGMENTS

The authors thank Georg Dreissen and Daniel Gorman for their help with the acquisition of the plant leaf and the arm sequence. Tobias Schuchert was with the Institute for Chemistry and Dynamics of the Geosphere, Forschungszentrum Jülich, Germany.

## REFERENCES

- [1] H. Spies, B. Jähne, and J. Barron, "Range Flow Estimation," *Computer Vision and Image Understanding*, vol. 85, no. 3, pp. 209-231, Mar. 2002.
- [2] M. Yamamoto, P. Boulanger, J.A. Beraldin, and M. Rioux, "Direct Estimation of Range Flow on Deformable Shape from a Video Rate Range Camera," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 1, pp. 82-89, Jan. 1993.
- [3] H. Gharavi and S. Gao, "3d Motion Estimation Using Range Data," *IEEE Trans. Intelligent Transportation Systems*, vol. 8, no. 1, pp. 133-143, Mar. 2007.
- [4] H. Spies, H. Haußecker, B. Jähne, and J. Barron, "Differential Range Flow Estimation," *Proc. DAGM Symp.*, pp. 309-316, 1999.
- [5] Y. Zhang and C. Kambhamettu, "Integrated 3D Scene Flow and Structure Recovery from Multiview Image Sequences," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 2674-2681, 2000.
- [6] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade, "Three-Dimensional Scene Flow," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 475-480, Mar. 2005.
- [7] H. Schar, "Towards a Multi-Camera Generalization of Brightness Constancy," *Proc. Int'l Workshop Complex Motion*, 2005.
- [8] J. Barron, D. Fleet, and S. Beauchemin, "Performance of Optical Flow Techniques," *Int'l J. Computer Vision*, vol. 12, no. 1, pp. 43-77, 1994.
- [9] B. Jähne, H. Haußecker, and P. Geissler, *Handbook of Computer Vision and Applications*, first ed. Academic Press, 1999.
- [10] A. Bruhn, "Variational Optic Flow Computation, Accurate Modelling and Efficient Numerics," PhD thesis, Dept. of Math. and Computer Science, Saarland Univ., 2006.
- [11] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert, "Highly Accurate Optic Flow Computation with Theoretically Justified Warping," *Int'l J. Computer Vision*, vol. 67, no. 2, pp. 141-158, 2006.
- [12] T.S.J. Denney and J.L. Prince, "Optimal Brightness Functions for Optical Flow Estimation of Deformable Motion," *IEEE Trans. Image Processing*, vol. 3, no. 2, pp. 178-191, Mar. 1994.
- [13] H. Haußecker and D.J. Fleet, "Computing Optical Flow with Physical Models of Brightness Variation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 661-673, June 2001.
- [14] T. Schuchert and H. Schar, "Simultaneous Estimation of Surface Motion, Depth and Slopes Under Changing Illumination," *Proc. DAGM Symp.*, pp. 184-193, 2007.
- [15] D. Toth, T. Aach, and V. Metzler, "Illumination Invariant Change Detection," *Proc. IEEE Southwest Symp. Image Analysis and Interpretation*, pp. 3-7, 2000.
- [16] A. Oppenheim, R. Schafer, and T.G. Stockham, Jr., "Nonlinear Filtering of Multiplied and Convolved Signals," *Proc. IEEE*, vol. 56, no. 8, pp. 1264-1291, Aug. 1968.

