Range Flow for Varying Illumination

Tobias Schuchert[1], Til Aach[2], and Hanno Scharr[1]

[1] Institute for Chemistry and Dynamics of the Geosphere, ICG-3: Phytosphere, Forschungszentrum Jülich, Germany
{t.schuchert,h.scharr}@fz-juelich.de
[2] Institute of Imaging & Computer Vision, RWTH Aachen University, Germany
til.aach@lfb.rwth-aachen.de

Abstract. In this paper range flow estimation is extended to handle brightness changes in image data caused by inhomogeneous illumination. Standard range flow computes 3d velocity fields from range and intensity image sequences. To this end it combines a depth change model and a brightness constancy model. In this contribution, the brightness constancy model is exchanged by (1) a gradient constancy model, (2) a combination of gradient and brightness constancy constraint that has been used successfully for optical flow estimation in literature, and (3) a physics-based brightness change model. Insensitivity to brightness changes can also be achieved by prefiltering of the input intensity data. High pass or homomorphic filtering are the most well known approaches from literature. In performance tests therefore the well known version and the novel versions of range flow estimation are investigated on prefiltered or non-prefiltered data using synthetic ground-truth and real data from a botanical experiment.

1 Introduction

In this paper influences of brightness models in 3d velocity field estimation are investigated. The brightness model is only one module of usual motion estimation methods. Typical methods consist of data prefiltering, brightness constraint equations describing imaging physics locally (see e.g. [1,2]) calculated by suitable convolution filtering [3,4], parameter estimation scheme like local least squares [5], local total least squares [6], or variational approaches [7]. Occlusions and other model violations are typically handled using robust error norms instead of plain least squares [8,9]. Variational estimators in addition allow for incorporation of prior knowledge, e.g. in form of regularization terms closing holes or reducing so-called aperture problems.

Motivation of this work is a target application, namely plant growth estimation. 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* for this application, but isolates influences of brightness change models frequently occurring in such botanical but also other data.

A lot of work has been carried out on estimating 3d motion fields. Range flow estimation [10,11] uses data solely from range sensors whereas [12,13] incorporate information from both range and image sensors. Reconstruction of

D. Forsyth, P. Torr, and A. Zisserman (Eds.): ECCV 2008, Part I, LNCS 5302, pp. 509–522, 2008.

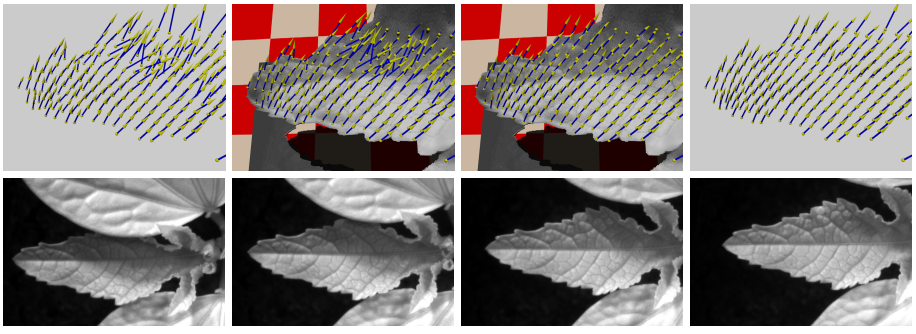

Fig. 1. Castor bean plant leaf. Top: Estimated motion vector fields. The two images to the left and to the right show the same results, respectively, but with and without 3d structure shown. Left images: standard range flow. Right: proposed TAYLOR model. Bottom: Images of the input sequence.

scene flow and 3d structure from the observed optical flow in several cameras has been proposed by [14,15,16]. These scene flow and most optical flow based approaches [17,18] imply brightness constancy and are therefore not suitable if substantial brightness changes are present in a sequence (cmp. Fig. 1, upper part of the shown leaf). For optical flow estimation less brightness sensitive models, e.g. constancy of intensity gradient vector [19,20] have been proposed, as well as physics-based brightness change models [21,1]. The physics-based models of [1] have been recently adapted and extended for moving surfaces under inhomogeneous illumination [2]. Suppressing brightness changes by prefiltering the intensity data has been shown to be very efficient. One of the simplest approaches is applying a spatial high-pass filter to minimize the effect of global brightness inhomogeneities. Toth et al [22] show that using homomorphic prefilters [23] can highly improve motion detection in image sequences with inhomogeneous illumination. More recent approaches for scene flow estimation use statistical similarity measures [24], a gradient constancy constraint [25] or probability distributions for optical flow and disparity [26] 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 is extended to be able to handle inhomogeneous illumination. To this end the techniques known for optical flow estimation are introduced in the range flow constraints, namely (1) gradient constancy [19] and (2) mixture of brightness and gradient constancy [20] as well as physics-based brightness modeling [2]. On the other hand performance of standard and the novel range flow models is investigated. To this end the models are tested on two kinds of synthetic data sets with ground truth available. These sequences are either used with or without suppression of illumination inhomogeneities by high-pass or homomorphic filtering. A motion estimation result for a "real" image sequence from a botanical experiment on leaf growth is shown in Fig. 1. Performance experiments focus on and therefore are especially designed for *brightness model influences*.

