

Complementary Optic Flow

Henning Zimmer^{1,4}, Andrés Bruhn¹, Joachim Weickert¹, Levi Valgaerts¹,
Agustín Salgado², Bodo Rosenhahn³, and Hans-Peter Seidel⁴

¹ Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University, Saarbrücken, Germany
{zimmer,bruhn,weickert,valgaerts}@mia.uni-saarland.de

² Departamento de Informática y Sistemas
Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain
asalgado@dis.ulpgc.es

³ Institut für Informationsverarbeitung, University of Hannover
Hannover, Germany
rosenhahn@tnt.uni-hannover.de

⁴ Max-Planck Institute for Informatics, Saarbrücken, Germany
hpseidel@mpi-sb.mpg.de

Abstract. We introduce the concept of complementarity between data and smoothness term in modern variational optic flow methods. First we design a sophisticated data term that incorporates HSV colour representation with higher order constancy assumptions, completely separate robust penalisation, and constraint normalisation. Our anisotropic smoothness term reduces smoothing in the data constraint direction instead of the image edge direction, while enforcing a strong filling-in effect orthogonal to it. This allows optimal complementarity between both terms and avoids undesirable interference. The high quality of our complementary optic flow (COF) approach is demonstrated by the current top ranking result at the Middlebury benchmark.

1 Introduction

In spite of the fact that variational optic flow methods are around for almost three decades and that they mark the state-of-the-art in terms of accuracy, there has been remarkably little research on the compatibility of their two ingredients: the data term and the smoothness term. The data term models constancy assumptions on certain image properties, e.g. grey value constancy in the seminal Horn and Schunck model [1]. The smoothness term penalises fluctuations in the flow field. However, these terms may contradict each other: While the brightness constancy assumption constrains the flow only along the image gradient but not across it (aperture problem), most smoothness terms enforce their constraints also along the image gradient. One notable exception is the Nagel-Enkelmann model [2] where the homogeneous Horn and Schunck smoothness term is replaced by an anisotropic one. For large image gradients the latter one

works solely orthogonal to the image gradient. Thus, both terms complement each other without undesirable interference. The fact that the Nagel-Enkelmann model outperforms the Horn and Schunck approach demonstrates the high potential of such a complementarity.

Unfortunately, this paradigm of complementary behaviour has not been explored further after 1986. Instead of this, research has focussed on improving the data or smoothness constraints independently. The goal of our paper is to propose a synergistic model for variational optic flow computation that integrates state-of-the-art data and smoothness assumptions in such a way that both terms work complementary. We will see that this can still lead to a very substantial gain in accuracy. This is demonstrated by the fact that our so-called *complementary optic flow (COF)* method ranks number one in the widely-used Middlebury benchmark.

Our paper is organised as follows: In Sec. 2 we review variational optic flow. Our data term is derived in Sec. 3 and is then used to complement the smoothness term in Sec. 4. After discussing implementation issues in Sec. 5, we show experiments proving the favourable performance of our method in Sec. 6. We then conclude with a summary and an outlook on future work in Sec. 7.

Related Work. Our model naturally incorporates many concepts that have demonstrated their usefulness over the years. Therefore let us briefly sketch the advances in data and smoothness terms that have been most influential for us.

For the data term, Black and Anandan [3] replaced the quadratic penalisation from [1] by a robust one which helps to cope with outliers caused by noise or occlusions. More recently, in order to make the data term robust under additive illumination changes, Brox et al. [4] successfully combined the classical *brightness constancy assumption (BCA)* with the *gradient constancy assumption (GCA)* [4–6]. Bruhn and Weickert [7] later improved this idea by introducing a separate robust penalisation of the BCA and the GCA. This gives advantages in those situations where one of the two constraints produces an outlier. Moreover, in realistic scenarios, one also has to deal with multiplicative illumination changes [8]. If colour image sequences are available, one solution to this issue can be the use of alternative colour spaces with photometric invariances, see [9] and the references therein. Besides the discussed robustification efforts, successful modifications of the data term have been reported by normalising the data term [10, 11]. It prevents an undesirable overweighting of the data term at large image gradient locations.

