Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Probabilistic Color Optical Flow

Volker Willert, Julian Eggert, Sebastian Clever, Edgar Körner

2005

Preprint:

This is an accepted article published in Pattern Recognition: 27th DAGM Symposium. The final authenticated version is available online at: https://doi.org/[DOI not available]

Probabilistic Color Optical Flow

Volker Willert[1], Julian Eggert[2], Sebastian Clever[1], and Edgar Körner[2]

[1] TU Darmstadt, Institute of Automatic Control, Control Theory and Robotics Lab,
Landgraf-Georg-Str.04, 64283 Darmstadt,
volker@rtr.tu-darmstadt.de, mail@sebastian-clever.de
[2] HRI Honda Research Institute GmbH,
Carl-Legien-Straße 30, 63073 Offenbach/Main,
{Julian.Eggert, Edgar.Koerner}@honda-ri.de

Abstract. Usually, optical flow computation is based on grayscale images and the brightness conservation assumption. Recently, some authors have investigated in transferring gradient-based grayscale optical flow methods to color images. These color optical flow methods are restricted to brightness and color conservation over time. In this paper, a correlation-based color optical flow method is presented that allows for brightness and color changes within an image sequence. Further on, the correlation results are used for a probabilistic evaluation that combines the velocity information gained from single color frames to a joint velocity estimate including all color frames. The resulting color optical flow is compared to other representative multi-frame color methods and standard single-frame grayscale methods.

1 Introduction

The optical flow is an approximation for a 2D motion field of the velocity vectors of each pixel of an image, with every vector being a projection of the real 3D velocity of a corresponding surface point [4]. In the literature, the term *optical flow* is usually related to image motion fields that are computed purely based on luminance information. To the contrary, *color optical flow* fields are image motion fields that are estimated based on color images which is also often termed multi-channel optical flow, image flow estimation and photometric invariant optical flow etc.

The usual way for motion estimation is using grayscale images and assuming constant brightness over time. The brightness constraint equation is a quite strong assumption which is only appropriate for high frame-rates with small changes between consecutive frames, so that it often does not hold for real world sequences. In fact, the luminance information is highly dependent on moving shadows, varying shading, moving specularities and fluctuations in the light source intensity [9]. To account for the luminance problem, some authors have extended standard optical flow estimation algorithms for the use of color images instead of grayscale ones. To do so, they replace the brightness assumption with a less restrictive constant chromaticity assumption, also called color invariance assumption, meaning that the color stays constant over time [6],[9]. Obviously, color images contain more scene information than grayscale images and therefore an improvement should be expected for color-based optical flow estimation. Nevertheless, the constant color assumption seems to be nearly as restrictive as the

W. Kropatsch, R. Sablatnig, and A. Hanbury (Eds.): DAGM 2005, LNCS 3663, pp. 9–16, 2005.
© Springer-Verlag Berlin Heidelberg 2005

constant brightness assumption, since the problem of varying information over time is still present.

A simple experimental examination shown in Fig. 1 indicates that also color information varies to a large extent within consecutive images of a real world sequence. A person holding a yellow book in front of his body is moving towards the camera under a fluorescent tube illumination. Measuring the RGB/HSV values of two consecutive images $\mathbf{I}^t$ and $\mathbf{I}^{t+\Delta t}$ shows that the color information in the different color channels within the black circle is 1) not constant across different pixels within the circle, mainly because of chroma noise, and 2) also not constant for corresponding pixels within a color frame over time, mainly because of luminance and reflection changes. The variances for the different color channels within the circle and over the two consecutive images are listed in the table of Fig. 1.

$\mathbf{I}^t$  $\mathbf{I}^{t+\Delta t}$ 

	$\mathbf{I}^t$	$\mathbf{I}^t \rightarrow \mathbf{I}^{t+\Delta t}$
RGB	$\sigma_r = 13$ $\sigma_g = 14$ $\sigma_b = 21$	$\sigma_r = 12$ $\sigma_g = 7$ $\sigma_b = 3$
HSV	$\sigma_h = 5$ $\sigma_s = 22$ $\sigma_v = 6$	$\sigma_h = 8$ $\sigma_s = 13$ $\sigma_v = 4$

Fig. 1. Example showing the change of color over time because of independent changes of brightness and contrast within different color channels. The color components of an image pixel can change over time because of reflection and illumination changes.