Input data without occlusions, inner borders, holes or aperture problems has been selected. Thus *no robust estimators* handling occlusion or *variational estimators* closing holes are needed, but local least squares estimation on a large neighborhood is suitable. In a final system robust and/or variational estimators may be used of course if needed, but not for investigations of *model* influences.

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

2 Range Flow

Range flow as used here is based on two motion constraints: one for range data and one for intensity data. Following [13], we briefly derive these two constraints in this section.

2.1 Range Constraint

Let a surface be described by its depth Z as a function of space and time $Z = Z(X, Y, t)$, where X, Y and Z are spatial coordinates and t denotes time. We select X and Y being aligned with camera sensor coordinates x and y, respectively. Z-axis points along the principal axis of the assumed projective camera. The total derivative of Z with respect to time then yields the so called range flow motion constraint equation

$$\frac{dZ}{dt} = \partial_X Z \frac{dX}{dt} + \partial_Y Z \frac{dY}{dt} + \partial_t Z \tag{1}$$

where partial derivatives are denoted as e.g. $\partial_X Z = \frac{\partial Z}{\partial X}$. Range flow is defined to be $f = [U, V, W]^{\mathrm{T}} := [\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt}]^{\mathrm{T}}$.

Range data is given as data sets $X = X(x, y, t)$, $Y = Y(x, y, t)$ and $Z = Z(x, y, t)$ on the sampling grid. Partial derivatives are not computed on world coordinate data but directly on the sensor grid in order to avoid 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

$$U = \partial_x X \dot{x} + \partial_y X \dot{y} + \partial_t X \tag{2}$$

$$V = \partial_x Y \dot{x} + \partial_y Y \dot{y} + \partial_t Y \tag{3}$$

$$W = \partial_x Z \dot{x} + \partial_y Z \dot{y} + \partial_t Z \tag{4}$$

where total derivatives with respect to time are indicated by a dot. Being not interested in the change on the sensor grid, i.e. optical flow, $\dot{x}$ and $\dot{y}$ can be eliminated. This yields

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

where e.g.

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

is the Jacobian of Z, Y with respect to x, y. Equation (5) only depends 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

$$\begin{aligned}(\partial_y Y \partial_x Z)U + (\partial_x X \partial_y Z)V - (\partial_x X \partial_y Y)W \\ + (\partial_x X \partial_y Y \partial_t Z - \partial_x X \partial_t Y \partial_y Z - \partial_t X \partial_y Y \partial_x Z) = 0\,. \end{aligned} \tag{7}$$

2.2 Intensity Constraint

The range flow constraint is solely for range data and full flow can only be estimated for corners or point-like structures. Plant surfaces are often nearly planar, smooth surfaces resulting in aperture problems almost everywhere when using solely the range flow constraint. As proposed in [12] intensity data should be 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 (cmp. e.g. [17]) is valid. Linearization of this constraint yields

$$\frac{dI}{dt} = \partial_x I \dot{x} + \partial_y I \dot{y} + \partial_t I = 0\,. \tag{8}$$

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

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

The estimated motion $[U, V, W]$ has to fulfill both constraints, the range flow constraint (5) and the intensity constraint (9). The intensity constraint more reliably yields point-to-point correspondences and therefore often solves the aperture problem. Together with the range constraint it allows also to solve for the vertical motion W. Combining range and intensity constraint and estimation of f via total least squares is shown in Sec. 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 well fulfilled anymore. In the following section different approaches to handle illumination changes are presented. Three of them lead to constraints novel in range flow estimation.

3.1 Prefiltering

A well known technique for illumination change suppression is suitable prefiltering making data illumination invariant.

Temporal and/or spatial high-pass filtering eliminates slow brightness changes in the data. However faster illumination changes both in spatial and temporal domain still remain in the data. In our experiments we test high-pass filtering with the filter

$$\tilde{I} = (1 - B_{11}) * I \tag{10}$$

where B_{11} denotes a spatial 11-tab binomial filter (see e.g. [27]) and $*$ is convolution.

