# ESTIMATION OF TRANSPARENT MOTIONS WITH PHYSICAL MODELS FOR ADDITIONAL BRIGHTNESS VARIATION

*Hanno Scharr[1], Ingo Stuke[2], Cicero Mota[3], and Erhardt Barth[3]*

[1]ICG III, Forschungszentrum Jülich GmbH, Jülich, Germany
H.Scharr@FZ-Juelich.de
[2]Institute for Signal Processing, University of Lübeck, Germany
[3]Institute for Neuro- and Bioinformatics, University of Lübeck, Germany

## ABSTRACT

The paper deals with the estimation of complex motion patterns. The complexity is due to (i) the motions of two transparent layers, and (ii) an additional change of brightness in the layers, which can be due to an additive source term, an exponential decay, or diffusion. We present new models and constraints for such complex motion patterns. Experiments on synthetic image sequences demonstrate the performance of our models in conjunction with a total least-squares parameter estimation scheme. Crucial ingredients of this scheme are new filter families of derivative filters of up to fourth order. We present a procedure for how to construct appropriate filter families for the introduced models.

## 1. INTRODUCTION

We present linear models for the estimation of transparent motions and physically motivated brightness changes. This work combines model design as presented in [3] and [7] with estimation approaches from [5] and discretization via optimised filter families [6].

The models introduced in Sec. 2 are all of the form $\mathbf{d}^T\mathbf{p}=0$ with a parameter vector $\mathbf{p}$ and a data vector $\mathbf{d}$. This parameter vector does not contain the model parameters directly, but in the form of *mixed motion parameters* [5]. We show how to disentangle the parameters, given an estimated parameter vector $\mathbf{p}$. Sec. 3 introduces the so called structure tensor and gives a rule for how to construct suitable filter families. Finally, the experimental results are presented in Sec. 4.

## 2. CONSTRUCTION OF THE MODELS

We will here derive brightness-change constraint equations (BCCE) from motion models by combining well known BCCE (see e.g. [3]) with the model of transparent motions first presented in [7]. We start with the simplest case, two transparent motions without brightness changes.

### 2.1 Transparent Motions

The BCCE for standard, single motion optical flow is

$$(\mathbf{v})f=0 \text{ where } \quad (\mathbf{v}) := v_x\ {}_x + v_y\ {}_y + \ {}_t \tag{1}$$

with image intensities $f$, partial derivatives ${}_x$, ${}_y$, ${}_t$ in $x$-, $y$- and $t$-directions, respectively, and displacement vector $\mathbf{v}=[v_x,v_y]^T$. We construct a BCCE for transparent motions by successively applying to the image sequence [7]. Thus if

$$f(\mathbf{x},t)=f_1(\mathbf{x}-\mathbf{u}t)+f_2(\mathbf{x}-\mathbf{v}t), \tag{2}$$

a basic calculation reveals that $(u)\ (v)f=0$ or equivalently

$$\begin{array}{l} u_xv_x\ {}_{xx}f+(u_xv_y+u_yv_x)\ {}_{xy}f+u_yv_y\ {}_{yy}f+ \\ (u_x+v_x)\ {}_{xt}f+(u_y+v_y)\ {}_{yt}f+\ {}_{tt}f=0 \end{array} \tag{3}$$

where ${}_{xx}f$ denotes the second order partial derivative in direction $x$, etc. We now define the *mixed motion parameters* [5]

$$\begin{array}{ccc} c_{xx}:=u_xv_x & c_{xy}:=u_xv_y+u_yv_x & c_{yy}:=u_yv_y \\ c_{xt}:=u_x+v_x & c_{yt}:=u_y+v_y & \end{array} \tag{4}$$

and rewrite Eq. 3 as

$$\mathbf{d}^T\mathbf{p}=0 \tag{5}$$

with

$$\begin{array}{l} \mathbf{p}:=[c_{xx},c_{xy},c_{yy},c_{xt},c_{yt},1]^T \\ \mathbf{d}:=[\ {}_{xx}f,\ {}_{xy}f,\ {}_{yy}f,\ {}_{xt}f,\ {}_{yt}f,\ {}_{tt}f]^T. \end{array} \tag{6}$$