Besides varying color information, it is not clear how existing correlations between the different color channels, mentioned by Madjidi and Negahdaripour in [7], can be used for optical flow computation and how to handle brightness and color assumptions concurrently. In spite of the many open questions regarding color optical flow, researches in that field are still quite sparse [6],[9],[1],[7],[3].

Looking at standard applications, the optical flow is usually gained by comparing brightness patterns of two consecutive images $\mathbf{I}^t$ and $\mathbf{I}^{t+\Delta t}$, where $\mathbf{I}^t$ is an image consisting of pixels at locations $\mathbf{x}$ at time t. Comparing brightness patterns often leads to assume brightness invariance of particular patterns under motion. Let $\hat{\mathbf{W}} \odot \mathbf{G}^{t,\mathbf{p}}$ be a weighted patch of gray values of image $\mathbf{I}^t$ centered and windowed about $\mathbf{p}$ (with $\odot$ symbolising the componentwise multiplication of two vectors). $\mathbf{G}^t$ denotes the vector of intensities linked to the particular image and $\hat{\mathbf{W}}$ defines a window function, which restricts the pattern size. Assuming brightness conservation the standard correlation-based optical flow equation can be formulated following [5] as:

$$\hat{\mathbf{W}} \odot \mathbf{G}^{t,\mathbf{p}} = \hat{\mathbf{W}} \odot \mathbf{G}^{t+\Delta t,\mathbf{p}+\Delta \mathbf{p}} \quad . \tag{1}$$

In order to obtain color optical flow Golland and Bruckstein [6] used a standard gradient-based approach with the same assumptions as in Eq. 1 and applied the resulting brightness constraint equation to each channel of the RGB color space. Then

they solved the resulting overdetermined system of linear equations by a standard least-squares algorithm. In an alternative approach they propose that in pure color channels changes in brightness do not appear and the resulting color conservation assumption would be more appropriate. In this approach, they eliminate the brightness containing channel, e.g. the Value channel of the HSV color space, and applied the two pure color channels, e.g. the Hue and Saturation channel, to the mentioned constraint.

Starting from the dichromatic reflection model, van de Weijer and Gevers [9] derive photometric invariants to improve optical flow estimation that in their approach is independent of shadow-shading and specular reflectance changes. They propose the combination of a reliability measure that considers instabilities of the photometric invariants and optical flow estimation to increase robustness.

Based on Gollands and Bruckstein's first proposal, Barron and Klette [2],[3] analysed different standard differential techniques for computing the optical flow, e.g. the Lucas and Kanade method or the Horn and Schunck regularization. For more details of standard methods see [4]. Furthermore, they recast the Horn and Schunck regularization adding a directional constraint that depends on one knowing whether the camera is panning or zooming in the standard minimization formula.

Andrews and Lovell [1] developed some faster algorithms for solving the color optical flow equations proposed by Golland, Bruckstein and Barron, Klette.

In this paper, we investigate the formulation of a local linear generative model that approximately describes the correlation between colored image patches within two consecutive multi-frame color images. This model is used to generate locally as well as channel independent measurements that are interpreted as discrete conditional probability density functions (pdfs) holding the probabilities for several motion hypotheses given two consecutive color patches. The advantage of these pdfs given a set of discrete velocities is that a number of velocity hypotheses can be tested concurrently. This means that the velocity information derived from the single color channels can be combined without loosing information. From these velocity distributions velocity vectors are extracted to estimate the optical flow of a color image sequence. The main assumptions are 1) that due to illumination and reflection changes, the color as well as the brightness and the contrast within an image patch can vary systematically over time, especially when there are moving objects in a scene and 2) the information in the color channels can be treated as statistically independent and so can be combined to one joint velocity distribution for each image location.

In Sec. 2 a short introduction to color spaces and a motivation why we treat the color channels as statistically independent is given. In Sec. 3 a model that allows for local value changes over time within the color channels is proposed and a correlation-based probabilistic interpretation is presented that leads to a contrast and brightness invariant color optical flow estimation. Finally some quantitative results are given in Sec. 4 followed by some short conclusion in Sec. 5.