A more sophisticated approach is using a homomorphic filter [23]. Following [22] we briefly derive a simple implementation of homomorphic filtering, which proved to be very successful in suppressing illumination changes. Homomorphic filtering uses the fact that image intensity $I(x, y)$ is proportional to incident illumination intensity $E(x, y)$, which is reflected by object surfaces with reflectance $R(x, y)$ in the observed scene. For Lambertian surfaces image intensity can be modeled as

$$I(x, y) \propto E(x, y) \cdot R(x, y) \; . \tag{11}$$

Reflectance R is the desired component for motion estimation as this part contains the structure of the scene and is temporally invariant. The logarithm transforms the multiplicative relation between illumination intensity E and reflectance R into an additive one

$$log(I(x, y)) \propto log(E(x, y)) + log(R(x, y)) \; . \tag{12}$$

Ideally $log(E)$ and $log(R)$ are separated in frequency domain as E is assumed to be low-frequent and R mainly high-frequent. In practice the two components overlap. Thus a tradeoff between suppressing brightness changes and loss of signal has to be made. However high-pass filtering more efficiently suppresses the illumination component after taking the logarithm of the signal. However, as the nonlinear log-operation makes the camera noise variance signal-dependent, therefore influencing parameter estimation, exponentiation returns an approximation of the sought for reflectance component.

3.2 Gradient Constancy Constraint

The intensity constraint (9) can be exchanged by an illumination invariant or insensitive constraint instead of or in addition to prefiltering the data.

A method reported to be successful in the pure optical flow case [20] is to assume that 2d image intensity gradient should not change along the motion trajectory. Derivative filtering is a highpass operation therefore also suppressing illumination changes (see Sec. 3.1). This leads to two linearized gradient constancy constraints

$$\frac{dI_x}{dt} = \partial_x I_x \dot{x} + \partial_y I_x \dot{y} + \partial_t I_x = 0 \tag{13}$$

$$\frac{dI_y}{dt} = \partial_x I_y \dot{x} + \partial_y I_y \dot{y} + \partial_t I_y = 0 \tag{14}$$

where lower indices indicate partial derivatives, e.g. $I_x := \partial_x I$, as before. Analogous to derivation of (9) optical flow $(\dot{x}, \dot{y})$ is eliminated using (2) and (3)

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

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

3.3 Combined Intensity and Gradient Constancy Constraint

A known drawback of the gradient constancy constraint proposed in Sec. 3.2 is that it noticeably reduces the structure in the image and leads to aperture problems. Using both the intensity constraint and the gradient constraint simultaneously has reduced this effect in the optical flow case [20]. This leads to three constraint equations ((9), (15) and (16)) which should be fulfilled simultaneously for the horizontal and vertical range flow components U and V.

3.4 Physics Based Brightness Change Model

A different approach to handle brightness changes is to model them explicitly and estimate both optical flow and brightness change parameters. Haussecker and Fleet [1] 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 $\mathbf{x}(t) = (x(t), y(t))^{\mathrm{T}}$. This is described by a parameterized function h_I

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

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

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

Assuming brightness constancy, i.e. $h_I(I_0, t, \mathbf{a}) = c$, (18) reduces to (8). 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 illumination change models are proposed in [1], i.e. changing surface orientation, motion of the illuminant, and physical models of heat transport in infrared images. We use a brightness change model presented in [2] which handles spatially varying time-dependent illumination changes coming from directed, inhomogeneous illumination and changing surface orientation. In [2] the brightness change function is set to

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

The incident irradiance caused by the moving illuminant is assumed to be spatially inhomogeneous, therefore changing not only by a time dependent factor, but also varying smoothly in space. Approximating these brightness changes by a second order Taylor series yields

$$h(\Delta X, \Delta Y, t, \mathbf{a}) := \sum_{i=1}^{2} \left(a_i + a_{i,x}\Delta X + a_{i,y}\Delta Y\right) t^i \tag{20}$$

with spatial neighborhood $\Delta X, \Delta Y$ and temporal derivative

$$\dot{h}(\Delta X, \Delta Y, t, \mathbf{a}) = \sum_{i=1}^{2} i \left(a_i + a_{i,x}\Delta X + a_{i,y}\Delta Y\right) t^{i-1} \tag{21}$$

using the notation $\mathbf{a} = [a_1, a_{1,x}, a_{1,y}, a_2, a_{2,x}, a_{2,y}]^{\mathrm{T}}$. Analogous to Sec. 2.2 we get the total differential

$$\frac{dI}{dt} = \partial_x I\dot{x} + \partial_y I\dot{y} + \partial_t I = I\dot{h}(\Delta X, \Delta Y, t, \mathbf{a}) \tag{22}$$