We observe that Eq. 5 is a linear model with parameter vector $\mathbf{p}$ and data vector $\mathbf{d}$. Parameters in such a linear model can be estimated by a number of standard parameter-estimation schemes used for optical flow estimation (see e.g. [1, 4] for surveys on such schemes). We implemented a total least-squares scheme, based on the so called extended and generalised structure tensor [3, 5]. We describe this scheme in Sec. 3. From the parameter vector $\mathbf{p}$ we can infer the velocities $\mathbf{u}$ and $\mathbf{v}$ by using the following 'trick' [5]. First, we interpret the vectors $\mathbf{u}$ and $\mathbf{v}$ as complex numbers. Then, with the definitions

$$A_0:=c_{xx}-c_{yy}+\mathbf{i}c_{xy} \qquad A_1:=c_{xt}+\mathbf{i}c_{yt} \tag{7}$$

$\mathbf{u}$ and $\mathbf{v}$ are the solutions of the complex polynomial

$$z^2-A_1z+A_0=0. \tag{8}$$

### 2.2 Transparent Motions plus Additive Source

A locally constant additive source in the image sequence $f$ implies that $f$ is given as

$$f(\mathbf{x},t)=f_1(\mathbf{x}-\mathbf{u}t)+f_2(\mathbf{x}-\mathbf{v}t)+k(t). \tag{9}$$

By using as defined in Eq. 1 we obtain

$$(u)\ (v)f=k''. \tag{10}$$

In analogy to Eq. 6 above, we introduce the vectors

$$\begin{array}{l} \mathbf{p}:=[c_{xx},c_{xy},c_{yy},c_{xt},c_{yt},1,k'']^T \\ \mathbf{d}:=[\ {}_{xx}f,\ {}_{xy}f,\ {}_{yy}f,\ {}_{xt}f,\ {}_{yt}f,\ {}_{tt}f,-1]^T \end{array} \tag{11}$$

This work has partly been funded by the DFG SPP1114 under SCHA 927/1-2 and BA 1176/7-2.

and again obtain a linear model of the form $\mathbf{d}^T\mathbf{p}=0$. Once the parameter vector $\mathbf{p}$ has been determined by using the estimation scheme above, $k''$ is given as its last component. Since we can neither retrieve $k$, nor $k'$, but only the second derivative $k''$, motion in an image sequence with a linear brightness change can be estimated via the model given in Eq. 3.

### 2.3 Transparent Motions plus Exponential Decay

In case of multiplicative brightness changes, the image sequence $f$ is given as

$$f(\mathbf{x},t)=f_1(\mathbf{x}-\mathbf{u}t)k_1(t)+f_2(\mathbf{x}-\mathbf{v}t)k_2(t). \tag{12}$$

If

$$k_1(t) \quad \exp(c_1 t) \qquad k_2(t) \quad \exp(c_2 t) \tag{13}$$

then

$$(u)f_1 = c_1 f_1 \qquad (v)f_2 = c_2 f_2. \tag{14}$$

This is the partial differential equation for an exponentially decaying signal. By defining

$$(\mathbf{v},c) := v_x\ _x + v_y\ _y + \ _t - c \tag{15}$$

we obtain

$$(\mathbf{u},c_1)\ \ (\mathbf{v},c_2)f=0. \tag{16}$$

The above equation is linearised by introducing

$$\begin{aligned}\mathbf{p} &:= [c_{xx},c_{xy},c_{yy},c_{xt},c_{yt},1,-u_xc_2-v_xc_1,\\ &\qquad -u_yc_2-v_yc_1,-c_1-c_2,c_1c_2]^T\\ \mathbf{d} &:= [\ _{xx}f,\ _{xy}f,\ _{yy}f,\ _{xt}f,\ _{yt}f,\ _{tt}f,\ _xf,\ _yf,\ _tf,f]^T.\end{aligned} \tag{17}$$

which leads again to a constraint of the form $\mathbf{d}^T\mathbf{p}=0$. The parameters $c_1$ and $c_2$ are then disentangled as in Sec. 2.1: with