2 Color Spaces

The representation of color in so called color spaces like the RGB, HSV or YCbCr follows the trichromatic theory of color, whereby every color can be specified by an

additive composition of appropriate features like, e.g. the quota of red, green and blue in a particular color. Hence every color can be represented by a linear combination of those features each presenting a basis vector, since they are linearly independent and their span describes the whole space.

$$\mathbf{c}_i = x_i\mathbf{e}_x + y_i\mathbf{e}_y + z_i\mathbf{e}_z\,, \quad \text{with} \;\; 0 \leq x_i, y_i, z_i \leq 1 \;\; \text{and} \;\; i \in \{\text{RGB,HSV,YCbCr}\} \quad (2)$$

The vector $\mathbf{c}_i$ represents a particular color in space i, x_i, y_i, z_i are the corresponding values of the associated features and $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ are the standard basis vectors. Thus, the whole color picture is described by

$$\mathbf{C}_i = (\mathbf{X}_i; \mathbf{Y}_i; \mathbf{Z}_i)^T\,, \quad (3)$$

where $\mathbf{X}_i$, $\mathbf{Y}_i$, $\mathbf{Z}_i$ are the single color channels of the frame.
An example of a typical three channel HSV color frame is given in Fig. 2. Remarkable are the large differences according to contrastive and homogenous areas between the channels. Every color space is spanned by a 3D coordinate system which describes the color gamut of a particular device [8].

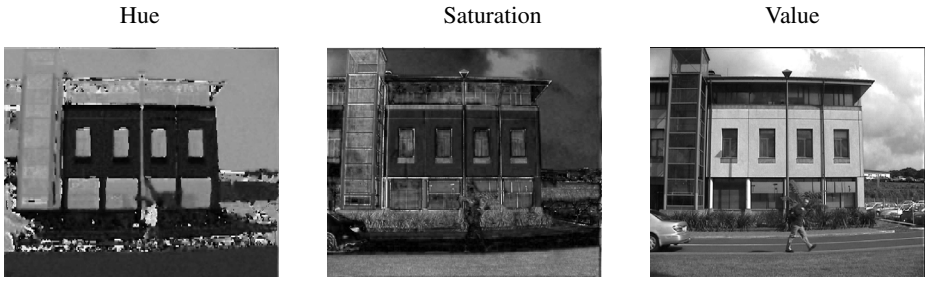

Fig. 2. Example of a three channel HSV color frame with remarkable differences between the channels

The spectral responsivity of the used sensors for measurement, e.g. of the red, green and blue component, will overlap in most cases. Thus the measured components will be correlated. Nevertheless we assume that the measured data of the different channels are statistical independent, since 1) the spectral overlapping appears only in subareas of the spectral responses and 2) the measured data is mostly preprocessed with device specific parameters that differs between the different sensor types meaning the dependence between the channels is unknown. The assumption of statistical independent spectral responses seems to be a good approximation, since in most cases knowing the spectral response of one channel does not allow to draw any conclusions regarding the responses of the other channels. Even if the value of one pixel in a particular channel is measured, this does not constrain the set of possible values of another channel.

3 Correlation-Based Color Optical Flow

Our approach for calculating color optical flow bases on a formulation for computing the optical flow of intensity images [5]. The used notation is illustrated and described in

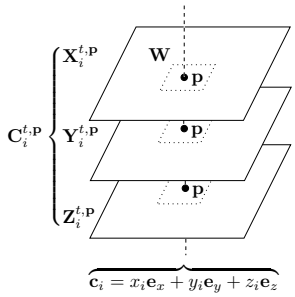

Fig. 3. This figure shows the used notation. A color picture $\mathbf{C}_i^{t,\mathbf{p}}$ consists of three color channels $\mathbf{X}_i^{t,\mathbf{p}}$, $\mathbf{Y}_i^{t,\mathbf{p}}$, $\mathbf{Z}_i^{t,\mathbf{p}}$ with each channel containing a component, e.g. x_i, y_i or z_i (See Eq. 2), for each pixel $\mathbf{p}$. A particular weighted 3D-patch of an image at position $\mathbf{p}$ at time t is achieved with the $\odot$ operation of the window vector $\mathbf{W} := (\hat{\mathbf{W}}; \hat{\mathbf{W}}; \hat{\mathbf{W}})^T$ on the vector $\mathbf{C}_i^{t,\mathbf{p}}$ in the sense $\mathbf{W}\odot\mathbf{C}_i^{t,\mathbf{p}} = (\mathbf{W}\odot\mathbf{X}_i^{t,\mathbf{p}}; \mathbf{W}\odot\mathbf{Y}_i^{t,\mathbf{p}}; \mathbf{W}\odot\mathbf{Z}_i^{t,\mathbf{p}})^T$, denoting elementwise multiplication.