Regarding the smoothness term, first ideas go back to Horn and Schunck [1] who used a *homogeneous* regulariser that does not respect any flow discontinuities. However, since different image objects may move in different directions, it is desirable to also permit discontinuities. This can for example be achieved by using *image-driven* regularisers that take into account image discontinuities. Alvarez et al. [12] proposed an isotropic model with a scalar-valued weight function that reduces the regularisation at image edges. An anisotropic counterpart that also exploits the directional information of image discontinuities was introduced by Nagel and Enkelmann [2]. Their method regularises the flow field

along image edges but not across them. However, as not every image edge coincides with a flow edge, image-driven methods are prone to oversegmentation artifacts in textured image regions. To avoid this, *flow-driven* regularisers have been proposed that respect discontinuities of the evolving flow and are thus not misled by image textures. In the isotropic setting this comes down to the use of robust, nonquadratic penalisers, which for discrete energy functions have been proposed by Black and Anandan [3]. In the context of rotationally invariant variational methods they go back to Schnörr [5], and Weickert and Schnörr [13] later presented an anisotropic extension. Nevertheless, the problem of flow-driven regularisers lies in less sharp and badly localised flow edges compared to image-driven approaches. The recent discrete method of Sun et al. [14] incorporates an anisotropic regulariser based on a Steerable Random Field [15] that uses directional flow derivatives steered by image structures. It can thus be classified as *joint image- and flow-driven (JIF)*, allowing to obtain sharp flow edges without oversegmentation problems.

2 Variational Optic Flow

Let $f(\mathbf{x})$ be an image sequence with $\mathbf{x} := (x, y, t)^\top$, where $(x, y)^\top \in \Omega$ denotes the location within a rectangular image domain $\Omega \subset \mathbb{R}^2$ and $t \geq 0$ denotes time. We further assume that f is presmoothed by a Gaussian convolution of standard deviation σ . The optic flow field $\mathbf{w} := (u, v, 1)^\top$ describes the displacement vector field between two frames at time t and $t + 1$. It is found by minimising a global energy functional of the general form

$$E(u, v) = \int_{\Omega} [M(u, v) + \alpha V(\nabla_2 u, \nabla_2 v)] \, dx \, dy \, , \quad (1)$$

where $\nabla_2 := (\partial_x, \partial_y)^\top$ denotes the spatial gradient operator. The term $M(u, v)$ denotes the data term, $V(\nabla_2 u, \nabla_2 v)$ the smoothness term, and $\alpha > 0$ is a smoothness weight. According to the calculus of variations, a minimiser (u, v) of the energy (1) necessarily has to fulfil the associated Euler-Lagrange equations

$$\partial_u M - \alpha (\partial_x (\partial_{u_x} V) + \partial_y (\partial_{u_y} V)) = 0 \, , \quad (2)$$

$$\partial_v M - \alpha (\partial_x (\partial_{v_x} V) + \partial_y (\partial_{v_y} V)) = 0 \, , \quad (3)$$

with homogeneous Neumann boundary conditions.

3 Data Term

Let us now derive our data term in a systematic way. The classical starting point is the *brightness constancy assumption (BCA)* used by Horn and Schunck [1]. It states that image intensities remain constant under their displacement, i.e., $f(\mathbf{x} + \mathbf{w}) = f(\mathbf{x})$. Assuming that the displacement is sufficiently small, we can

perform a first-order Taylor linearisation that yields the *optic flow constraint (OFC)*

$$0 = f_x u + f_y v + f_t = \nabla_3 f^\top \mathbf{w} , \quad (4)$$

where the subscripts denote partial derivatives and $\nabla_3 := (\partial_x, \partial_y, \partial_t)^\top$ denotes the spatio-temporal gradient operator. For a quadratic penalisation the corresponding data term is given by

$$M_1(u, v) = (\nabla_3 f^\top \mathbf{w})^2 = \mathbf{w}^\top J_0 \mathbf{w} , \quad (5)$$

with the tensor $J_0 := \nabla_3 f \nabla_3 f^\top$.

The OFC is not sufficient to compute a unique solution (aperture problem), but only allows to compute the flow component orthogonal to the image edges, the so-called *normal flow*. It is defined as

$$\mathbf{w}_n := (\mathbf{u}_n^\top, 1)^\top := \left(-\frac{f_t}{|\nabla_2 f|} \frac{\nabla_2 f^\top}{|\nabla_2 f|}, 1 \right)^\top . \quad (6)$$

Normalisation. Our experiments will show that normalising the data term can be beneficial. Following [10, 11] and using the abbreviation $\mathbf{u} := (u, v)^\top$, we rewrite the data term M_1 as

$$M_1(u, v) = (\nabla_2 f^\top \mathbf{u} + f_t)^2 = |\nabla_2 f|^2 \left(\left(\frac{\nabla_2 f}{|\nabla_2 f|} \right)^\top (\mathbf{u} - \mathbf{u}_n) \right)^2 =: |\nabla_2 f|^2 d^2 . \quad (7)$$