$$A_0 := c_1c_2 \qquad A_1 := -c_1-c_2, \tag{18}$$

$c_1$ and $c_2$ are the solutions of the real polynomial

$$x^2+A_1x+A_0=0. \tag{19}$$

We have determined the two motion vectors and the two decay rates, but still do not know which decay rate belongs to which motion vector. We use the remaining components of $\mathbf{p}$ for the correct assignment. If the assignment is correct, $[u_x,u_y,c_1]$ and $[v_x,v_y,c_2]$ should fulfill the conditions

$$-u_xc_2-v_xc_1-p_7\approx 0 \qquad -u_yc_2-v_yc_1-p_8\approx 0. \tag{20}$$

Otherwise we swap $c_1$ and $c_2$.

### 2.4 Transparent Motions plus Diffusion

To obtain a model for transparent motions plus diffusion, we start with the following partial differential equation

$$(\mathbf{v},c)f=0 \text{ with } \ (\mathbf{v},c)=v_x\ _x+v_y\ _y+\ _t-c\triangle, \tag{21}$$

and $\triangle = \ _{xx}+\ _{yy}$. For two image sequences $f_1$ and $f_2$ that are affected by diffusion, i.e.

$$(\mathbf{u},c_1)f_1=0 \qquad (\mathbf{v},c_2)f_2=0 \tag{22}$$

and $f=f_1+f_2$, we obtain the constraint

$$(\mathbf{u},c_1)\ \ (\mathbf{v},c_2)f=0. \tag{23}$$

Again we can linearise this equation, by using

$$\begin{aligned}\mathbf{p} &:= [c_{xx},c_{xy},c_{yy},c_{xt},c_{yt},1,-u_xc_2-v_xc_1,\\ &\qquad -u_yc_2-v_yc_1,-c_1-c_2,c_1c_2]^T\\ \mathbf{d} &:= [\ _{xx}f,\ _{xy}f,\ _{yy}f,\ _{xt}f,\ _{yt}f,\ _{tt}f,\ _x\triangle f,\ _y\triangle f,\ _t\triangle f,\triangle\triangle f]^T,\end{aligned} \tag{24}$$

to $\mathbf{d}^T\mathbf{p}=0$. The parameter vector is identical to the one from the model for exponential decay (Eq. 17).

## 3. ESTIMATION SCHEME

### 3.1 Structure Tensor

All the constraints derived in Sec. 2 are of the form $\mathbf{d}^T\mathbf{p}=0$ with the data vector $\mathbf{d}$ and the parameter vector $\mathbf{p}$. We have thus one equation per pixel, but need to estimate several parameters (the components of $\mathbf{p}$) per equation, which is an under determined system of equations. In order to solve this system we assume, that within a small spatio-temporal neighborhood of a pixel $i$ all equations are approximately solved by the same set of parameters $\mathbf{p}$. The model $\mathbf{d}^T\mathbf{p}=0$ therefore becomes

$$\mathbf{d}_i^{\mathrm{T}}\mathbf{p}=e_i \text{ for all pixels } i \text{ in} \tag{25}$$

with errors $e_i$ which have to be minimized by the sought for solution $\tilde{\mathbf{p}}$. Using a matrix $\mathbf{D}$ composed of the vectors $\mathbf{d}_i$ via $\mathbf{D}_{ij}=(\mathbf{d}_i)_j$ Eq. 25 becomes $\mathbf{D}\mathbf{p}=\mathbf{e}$. We minimize $\mathbf{e}$ using a weighted 2-norm

$$||\mathbf{e}|| = ||\mathbf{D}\mathbf{p}|| = \mathbf{p}^{\mathrm{T}}\mathbf{D}^{\mathrm{T}}\mathbf{W}\mathbf{D}\mathbf{p} =: \mathbf{p}^{\mathrm{T}}\mathbf{J}\mathbf{p} \stackrel{!}{=} \min \tag{26}$$