Fig. 3. Here, extending the brightness constancy assumption of Eq. 1, it is assumed that during motion the contributions of the single color channels are independently jittered by noise η and that brightness and contrast variations may occur over time, allowing for systematic brightness, contrast and in addition color changes. The assumed color changes imply e.g. changes of brightness or changes of the spectrum of the illumination. These color changes are accounted for by a scaling vector $\boldsymbol{\lambda}_i$ and a bias vector $\boldsymbol{\kappa}_i$. For the sake of simplicity, we now neglect index i for the different color spaces and arrive in analogy to Eq. 1 at

$$\mathbf{W}\odot\mathbf{C}^{t,\mathbf{p}} = \mathbf{W}\odot\left[\boldsymbol{\lambda}\odot\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}} + \boldsymbol{\kappa}\right] + \boldsymbol{\eta} \quad \text{, with} \tag{4}$$
$$\boldsymbol{\lambda} := (\lambda_x\mathbf{1}; \lambda_y\mathbf{1}; \lambda_z\mathbf{1})^T\,, \quad \boldsymbol{\kappa} := (\kappa_x\mathbf{1}; \kappa_y\mathbf{1}; \kappa_z\mathbf{1})^T \quad \text{and} \quad \boldsymbol{\eta} := (\eta_x\mathbf{1}; \eta_y\mathbf{1}; \eta_z\mathbf{1})^T$$

In this expression the patch $\mathbf{W}\odot\mathbf{C}^{t,\mathbf{p}}$ contains in accordance to Eq. 3 for each pixel a vector.

Assuming that the image noise of each channel is zero mean Gaussian and statistically independent from the noise in the other channels, with variances $\sigma_{\eta x}$, $\sigma_{\eta y}$ and $\sigma_{\eta z}$, we get a covariance matrix $\boldsymbol{\Sigma_\eta}$, where only the elements of the leading diagonal are different from zero. Thus, the likelihood that $\mathbf{C}^{t,\mathbf{p}}$ is a match for $\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}$, given a velocity $\mathbf{v} = \Delta\mathbf{p}/\Delta t$ and the parameter vectors $\boldsymbol{\lambda}, \boldsymbol{\kappa}$, the window function $\mathbf{W}$ and the covariance matrix $\boldsymbol{\Sigma_\eta}$, can be written down as:

$$\rho\left(\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}, \mathbf{C}^{t,\mathbf{p}} \mid \mathbf{v}\right) \sim \tag{5}$$
$$\sim e^{-\frac{1}{2}\left(\mathbf{W}\odot\left(\boldsymbol{\lambda}\odot\mathbf{C}^{t,\mathbf{p}}+\boldsymbol{\kappa}\,-\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}\right)\right)\boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{-1}\left(\mathbf{W}\odot\left(\boldsymbol{\lambda}\odot\mathbf{C}^{t,\mathbf{p}}+\boldsymbol{\kappa}\,-\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}\right)\right)^T}$$

We now proceed to make Eq. 5 independent on $\boldsymbol{\lambda}$ and $\boldsymbol{\kappa}$. Thus we maximize Eq. 5 with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{\kappa}$. This leads to a minimization of the exponent of Eq. 5:

$$\{\boldsymbol{\lambda}^*, \boldsymbol{\kappa}^*\} := \min_{\boldsymbol{\lambda},\boldsymbol{\kappa}} F, \quad \text{with} \tag{6}$$
$$F := \left(\mathbf{W}\odot\left(\boldsymbol{\lambda}\odot\mathbf{C}^{t,\mathbf{p}}+\boldsymbol{\kappa}\,-\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}\right)\right)\boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{-1}\left(\mathbf{W}\odot\left(\boldsymbol{\lambda}\odot\mathbf{C}^{t,\mathbf{p}}+\boldsymbol{\kappa}\,-\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}\right)\right)^T.$$