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

$$\begin{aligned}\frac{\partial(I,Y)}{\partial(x,y)}U + \frac{\partial(X,I)}{\partial(x,y)}V + \frac{\partial(X,Y,I)}{\partial(x,y,t)} - Ia_1 - Ia_{1,x}\Delta X \\ - Ia_{1,y}\Delta Y - 2Ia_2 t - 2Ia_{2,x}\Delta X t - 2Ia_{2,y}\Delta Y t = 0\,.\end{aligned} \tag{23}$$

4 Parameter Estimation

For parameter estimation in a total least squares framework we closely follow [13]. The range constraint (see Sec. 2.1) yields for every pixel an equation of the form $\mathbf{d}_{rc}^{\mathrm{T}}\mathbf{p} = 0$ with

$$\mathbf{d}_{rc} = \left[\frac{\partial(Z,Y)}{\partial(x,y)}, \frac{\partial(X,Z)}{\partial(x,y)}, \frac{\partial(Y,X)}{\partial(x,y)}, \frac{\partial(X,Y,Z)}{\partial(x,y,t)}\right]^{\mathrm{T}} \tag{24}$$

and $\mathbf{p} = [U, V, W, 1]^{\mathrm{T}}$. To solve this equation containing three unknowns, we assume that within a local neighborhood Ω one parameter vector $\mathbf{p}$ solves all equations but for an error $\mathbf{e}$. Minimizing error $\mathbf{e}$ in weighted L_2-norm yields

$$||\mathbf{e}|| = \mathbf{p}^{\mathrm{T}}\mathbf{J}_{\mathbf{rc}}\mathbf{p} \stackrel{!}{=} \min \tag{25}$$

with structure tensor $\mathbf{J}_{\mathbf{rc}} = \mathbf{W} * (\mathbf{d}_{rc}\mathbf{d}_{rc}^{\mathrm{T}})$ and averaging filter $\mathbf{W}$ defining the neighborhood Ω. As described in Secs. 2.2 and 3.2 to 3.4 we have more than one constraint for the U and V component of the range flow. Analogous to the range constraint $\mathbf{d}_{rc}^{\mathrm{T}}\mathbf{p} = 0$ the intensity (9) and the gradient constancy constraints (15) and (16) may be formulated as $\mathbf{d}_{Q}^{\mathrm{T}}\mathbf{p} = 0$ using

$$\mathbf{d}_Q = \left[\frac{\partial(Q,Y)}{\partial(x,y)}, \frac{\partial(X,Q)}{\partial(x,y)}, 0, \frac{\partial(X,Y,Q)}{\partial(x,y,t)}\right]^{\mathrm{T}} \tag{26}$$

with $Q = \{I, I_x, I_y\}$ respectively. By inserting zeros into the appropriate places of data vector $\mathbf{d}$ we adapt all constraints to the same dimensions. The brightness change model presented in Sec. 3.4 contains motion and brightness change parameters. Therefore both data vectors for the range and the brightness change constraint have to be enlarged by zeros appropriately. Parameter vector $\mathbf{p}$ then becomes $\mathbf{p} = [U, V, W, 1, \mathbf{a}^{\mathrm{T}}]^{\mathrm{T}}$.

As in [13] we combine the different constraints yielding a combined structure tensor that is simply the weighted sum of the different tensors of the depth and the intensity channels

$$\mathbf{J} = \mathbf{J_{rc}} + \sum_{i=1}^{j} \beta_i \mathbf{J_i} \tag{27}$$

weighted with constants β_i and 3 possible choices for j, namely for $j = 1, 2$, or 3, depending on the number of brightness constraint equations used and corresponding to the constraints proposed in Secs. 2.2 and 3.2 to 3.4. Constants β_i may be used to account for different signal-to-noise-ratios of the structure tensors. Furthermore the data channels should be scaled to same mean and variance before they are combined.