The term d constitutes a projection of the difference between the estimated flow \mathbf{u} and the normal flow \mathbf{u}_n in the direction of the image gradient $\nabla_2 f$. Hence, this rewriting allows a geometric interpretation of the data constraint in terms of the distance from \mathbf{u} to the line described by the OFC (4). Ideally, we would like to penalise this distance d , but in the data term M_1 it is weighted by the squared spatial image gradient. This results in a stronger enforcement of the data constraint at high gradient locations. Such an overweighting may be inappropriate as large gradients can be caused by unreliable structures, such as noise or occlusions.

As a remedy, we normalise the data term M_1 by multiplying it with a factor [10, 11]

$$\theta_0 := \frac{1}{|\nabla_2 f|^2 + \zeta^2} , \quad (8)$$

where the regularisation parameter $\zeta > 0$ avoids division by zero. The normalised version of M_1 can be written as

$$M_2(u, v) = \mathbf{w}^\top \bar{J}_0 \mathbf{w}, \text{ with } \bar{J}_0 := \theta_0 J_0 . \quad (9)$$

Gradient Constancy Assumption. To cope with the problem of additive illumination changes, the *gradient constancy assumption (GCA)* has been proposed [4–6]. It states that image gradients remain constant under their displacement, i.e., $\nabla_2 f(\mathbf{x} + \mathbf{w}) = \nabla_2 f(\mathbf{x})$. A data term that combines both BCA and

GCA is

$$M_3(u, v) = \mathbf{w}^\top J \mathbf{w} , \quad (10)$$

where we use the *motion tensor notation* [16]

$$J := \nabla_3 f \nabla_3 f^\top + \gamma (\nabla_3 f_x \nabla_3 f_x^\top + \nabla_3 f_y \nabla_3 f_y^\top) . \quad (11)$$

Here, the parameter $\gamma > 0$ steers the contribution of the GCA.

To normalise M_3 , we replace the motion tensor J by

$$\bar{J} := \theta_0 \nabla_3 f \nabla_3 f^\top + \gamma (\theta_x \nabla_3 f_x \nabla_3 f_x^\top + \theta_y \nabla_3 f_y \nabla_3 f_y^\top) , \quad (12)$$

and obtain the data term $M_4(u, v) = \mathbf{w}^\top \bar{J} \mathbf{w}$. The two additional normalisation factors are defined as

$$\theta_x := \frac{1}{|\nabla_2 f_x|^2 + \zeta^2} , \quad \text{and} \quad \theta_y := \frac{1}{|\nabla_2 f_y|^2 + \zeta^2} . \quad (13)$$

Postponing the Linearisation. Linearisation of the data term with respect to u and v is only valid for small displacements. In order to handle large displacements correctly, Brox et al. [4] postpone any linearisation to the numerical scheme. Applying this strategy within the data term M_4 yields

$$M_5(u, v) = \left| \sqrt{\theta_0} (f(\mathbf{x} + \mathbf{w}) - f(\mathbf{x})) \right|^2 + \gamma \left(\left| \text{diag} \left(\sqrt{\theta_x}, \sqrt{\theta_y} \right) (\nabla_2 f(\mathbf{x} + \mathbf{w}) - \nabla_2 f(\mathbf{x})) \right|^2 \right) , \quad (14)$$

where $\text{diag}(a, b)$ denotes the 2×2 the diagonal matrix with the entries a and b .

We wish to emphasise that the numerical solution for large displacement optic flow proceeds by computing flow increments in a multiresolution framework. The linearisation of the data term M_5 w.r.t. these small increments will give rise to the motion tensor (12) on every image scale.

Colour Image Sequences. In a next step we extend the data term M_5 to multi-channel sequences by coupling three colour channels $(f^1(\mathbf{x}), f^2(\mathbf{x}), f^3(\mathbf{x}))$. A natural formulation for this is

$$M_6(u, v) = \sum_{i=1}^3 \left(\left| \sqrt{\theta_0^i} (f^i(\mathbf{x} + \mathbf{w}) - f^i(\mathbf{x})) \right|^2 + \gamma \left| \text{diag} \left(\sqrt{\theta_x^i}, \sqrt{\theta_y^i} \right) (\nabla_2 f^i(\mathbf{x} + \mathbf{w}) - \nabla_2 f^i(\mathbf{x})) \right|^2 \right) . \quad (15)$$

Photometric Invariant Colour Spaces. Realistic illumination models encompass a multiplicative influence [8], which cannot be captured by the GCA. This problem can be tackled by replacing the RGB colour space by the *Hue Saturation Value (HSV)* colour space [17] instead. The hue channel is invariant

under multiplicative illumination changes and in particular under shadow, shading, highlights and specularities. The saturation channel is only invariant w.r.t. shadow and shading and the value channel exhibits none of these invariances. In [9], only the hue channel was used for optic flow computation. We will additionally use the saturation and value channel, because they contain information that is not encoded in the hue channel.