This amounts to solve the homogeneous equation system which we get from setting the derivatives of Eq. 6 in direction of $(\boldsymbol{\lambda}, \boldsymbol{\kappa})$ to zero:

$$\frac{d}{d(\boldsymbol{\lambda}, \boldsymbol{\kappa})}F = \mathbf{0} \quad , \quad \frac{d^2}{d(\boldsymbol{\lambda}, \boldsymbol{\kappa})^2}F > \mathbf{0} \quad . \tag{7}$$

In the case that each color channel has an independent λ, this minimization leads to three independent equation systems each containing two constraints with two unknowns, which can be solved separately.

Let $\boldsymbol{\lambda}^* = \left(\lambda_x^*, \lambda_y^*, \lambda_z^*\right)^T$, $\boldsymbol{\kappa}^* = \left(\kappa_x^*, \kappa_y^*, \kappa_z^*\right)^T$ be the parameters which minimize Eq. 6. Substituting $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$ and $\boldsymbol{\kappa} = \boldsymbol{\kappa}^*$ into Eq. 5, the final likelihood reads:

$$\rho\left(\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}, \mathbf{C}^{t,\mathbf{p}} \mid \mathbf{v}\right) \sim e^{-\frac{1}{2}\sum_{\mathbf{A}=\mathbf{X},\mathbf{Y},\mathbf{Z}}\left(\frac{\sigma_{\mathbf{A}^{t,\mathbf{p}}}}{\sigma_{\eta\mathbf{A}}}\right)^2\left(1-\varrho^2_{\mathbf{A}^{t,\mathbf{p}},\mathbf{A}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}}\right)} \quad , \tag{8}$$

with $\sigma_{\mathbf{A}^{t,\mathbf{p}}}$ being the variance of the weighted patch of the momentary channel and $\varrho_{\mathbf{A}^{t,\mathbf{p}},\mathbf{A}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}}$ being the correlation coefficient between two patches of two consecutive channels over time. From here on, for the sake of brevity, we now write $\rho^t(\mathbf{p}|\mathbf{v}) := \rho\left(\mathbf{C}^{t+\Delta t,\mathbf{p}+\Delta\mathbf{p}}, \mathbf{C}^{t,\mathbf{p}} \mid \mathbf{v}\right)$, which expresses the joint likelihood of the image data at location $\mathbf{p}$ at time t given discrete motion hypotheses $\mathbf{v}$ for all color channels. We see that $\rho^t(\mathbf{p}|\mathbf{v})$ factorizes, so that the corresponding probability for the whole image can be written as

$$\rho^t(\mathbf{p}|\mathbf{v}) = \rho_x^t(\mathbf{p}|\mathbf{v}) \cdot \rho_y^t(\mathbf{p}|\mathbf{v}) \cdot \rho_z^t(\mathbf{p}|\mathbf{v}) \quad , \tag{9}$$

meaning that they can be calculated separately, with $\rho_x^t(\mathbf{p}|\mathbf{v}), \rho_y^t(\mathbf{p}|\mathbf{v}), \rho_z^t(\mathbf{p}|\mathbf{v})$ being the likelihoods of each channel. The final optical flow field can directly be estimated from the joint likelihood $\rho^t\left(\mathbf{p}|\mathbf{v}\right)$ and a given prior $\rho\left(\mathbf{v}\right)$ using Bayes' rule and the maximum a posteriori estimator:

$$\rho^t\left(\mathbf{v}|\mathbf{p}\right) = \rho\left(\mathbf{v}\right)\rho^t\left(\mathbf{p}|\mathbf{v}\right), \quad \text{and} \quad \mathbf{v} = \max_{\mathbf{v}}\left(\rho^t\left(\mathbf{v}|\mathbf{p}\right)\right) \quad . \tag{10}$$

In the following experiments the prior was chosen to be equally distributed.