As is well-known, this equation is minimized by the eigenvector $\mathbf{b}$ to the smallest eigenvalue of $\mathbf{J}$. Range flow is then given by

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \frac{1}{b_4} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} . \tag{28}$$

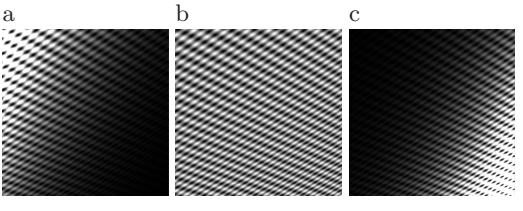

Fig. 2. Scaled first (a), central (b) and last (c) frame of sinusoidal sequence with illumination parameters $a_1 = 0$ and $a_{1,x} = 0.06$

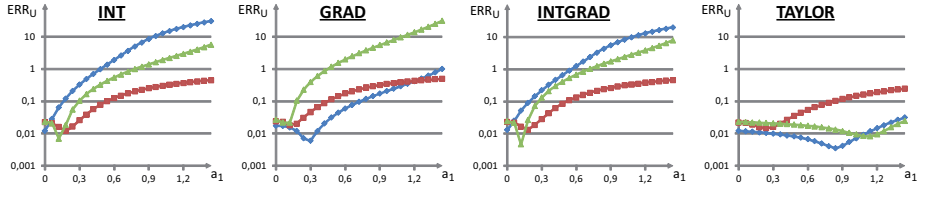

Fig. 3. Mean absolute value of relative error of U versus the brightness change parameter a_1 with no (-♦-), highpass (-▲-) and homomorphic (-■-) prefilters for different models

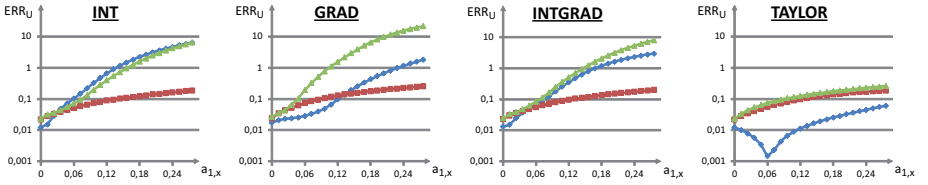

Fig. 4. Mean absolute value of relative error of U versus the brightness change parameter $a_{1,x}$ with no (-◆-), highpass (-▲-) and homomorphic (-■-) prefilters for different models

5 Experiments

For systematic error analysis of the different models sinusoidal patterns are used. The models compared are combinations of the range constraint with

1. the intensity constancy constraint (INT, (9)),
2. the gradient constancy constraint (GRAD, (15) and (16)),
3. the combined intensity and gradient constancy constraint (INTGRAD, (9), (15) and (16)) and
4. the intensity constraint with modeling of brightness changes by taylor series (TAYLOR, (23)).

Further accuracy of motion estimates on a rendered cube illuminated by a directed light source and motion estimation results on real data are demonstrated. For all experiments we generate range data by multi camera stereo reconstruction as presented in [28]. The weighting constants are set to $\beta_i = 1$ in all calculations.

5.1 Sinusoidal Pattern

For evaluation of the different models we modeled moving patches with sinusoidal patterns under varying illumination. Three frames of one test sequence are shown in Fig. 2. For our analysis we use patches translating with $U = 0.0073\,\text{mm/frame}$, $V = 0\,\text{mm/frame}$ and $W = 0.5\,\text{mm/frame}$ and angular velocity of $\omega = 0.002\,\text{radians/frame}$ around the Y-axis. For $t = 0$ the surface normal of the patch is $\mathbf{n} = (1, 2, -1)^{\mathrm{T}}$ and the distance of the patch center to the camera is $Z_0 = 100\,\text{mm}$. The synthetic sensor contains 301×301 pixels of size $0.0044\,\text{mm}^2$. Focal length of the synthetic projective camera is $f = 12\,\text{mm}$. For our analysis we compare the mean absolute value of the relative error of U

$$ERR_U = \frac{1}{N} \sum_{i}^{N} \frac{|U_{estimated} - U_{reference}|}{|U_{reference}|} \tag{29}$$

over all pixel N at a minimum distance of 60 pixel from the nearest image border. To reduce systematic errors the structure tensor weighting matrix $\mathbf{W}$ is realized by a 65-tab Gaussian with standard deviation $\sigma = 19$. Both prefilters

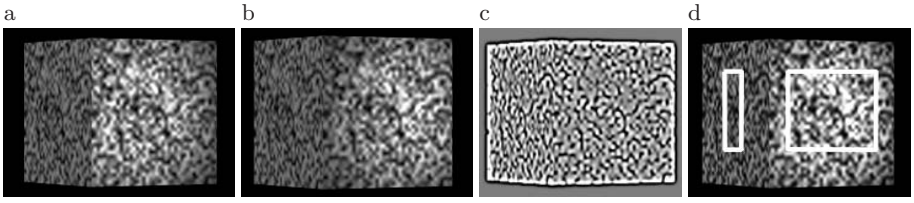

Fig. 5. First (a) and last (b) frame of cube sequence, (c) central frame of cube sequence after homomorphic prefiltering, (d) regions of cube used for error analysis

are generated as depicted in Sec. 3.1. Following [2] we compare the estimation of U for increasing illumination parameters $a_1|_{a_{1,x}=0}$ and $a_{1,x}|_{a_1=0}$ to simulate brightness changes. Only errors for U are presented, as errors of V and W showed similar characteristics.