where $\mathbf{W}$ is a diagonal matrix containing the weights. In our case Gaussian weights are used, implemented via a 15-tab filter with standard deviation 7. It multiplies each equation $i$ in Eq. 25 by a weight $w_i$. The matrix $\mathbf{J}=\mathbf{D}^{\mathrm{T}}\mathbf{W}\mathbf{D}$ is the so called structure tensor. The error $\mathbf{e}$ is minimized by introducing the assumption $|\tilde{\mathbf{p}}|=1$ in order to suppress the trivial solution $\tilde{\mathbf{p}}=\mathbf{0}$. Doing so the solutions $\tilde{\mathbf{p}}$ is given by the eigenvector to the smallest eigenvalue of $\mathbf{J}$. Moreover, for our models, this vector $\tilde{\mathbf{p}}$ must be normalized with respect to its $6^{th}$ component (cmp. eqs. 6, 11, 17 and 24). For further details on computational issues of TLS estimation, we refer to [3] and the citations therein.

### 3.2 Optimised Filter Sets

We implement all derivatives occurring in Eqs. 6, 11, 17, and 24 as separable spatio-temporal convolution filters. Each filter consists of first or second order derivatives, smoothed in the orthogonal directions. For filter design, we use the method presented in [6]. All of the models developed in Sec. 2.1 are of the form $\mathbf{d}^T\mathbf{p}=0$. Obviously the data vector $\mathbf{d}$ has to be perpendicular to the parameter vector $\mathbf{p}$. The length of $\mathbf{d}$ is of minor interest. Thus we need to design a filter set for each model separately that calculates the *orientation* of $\mathbf{d}$ as good as possible. For the mixed second order derivatives from Eq. 6, we performed a rigorous optimisation following [6]. For 3-tab filters we obtain

$$\begin{array}{ll} I^1=[0.12026,0.75948,0.12026] & D^1=[0.5,0,-0.5]\\ I^2=[0.21478,0.57044,0.21478] & D^2=[1,-2,1]\end{array} \tag{27}$$

and for 5-tab filters

$$\begin{aligned} I^1 &= [0.01504,0.23301,0.50390,0.23301,0.01504]\\ I^2 &= [0.01554,0.23204,0.50484,0.23204,0.01554]\\ D^1 &= [0.06368,0.37263,0,-0.37263,-0.063681]\\ D^2 &= [0.20786,0.16854,-0.75282,0.16854,0.20786]\end{aligned} \tag{28}$$

Second order derivative filters, e.g. $_{xy}$ and $_{xx}$, are then generated via

$$_{xy} = D_x^1 * D_y^1 * I_t^1 \quad \text{and} \quad _{xx} = D_x^2 * I_y^2 * I_t^2 \tag{29}$$

where $*$ denotes convolution and lower indices denote the application direction. All filters not introduced above, can be

derived by exchange of lower indices. The filters are of size $3 \times 3 \times 3$ or $5 \times 5 \times 5$.

For the additive brightness model (Eq. 11), we use the same filters. In the multiplicative brightness model (Eq. 17), first-order derivatives and an identity operator (last component of $\mathbf{d}$) occur. They are build via

$$_x = D_x^1 * I_y^2 * I_t^2 \quad \text{and} \quad I = I_x^2 * I_y^2 * I_t^2 \tag{30}$$

These filters are of size $3 \times 3 \times 3$ or $5 \times 5 \times 5$ also.

The diffusion model is more complex because spatial Laplacians $\triangle$ occur. We use

$$L = D_x^2 * I_y^2 + D_y^2 * I_x^2 \tag{31}$$

as discrete Laplace operator. The $7^{th}$ to $10^{th}$ component of $\mathbf{d}$ from Eq. 24 are then

$$\begin{aligned} _x\triangle &= D_x^1 * I_y^2 * I_t^2 * L & \qquad _y\triangle &= I_x^2 * D_y^1 * I_t^2 * L \\ _t\triangle &= I_x^2 * I_y^2 * D_t^1 * L & \qquad \triangle\triangle &= I_t^2 * L * L. \end{aligned} \tag{32}$$

These are of size $5 \times 5 \times 3$ or $9 \times 9 \times 5$. The inherent smoothing of all filters of a set should be as similar as possible [6]. Thus the $1^{st}$ to $6^{th}$ component of $\mathbf{d}$ from Eq. 24 are not calculated via the second order derivative filters from Eq. 29 directly, but via smoothed versions

$$_{xy} = D_x^1 * D_y^1 * I_t^1 * I_x^2 * I_y^2 \quad \text{and} \quad _{xx} = D_x^2 * I_y^2 * I_t^2 * I_x^2 * I_y^2 \tag{33}$$

Note that additional smoothing is only applied in $x$- and $y$-directions. All filters of the set are of the same spatio-temporal size.