4 Results

To give a quantitative analysis and a comparison to other existing color optical flow methods we used two pan and zoom synthetic image sequences generated by John Barron and Reinhard Klette and added our results to the results presented in [3]. The quantitative error measurements can be seen in Table 1 and the corresponding optical flows

Fig. 4. Panning and Zooming color optical flow for the RGB color image

Table 1. Error measurements reported in [3] including our results for comparison

Panning		$e_M \pm \sigma_{e_M}$	$e_\varphi \pm \sigma_{e_\varphi}$	Zooming		$e_M \pm \sigma_{e_M}$	$e_\varphi \pm \sigma_{e_\varphi}$
Horn-Schunk	RGB	$24.27\% \pm 23.62\%$	$4.56^\circ \pm 6.19^\circ$	Horn-Schunk	RGB	$13.72\% \pm 14.92\%$	$6.54^\circ \pm 7.91^\circ$
Lucas-Kanade	RGB	$4.90\% \pm 9.31\%$	$1.00^\circ \pm 3.44^\circ$	Lucas-Kanade	RGB	$8.17\% \pm 12.04\%$	$4.01^\circ \pm 12.13^\circ$
Barron-Klette	RGB	$26.21\% \pm 19.29\%$	0.0°	Barron-Klette	RGB	$15.48\% \pm 23.89\%$	0.0°
Golland-Bruckstein	RGB	$11.38\% \pm 17.36\%$	$5.04^\circ \pm 11.80^\circ$	Golland-Bruckstein	RGB	$14.74\% \pm 19.37\%$	$7.56^\circ \pm 14.87^\circ$
Otha	RGB	$12.38\% \pm 18.91\%$	$6.04^\circ \pm 12.80^\circ$	Otha	RGB	$18.69\% \pm 27.87\%$	$9.78^\circ \pm 15.80^\circ$
Willert-Eggert-Clever	RGB	0.0%	0.0°	Willert-Eggert-Clever	RGB	$22.17\% \pm 18.77\%$	$6.34^\circ \pm 6.33^\circ$
Willert-Eggert-Clever	HSV	$1.97\% \pm 13.9\%$	$0.001^\circ \pm 0.41^\circ$	Willert-Eggert-Clever	HSV	$26.20\% \pm 21.39\%$	$7.79^\circ \pm 8.61^\circ$
Willert-Eggert-Clever	YCbCr	0.0%	0.0°	Willert-Eggert-Clever	YCbCr	$22.43\% \pm 18.91\%$	$6.52^\circ \pm 6.48^\circ$
Horn-Schunk	S	$43.69\% \pm 28.94\%$	$10.48^\circ \pm 11.61^\circ$	Horn-Schunk	S	$22.83\% \pm 11.23\%$	$11.23^\circ \pm 12.81^\circ$
Lucas-Kanade	S	$6.54\% \pm 13.19\%$	$1.39^\circ \pm 4.54^\circ$	Lucas-Kanade	S	$10.21\% \pm 14.66\%$	$5.32^\circ \pm 13.62^\circ$
Barron-Klette	S	$32.59\% \pm 20.19\%$	0.0°	Barron-Klette	S	$26.43\% \pm 39.99\%$	0.0°
Golland-Bruckstein	S	$16.97\% \pm 22.41\%$	$9.34^\circ \pm 20.80^\circ$	Golland-Bruckstein	S	$20.08\% \pm 24.05\%$	$12.99^\circ \pm 23.98^\circ$
Willert-Eggert-Clever	S	$10.77\% \pm 31.00\%$	$0.002^\circ \pm 0.70^\circ$	Willert-Eggert-Clever	S	$37.69\% \pm 24.68\%$	$16.55^\circ \pm 21.10^\circ$
Horn-Schunk	Y	$22.90\% \pm 24.12\%$	$4.73^\circ \pm 7.28^\circ$	Horn-Schunk	Y	$21.47\% \pm 20.50\%$	$10.35^\circ \pm 11.73^\circ$
Lucas-Kanade	Y	$6.14\% \pm 11.95\%$	$1.63^\circ \pm 5.23^\circ$	Lucas-Kanade	Y	$9.46\% \pm 14.04\%$	$5.05^\circ \pm 13.61^\circ$
Barron-Klette	Y	$20.06\% \pm 20.67\%$	0.0°	Barron-Klette	Y	$23.67\% \pm 21.18\%$	0.0°
Golland-Bruckstein	Y	$21.08\% \pm 28.07\%$	$12.59^\circ \pm 24.99^\circ$	Golland-Bruckstein	Y	$18.91\% \pm 23.72\%$	$11.46^\circ \pm 21.84^\circ$
Willert-Eggert-Clever	Y	$0.005\% \pm 0.51\%$	$0.007^\circ \pm 0.89^\circ$	Willert-Eggert-Clever	Y	$23.01\% \pm 18.32\%$	$6.73^\circ \pm 6.72^\circ$