- [17] J.-P. Pons, R. Keriven, and O. Faugeras, "Multi-View Stereo Reconstruction and Scene Flow Estimation with a Global Image-Based Matching Score," *Int'l J. Computer Vision*, vol. 72, no. 2, pp. 179-193, 2007.
- [18] F. Huguet and F. Devernay, "A Variational Method for Scene Flow Estimation from Stereo Sequences," *Proc. IEEE Int'l Conf. Computer Vision*, 2007.
- [19] R. Li and S. Sclaroff, "Multi-Scale 3D Scene Flow from Binocular Stereo Sequences," *Proc. IEEE Workshop Computer Vision*, vol. 2, pp. 147-153, 2005.
- [20] B. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," *Proc. Defense Advanced Research Projects Agency Image Understanding Workshop*, pp. 121-130, 1981.
- [21] J. Bigün, G. Granlund, and J. Wiklund, "Multidimensional Orientation Estimation with Applications to Texture Analysis and Optical Flow," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 775-790, Aug. 1991.
- [22] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods," *Int'l J. Computer Vision*, vol. 61, no. 3, pp. 211-231, 2005.
- [23] M.J. Black and P. Anandan, "The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields," *Computer Vision and Image Understanding*, vol. 63, no. 1, pp. 75-104, 1996.
- [24] S.X. Ju, M.J. Black, and A.D. Jepson, "Skin and Bones: Multi-Layer, Locally Affine, Optical Flow and Regularization with Transparency," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 307-314, June 1996.
- [25] T. Schuchert, T. Aach, and H. Scharr, "Range Flow for Varying Illumination," *Proc. European Conf. Computer Vision*, pp. 509-522, 2008.
- [26] B. Jähne, *Digitale Bildverarbeitung*, fourth ed. Springer-Verlag, 1997.
- [27] H. Arsenault and M. Denis, "Image Processing in Signal-Dependent Noise," *Canadian J. Physics*, vol. 61, pp. 309-317, 1983.
- [28] H. Scharr and T. Schuchert, "Simultaneous Motion, Depth and Slope Estimation with a Camera-Grid," *Proc. Vision, Modeling, and Visualization '06*, pp. 81-88, 2006.
- [29] H. Scharr, "Optimal Filters for Extended Optical Flow," *Proc. Int'l Workshop Complex Motion*, pp. 14-29, 2004.
- [30] H. Spies, B. Jähne, and J.L. Barron, *Surface Expansion from Range Data Sequences*, pp. 163-169. Springer-Verlag, 2001.
- [31] O. Nestares, D.J. Fleet, and D.J. Heeger, "Likelihood Functions and Confidence Bounds for Total Least Squares Estimation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2000.



**Til Aach** received the diploma and doctoral degrees, both with honors, in electrical engineering from RWTH Aachen University in 1987 and 1993, respectively. While working toward the doctoral degree, he was a research scientist with the Institute for Communications Engineering, RWTH Aachen University, being in charge of several projects in image analysis, 3D television, and medical image processing. In 1984, he was an intern with Okuma Machinery Works, Ltd., Nagoya, Japan. From 1993 to 1998, he was with Philips Research Labs, Aachen, Germany, where he was responsible for several projects in medical imaging, image processing, and analysis. In 1996, he was also an independent lecturer with the University of Magdeburg, Germany. In 1998, he was appointed a full professor and director of the Institute for Signal Processing, University of Lübeck. In 2004, he became chairman of the Institute of Imaging and Computer Vision, RWTH Aachen University. His research interests include medical and industrial image processing, signal processing, pattern recognition, and computer vision. He has authored or coauthored more than 250 papers, and received several awards, among these the award of the German "Informationstechnische Gesellschaft" (ITG/VDE), for a paper published in the *IEEE Transactions on Image Processing* in 1998. He is a co-inventor for about 20 patents. From 2002 to 2008, he was an associate editor of the *IEEE Transactions on Image Processing*. He was a technical program cochair for the IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI) in 2000, 2002, 2004, and 2006. He is a member of the Bio-Imaging and Signal Processing Committee (BISP-TC) of the IEEE Signal Processing Society. He is a senior member of the IEEE.



**Hanno Scharr** received the diploma and doctoral degrees in physics (Dr. rer. nat.) from Ruprecht-Karls-University, Heidelberg, Germany, in 1996 and 2000, respectively. After receiving the PhD degree, he joined the Computational Nano-Vision Group at Intel Corporation, Santa Clara, California, where he was a senior researcher. Currently, he is the head of the Digital Image Processing Group at the Institute for Chemistry and Dynamics of the Geosphere, Institute 3: Phytosphere at Forschungszentrum Jülich, Germany. He has held a university teaching position at Johann Wolfgang Goethe University Frankfurt, Germany, since 2006, and was awarded a Junior Research Group in 2007. His research interests include quantitative signal and image processing in multidimensional data sets as an analysis tool for scientific, medical, and industrial data. He is the author or coauthor of more than 50 scientific publications and patents.

► For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/publications/dlib](http://www.computer.org/publications/dlib).



**Tobias Schuchert** received the diploma degree in electrical engineering from TU Dortmund, Germany, in 2005 and the doctoral degree from RWTH Aachen University, Germany, in 2010. While working toward the doctoral degree, he was a research scientist with the Digital Image Processing Group at the Institute for Chemistry and Dynamics of the Geosphere, Institute 3: Phytosphere at Forschungszentrum Jülich, Germany. Currently, he is a research scientist at the

Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB in the Department of Autonomous Systems and Machine Vision (ASM). His research interests include signal and image processing, pattern recognition, and computer vision. He is the author or coauthor of six scientific publications.