## 4. EXPERIMENTS

All the experimental results have been obtained by using synthetic image sequences with ground truth.

### 4.1 Synthetic Sequences

All image sequences used in the experiments consist of two motion layers. Each layer is generated by moving a given basic pattern or image and a subsequent brightness change. The first motion layer $f_1$ is generated by moving such a pattern with known 'actual' velocity $\mathbf{u}_a = [0,-1]^T$, the second layer $f_2$ via $\mathbf{v}_a = [1,1]^T$. We use integer shifts in order to avoid interpolation errors. From these 2 layers we generate 4 image sequences, one for pure transparent motions, one with additive brightness, one with exponential decay and one with diffusion. The first 3 sequences are defined by Eq. 2, Eq. 9 ($k'' = 8$) and Eq. 12 ($c_1 = -1, c_2 = -0.5$), respectively. The diffusion sequence is generated via successive convolution of each layer by a truncated and normalised Gaussian kernel $G \quad \exp(-x^2/(2 \quad ^2))$ before summing them up. Solving Eq. 21 (see e.g. [2])[1]reveals that we have to choose $= \sqrt{2ct}$. For diffusion we choose $c_1 = 1.0$ and $c_2 = 0.5$. Truncation took place above 6 in order to keep discretization errors low.

For illustration of the motions and brightness changes, we use smoothed delta-combs as basic patterns (cmp. Fig. 1), but for the error analysis smooth noise patterns as depicted in Fig. 2. Three subsequent images of the test sequences illustrating the brightness change models are shown in Fig. 1. The experiments are done with static noise patterns smoothed by a 5-tab binomial filter $[1,4,6,4,1]/16$ applied in $x$- and $y$-direction (cmp. Fig. 2). No time-varying noise has been added to the sequences, and therefore, errors presented in Sec. 4.3 are systematic errors only, coming from the discretization and estimation process.

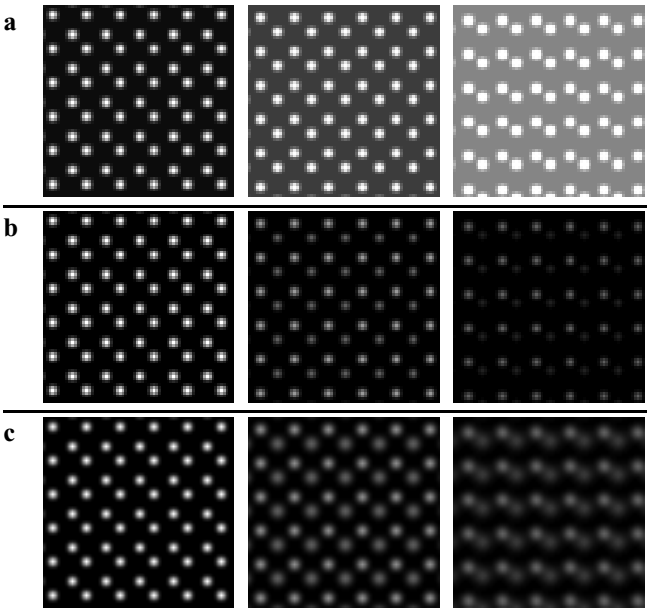


Figure 1: Images 1 to 3 of the 'delta-comb' test sequence: **a** quadratic additive source (Eq. 11), **b** exponential decay (Eq. 17), **c** diffusion (Eq. 24).