Fig. 5. Color optical flow for real-world sequences for a YCbCr and a RGB color sequence

are printed in Fig. 4. The same quantitative error measurements reported in [3], that is, the relative magnitude errors $e_M = 1/ij \sum_{ij} ||\mathbf{v}_{ij}^c - \mathbf{v}_{ij}^e||^2 / \sum_{ij} ||\mathbf{v}_{ij}^c||^2 \times 100\%$ and the angle errors $e\varphi = 1/ij \sum_{ij} \arccos(\mathbf{v}_{ij}^c \cdot \mathbf{v}_{ij}^e)$ with their corresponding variances σ_{e_M} and σ_{e_φ}, are used. Although the test sequences are synthetic sequences with no illumination and reflection changes at all, which means the extension to a channelwise contrast and brightness invariant measurement is not necessary, our method compares quite favourably to the others. Especially for the panning sequence our method outperforms the existing ones. Since we use the maximum a posteriori estimator we cannot reach subpixel accuracy and our results for the zooming sequence are not that convincing. To give a first impression on how our method works on real world sequences we

added two further example sequences shown in Fig. 5. The qualitative optical flow results for an YCbCr (left) color and a RGB (right) color sequence processing only two consecutive images of the sequence can be seen.

5 Conclusion

We have presented a probabilistic correlation-based color optical flow algorithm that allows for brighness, contrast and color changes over time. In the case of pixel accuracy it compares quite favourably to existing gradient-based approaches that assume brightness, contrast and color to be constant over time. Our generative model for color image formation can be extended straightforwardly to also allow for correlations between the color channels by adding crossterms to the scaling and bias parameters. To study the usefulness of such crossterms for optical flow estimation will be the topic of further research.

Acknowledgements. Special thanks to Prof. John L. Barron from the University of Western Ontario, Canada, and his college Prof. Dr. Reinhard Klette from the University of Auckland, New Zealand, who provided us with their test sequences.

References

1. J. Andrews and B.C. Lovell. Color optical flow. *Eds. Workshop on Digital Image Computing, Brisbane, Australia*, 1(1):135–139, 2003.
2. J. Barron and R. Klette. Experience with optical flow in colour video image sequences. *Image and Vision Computing'2001, Auckland University, New Zealand*, pages 195–200, 2001.
3. J. Barron and R. Klette. Quantitative color optical flow. *International Conference on Pattern Recognition, Vancouver, Canada*, pages 251–255, 2002.
4. S.S. Beauchemin and J.L. Barron. The computation of optical flow. *ACM Computing Surveys*, 27(3):433–467, 1995.
5. J. Eggert, V. Willert, and E. Körner. *LNCS 3175*, chapter Building a Motion Resolution Pyramid by Combining Velocity Distributions, pages 310–317. Springer, Berlin, Germany, 2004.
6. P. Golland and A.M. Bruckstein. Motion from color. *Computer Vision and Image Understanding*, 68(3):346–362, 1997.
7. H. Madjidi and S. Negahdaripour. On robustness and localization accuracy of optical flow computation from color imagery. *2nd International Symposium on 3D Data Processing, Visualization, and Transmission, Thessaloniki, Greece*, pages 317–324, 2004.
8. S. Süsstrunk, R. Buckley, and S. Swen. Standard rgb color spaces. In *Color Imaging Conference*, pages 127–134. IS&T - The Society for Imaging Science and Technology, 1999.
9. J. van de Weijer and Th. Gevers. Robust optical flow from photometric invariants. *IEEE International Conference on Image Processing, Singapore*, pages 251–255, 2004.