In Fig. 3 and Fig. 4 errors of U for the proposed models in combination with prefilters are shown for increasing a_1 and $a_{1,x}$ respectively. Using homomorphic prefilters performances of all models are comparable. For models INT and INTGRAD homomorphic prefiltering is the best choice. For small brightness changes model GRAD with no prefilter produces even lower errors than using the homomorphic one. Lowest errors can be observed using model TAYLOR without prefilter. Overall these errors are about one order of magnitude smaller than for all other model and prefilter combinations over a wide range of brightness changes.

For models INT and INTGRAD the positive effect of reducing brightness changes using homomorphic prefiltering apparently overcompensates the negative effect of losing signal. For models GRAD and TAYLOR, which can handle brightness changes better, it seems to be the other way round. Reduced accuracy for non-changing brightness ($a_1 = 0$ and $a_{1,x} = 0$) proves the negative effect of loss of signal due to prefiltering.

5.2 Synthetic Cube

The synthetic cube sequence allows to test the models on more realistic data with groundtruth available. The cube moves with $U = -0.2\,\mathrm{mm/frame}$, $V = 0\,\mathrm{mm/frame}$ and $W = -2\,\mathrm{mm/frame}$. In addition to ambient light the cube is illuminated by a fixed light spot from the right. Figure 5 shows two frames of the cube sequence, a frame after homomorphic prefiltering as well as the investigated regions on the left and right side of the cube. The size of the weighting matrix $\mathbf{W}$ is realized by a 31-tab Gaussian with standard deviation $\sigma = 11$.

In Fig. 6 motion estimates of the standard range flow model INT is compared with the three models, which performed best on sinusoidal test sequences, i.e. model INTGRAD with homomorphic prefilter and both models GRAD and TAYLOR without prefilter. As errors are too small to be visible for all models the erroneous motion estimates are amplified for U and V by 50 and W by 20.

Table 1 shows numerical errors of the different models for the regions on the left and right side of the cube. We compare the average angular error [17]

$$AAE = \arccos\left(\frac{f_c f_e}{|f_c|\,|f_e|}\right)\,[^\circ] \tag{30}$$

and its standard deviation with the true and the estimated flow f_c and f_e respectively. Interested in estimating plant growth we furthermore compare the average relative growth rate and its standard deviation. According to [29] the relative change in size dA of a local surface area s parameterized in sensor coordinates x and y is calculated using the 3d displacement vector field f

$$dA = \frac{|s(x+1,y)+f(x+1,y) - (s(x,y)+f(x,y))| \times |s(x,y+1) + f(x,y+1) - (s(x,y)+f(x,y))|}{|s(x+1,y) - s(x,y)| \times |s(x,y+1) - s(x,y)|} \tag{31}$$

with $s(x,y) = [X(x,y), Y(x,y), Z(x,y)]^{\mathrm{T}}$. The relative growth rate is determined by $RGR = (dA-1)\cdot 100\%$. In the experiment the 3d structure of the cube remains constant in time, i.e. $RGR = 0\%$.

Additionally we show results obtained by the recent scene flow approach of [25]. We apply the algorithm the authors provide with the parameters of the *rotating sphere* experiment. Only the weighting parameter of the gradient constraint was increased to $\gamma = 30$ to reduce the effect of the severe brightness variations in the data.

Standard range flow estimation, i.e. model INT without prefilter (Fig. 6a) yields highly corrupted estimation results on the side of the cube where illumination changes due to the fixed spotlight. The other side does not suffer from illumination changes. There motion estimates are much more accurate. As expected estimates are well improved by the other models when brightness changes are present.

But using a homomorphic prefilter, e.g. with model INTGRAD (Fig. 6c), results on the left side of the cube are visibly worse than using standard range flow. Models without prefilter (Fig. 6b+d) or with a highpass yield more or less the same good results as standard range flow on the left side of the cube. These observations coincide with the errors in Table 1.