### 4.2 Error Measures

For the estimation of optical flow the most popular error measure is the *angular error* $E_\mathbf{v}$ (see [1], eq. 3.38) defined by

$$E_\mathbf{v} = \arccos(\mathbf{r}_a^T \mathbf{r}_e) \tag{34}$$

where the lower index of $E$ indicates, which velocity is used to obtain this error, $\mathbf{r}_a = [v_{x,a}, v_{y,a}, 1]/(v_{x,a}^2 + v_{y,a}^2 + 1)^{1/2}$ is the known ('actual') ground truth spatio-temporal velocity vector of length 1 and $\mathbf{r}_e = [v_{x,e}, v_{y,e}, 1]/(v_{x,e}^2 + v_{y,e}^2 + 1)^{1/2}$ is the estimated velocity vector. The same definition holds for velocities $\mathbf{u}$. For the brightness change parameters $c$, we use the relative deviation $E_c$ between known and estimated $c$'s giving the systematic error. These errors can be minimised by an optimal choice of filter families as demonstrated in [6].

### 4.3 Results

We tested how well the optical flow fields and the brightness parameters are estimated by using a TLS scheme (Sec. 3.1). We used 3 filter sets: central differences ($3 \times 1 \times 1$), 3-tab optimised ($3 \times 3 \times 3$), and 5-tab optimised ($5 \times 5 \times 5$) filters (cmp. Sec. 3.2).[2] In the first test we combined models with suitable test sequences, see Fig. 2. We observe that all errors are rather high for central differences and 3-tab optimised filters.Using 5-tab optimised filters, the errors drop

[1]More exactly: If $_t f = c\triangle f$ and $f = F(\mathbf{x})T(t)$, then $T = \exp(-k^2 ct)$ if $F = \exp(-kx)$. Thus if $F$ is more complicated than this, we have to multiply the Fourier transform of $F$ with $T$. This means, in spatial domain we have to convolve $F$ with the Fourier transform of $T$ which is proportional to $\exp(-x^2/(4ct))$.

[2]Sizes for the diffusion case are ($5 \times 1 \times 1$), ($5 \times 5 \times 3$), and ($9 \times 9 \times 5$).

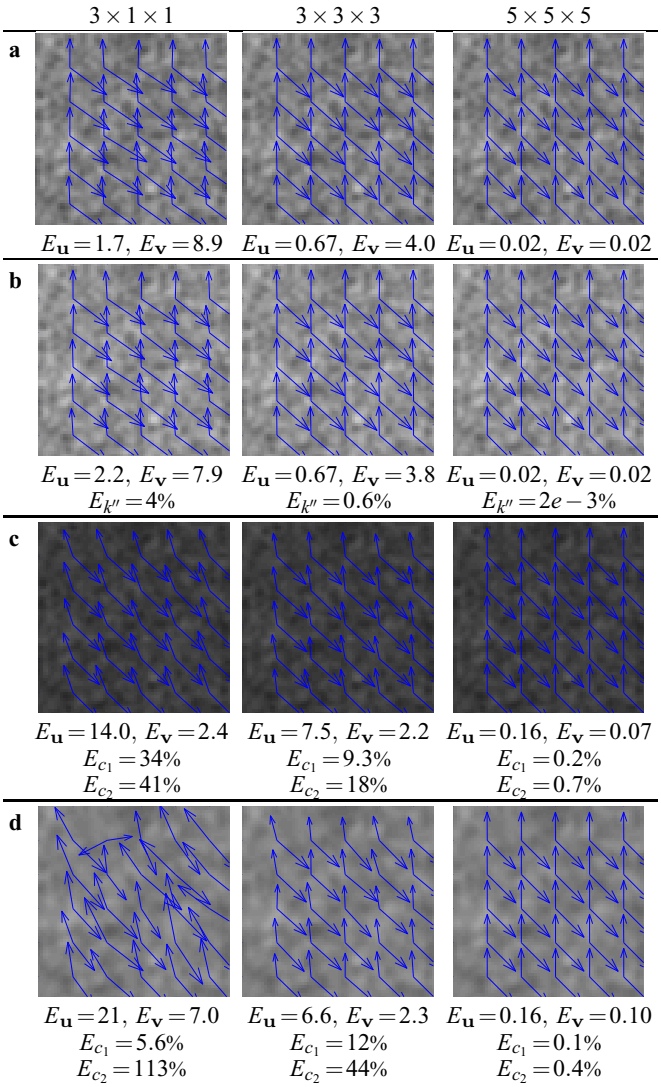


Figure 2: Flow fields using different filter sets, models and suitable test sequences: **a** no brightness change, **b** additive quadratic, **c** multiplicative exponential, and **d** diffusion model. Filter sizes **left**: $3 \times 1 \times 1$, **middle**: $3 \times 3 \times 3$, **right**: $5 \times 5 \times 5$. Motion vectors are scaled by a factor 10. Errors $E_{\mathbf{u}}$ and $E_{\mathbf{v}}$ are angular errors (Eq. 34) in degree, $E_{c_1}$, $E_{c_2}$, and $E_{k''}$ are the absolute values of the relative error of the brightness parameter estimate.