On its right side, where the brightness changes dominate, model TAYLOR (Fig. 6d) yields slightly more uniform, more accurate motion vectors than model GRAD (Fig. 6b). Moreover Table 1 shows that model TAYLOR is more robust with respect to prefiltering. With no prefilter or using a highpass rarely changes results, whereas the other models yield much more varying errors depending on the prefilter. The errors of [25] are much worse than any other used model, despite the elaborate estimator. We assume that tuning parameters may reduce errors a bit but did not find a good parameter set. This shows that using an elaborate estimator does not necessarily help, if the brightness model does not fit.

We conclude that homomorphic prefiltering should be avoided if possible and modeling brightness changes yields slightly more accurate results than using a model suppressing effects of brightness changes.

Table 1. Average angular error, average relative growth rate and their standard deviations of regions on left and right side of the cube (see Fig. 5d). High and low errors indicated in red and green respectively.

model	prefilter	left region AAE	left region RGR	right region AAE	right region RGR
INT	NO	0.148 ± 0.074	0.093 ± 0.092	4.311 ± 4.024	4.372 ± 9.211
	HP	0.137 ± 0.074	0.105 ± 0.097	0.069 ± 0.045	0.017 ± 0.057
	HOM	0.322 ± 0.216	0.087 ± 0.429	0.096 ± 0.148	-0.175 ± 0.564
GRAD	NO	0.142 ± 0.075	0.097 ± 0.094	0.120 ± 0.111	0.042 ± 0.102
	HP	0.135 ± 0.072	0.110 ± 0.098	0.046 ± 0.017	-0.003 ± 0.057
	HOM	0.664 ± 0.471	0.039 ± 0.862	0.139 ± 0.280	-0.29 ± 0.831
INTGRAD	NO	0.147 ± 0.074	0.094 ± 0.092	4.400 ± 4.081	4.500 ± 9.356
	HP	0.137 ± 0.073	0.107 ± 0.097	0.085 ± 0.066	0.028 ± 0.061
	HOM	0.240 ± 0.146	0.090 ± 0.316	0.084 ± 0.110	-0.14 ± 0.465
TAYLOR	NO	0.0149 ± 0.074	0.090 ± 0.094	0.055 ± 0.014	-0.009 ± 0.052
	HP	0.139 ± 0.074	0.105 ± 0.097	0.047 ± 0.016	-0.012 ± 0.051
	HOM	0.323 ± 0.216	0.087 ± 0.431	0.096 ± 0.149	-0.02 ± 0.565
[25]	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

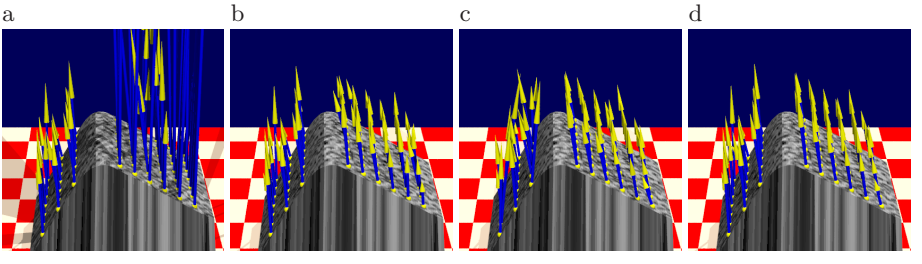

Fig. 6. Scaled motion estimates with amplified errors for different models: (a) INT without prefiltering, (b) GRAD without prefiltering, (c) INTGRAD with homomorphic prefiltering, (d) TAYLOR without prefiltering

5.3 Real Data

The previous experiments showed that model TAYLOR yields most reliable estimates. In Fig. 1 we show estimated 3d velocity fields for a freely moving castor bean plant 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 1 frame per 2 minutes was taken, i.e. acquisition time for the used images was 16 minutes.

The leaf rotates around the node where it is attached to the stem. This results in a visible motion towards 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. As expected for a rigid motion, which can be assumed using a high temporal resolution we obtain a smoothly varying vector field.

6 Summary and Outlook

In this paper we extended range flow estimation presented in [13] with different approaches to handle inhomogeneous illumination. We presented a detailed error analysis for four different model constraints in combination with highpass and homomorphic prefilters on synthetic sequences with sinusoidal patterns and a translating cube. While prefilters improved estimation results on data when illumination changes are present, they suppress information producing worse results when no changes are present. Modeling brightness changes instead of prefiltering provides equal or better estimation results whether illumination changes occur or not.

In future work we plan to further improve the accuracy of the model and the estimator.

Acknowledgments. The authors would like to thank Georg Dreissen for his help with the acquisition of the plant leaf sequence.