about 1-2 orders of magnitude. Consequently, we use these 5-tab filters in the remainder of this paper. As a second test, we combined the models with non-suitable image sequences. Obviously the brightness parameters estimated do not make any sense then. We neglect them and present the flow fields, only (see Fig. 3). Only the combination multiplicative model with additive brightness change in the data gives reasonable estimates. All other combinations lead to large errors.

## 5. CONCLUSION

We have presented linear models for the estimation of transparent motions with additional, physically motivated, brightness changes. When using the correct model, and 5-tab optimised filter families for the derivatives, the experiments

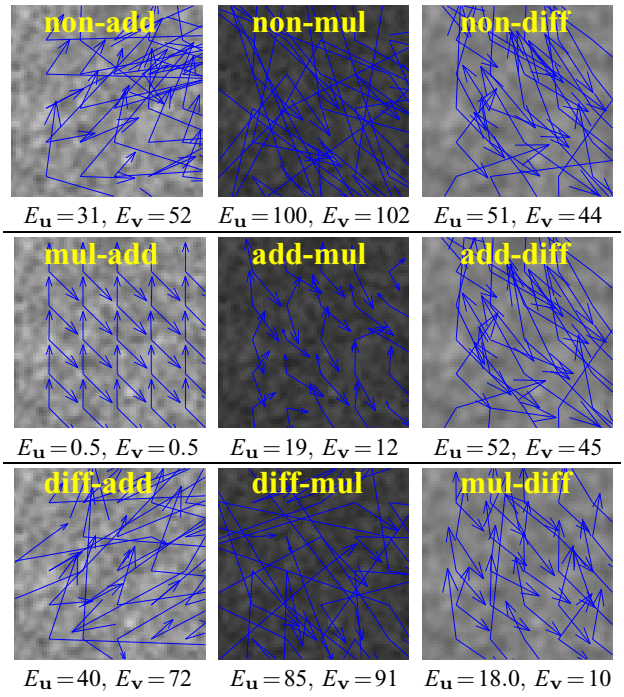


Figure 3: Flow fields obtained by using models that do not match the test sequences. The text on the images indicate the model-sequence combination: **non**: no , **add**: additive, **mul**: multiplicative, and **diff**: diffusive brightness change. Motion vectors are scaled by a factor 10. Errors $E_{\mathbf{u}}$ and $E_{\mathbf{v}}$ are angular errors (Eq. 34) in degree.

yielded low error rates. When using inappropriate motion models, estimates are poor, except when the multiplicative model is used for additive brightness change. We have thus presented a proof of concept for our new methods that deal with the estimation of complex motion patterns.

## References


[1] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[2] Mary L. Boas. *Mathematical Methods in the Physical Sciences*. Wiley, 2 edition, 1999.

[3] H. Haußecker and D. J. Fleet. Computing optical flow with physical models of brightness variation. *PAMI*, 23(6):661–673, June 2001.

[4] H. Haußecker and H. Spies. Motion. In B. Jähne, H. Haußecker, and P. Geißler, editors, *Handbook of Computer Vision and Applications*. Academic Press, 1999.

[5] Cicero Mota, Ingo Stuke, and Erhardt Barth. Analytic solutions for multiple motions. In *Proceedings of the International Conference on Image Processing*, pages 917–920, 2001.

[6] H. Scharr. Optimal filters for extended optical flow. In *International Workshop on Complex Motion, LNCS 3417*, 2005.

[7] M. Shizawa and K. Mase. Simultaneous multiple optical flow estimation. In *ICPR'90*, pages 274–278, 1990.