References

1. Haußecker, H., Fleet, D.J.: Computing optical flow with physical models of brightness variation. PAMI 23, 661–673 (2001)
2. Schuchert, T., Scharr, H.: Simultaneous estimation of surface motion, depth and slopes under changing illumination. In: Hamprecht, F.A., Schnörr, C., Jähne, B. (eds.) DAGM 2007. LNCS, vol. 4713, pp. 184–193. Springer, Heidelberg (2007)
3. Simoncelli, E.P.: Design of multi-dimensional derivative filters. In: ICIP (1), pp. 790–794 (1994)
4. Scharr, H.: Optimal filters for extended optical flow. In: International Workshop on Complex Motion, pp. 14–29 (2004)
5. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: DARPA Im. Underst. Workshop, pp. 121–130 (1981)
6. Bigün, J., Granlund, G., Wiklund, J.: Multidimensional orientation estimation with applications to texture analysis and optical flow. IEEE Trans. PAMI 13, 775–790 (1991)
7. Bruhn, A., Weickert, J., Schnörr, C.: Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. IJCV 61, 211–231 (2005)
8. Black, M.J., Anandan, P.: The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. Comput. Vis. Image Underst. 63, 75–104 (1996)
9. Ju, S.X., Black, M.J., Jepson, A.D.: Skin and bones: Multi-layer, locally affine, optical flow and regularization with transparency. In: Proc. Computer Vision and Pattern Recognition, CVPR-1996, San Francisco, pp. 307–314 (1996)
10. Yamamoto, M., Boulanger, P., Beraldin, J.A., Rioux, M.: Direct estimation of range flow on deformable shape from a video rate range camera. IEEE PAMI 15, 82–89 (1993)
11. Gharavi, H., Gao, S.: 3-d motion estimation using range data. IEEE Transactions on intelligent transportation systems 8, 133–143 (2007)
12. Spies, H., Haußecker, H., Jähne, B., Barron, J.: Differential range flow estimation. In: DAGM, pp. 309–316 (1999)
13. Spies, H., Jähne, B., Barron, J.: Range flow estimation. CVIU 85, 209–231 (2002)

14. Zhang, Y., Kambhamettu, C.: Integrated 3d scene flow and structure recovery from multiview image sequences. In: CVPR, pp. 2674–2681 (2000)
15. Vedula, S., Baker, S., Rander, P., Collins, R., Kanade, T.: Three-dimensional scene flow. IEEE PAMI 27, 475–480 (2005)
16. Scharr, H.: Towards a multi-camera generalization of brightness constancy. In: Jähne, B., Mester, R., Barth, E., Scharr, H. (eds.) IWCM 2004. LNCS, vol. 3417, pp. 78–90. Springer, Heidelberg (2007)
17. Barron, J., Fleet, D., Beauchemin, S.: Performance of optical flow techniques. IJCV 12(1), 43–77 (1994)
18. Jähne, B., Haußecker, H., Geissler, P.: Handbook of Computer Vision and Applications, 1st edn. Academic Press, London (1999)
19. Bruhn, A.: Variational optic flow computation, accurate modelling and efficient numerics (2006)
20. Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J.: Highly accurate optic flow computation with theoretically justified warping. IJCV 67, 141–158 (2006)
21. Denney, T.S.J., Prince, J.L.: Optimal brightness functions for optical flow estimation of deformable motion. IEEE Trans. Im. Proc. 3, 178–191 (1994)
22. Toth, D., Aach, T., Metzler, V.: Illumination-invariant change detection. In: SSIAI, pp. 3–7 (2000)
23. Oppenheim, A., Schafer, R., Stockham, T.G.: Nonlinear filtering of multiplied and convolved signals. In: Proceedings of the IEEE, vol. 56, pp. 1264–1291 (1968)
24. Pons, J.P., Keriven, R., Faugeras, O.: Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. IJCV 72(2), 179–193 (2007)
25. Huguet, F., Devernay, F.: A variational method for scene flow estimation from stereo sequences. In: ICCV (2007)
26. Li, R., Sclaroff, S.: Multi-scale 3d scene flow from binocular stereo sequences. wacv-motion 2, 147–153 (2005)
27. Jähne, B.: Digitale Bildverarbeitung, 4th edn. Springer, Heidelberg (1997)
28. Scharr, H., Schuchert, T.: Simultaneous motion, depth and slope estimation with a camera-grid. In: Vision, Modeling and Visualization 2006, pp. 81–88 (2006)
29. Spies, H., Jähne, B., Barron, J.L.: Surface expansion from range data sequences. In: Radig, B., Florczyk, S. (eds.) DAGM 2001. LNCS, vol. 2191, pp. 163–169. Springer, Heidelberg (2001)