Robust Penalisers. To provide robustness of the data term against outliers caused by noise and occlusions, Black and Anandan [3] proposed to refrain from a quadratic penalisation. Instead they use a non-quadratic penalisation function $\Psi_M(s^2)$, where s denotes the data constraint. Good results are reported in [4] for the regularised L_1 -norm, $\Psi_M(s^2) := \sqrt{s^2 + \varepsilon^2}$, with a small regularisation parameter $\varepsilon > 0$. Bruhn et al. [7] use a *separate* L_1 penalisation of the BCA and the GCA, which is advantageous if one assumption produces an outlier. In our variational framework we will go further by proposing a separate robustification of each HSV channel. It can be justified by the distinct information content of each of the three channels that drives the optic flow estimation in different ways.

Final Data Term. Incorporating our separate robustification idea into M_6 brings us to our final data term

$$M(u, v) = \sum_{i=1}^3 \Psi_M \left(\left| \sqrt{\theta_0^i} (f^i(\mathbf{x} + \mathbf{w}) - f^i(\mathbf{x})) \right|^2 \right) \quad (16)$$

$$+ \gamma \left(\sum_{i=1}^3 \Psi_M \left(\left| \text{diag} \left(\sqrt{\theta_x^i}, \sqrt{\theta_y^i} \right) (\nabla_2 f^i(\mathbf{x} + \mathbf{w}) - \nabla_2 f^i(\mathbf{x})) \right|^2 \right) \right).$$

This data term is (i) normalised, it (ii) combines the BCA and GCA, (iii) does not linearise the constancy assumptions and (iv) uses the HSV colour space with (v) a separate robustification of all colour channels.

To derive the contributions of the data term (16) to the Euler-Lagrange equations (2) and (3), we use the abbreviations from [4]:

$$f_{**} := \partial_{**} f(\mathbf{x} + \mathbf{w}), \quad f_z := f(\mathbf{x} + \mathbf{w}) - f(\mathbf{x}), \quad f_{*z} := \partial_* f(\mathbf{x} + \mathbf{w}) - \partial_* f(\mathbf{x}), \quad (17)$$

where $** \in \{x, y, xx, xy, yy\}$ and $* \in \{x, y\}$. Then we can write the contributions $\partial_u M$ and $\partial_v M$ as

$$\partial_u M = \sum_{i=1}^3 \Psi'_M \left(\theta_0^i (f_z^i)^2 \right) \cdot \theta_0^i f_z^i f_x^i \quad (18)$$

$$+ \gamma \left(\sum_{i=1}^3 \Psi'_M \left(\theta_x^i (f_{xz}^i)^2 + \theta_y^i (f_{yz}^i)^2 \right) \cdot (\theta_x^i f_{xz}^i f_{xx}^i + \theta_y^i f_{yz}^i f_{xy}^i) \right),$$

$$\partial_v M = \sum_{i=1}^3 \Psi'_M \left(\theta_0^i (f_z^i)^2 \right) \cdot \theta_0^i f_z^i f_y^i \quad (19)$$

$$+ \gamma \left(\sum_{i=1}^3 \Psi'_M \left(\theta_x^i (f_{xz}^i)^2 + \theta_y^i (f_{yz}^i)^2 \right) \cdot (\theta_x^i f_{xz}^i f_{xy}^i + \theta_y^i f_{yz}^i f_{yy}^i) \right),$$

where $\Psi'_M(s^2)$ denotes the derivative of $\Psi_M(s^2)$ w.r.t. its argument. Here we see that the separate robustification of the HSV channels makes sense: If a specific channel violates the imposed constancy assumption at a certain location, the corresponding argument of the decreasing function Ψ'_M will be large, yielding a downweighting of this channel. The other channels that satisfy the constancy assumption will then have a dominating influence on the solution.

4 Smoothness Term

4.1 Previous Smoothness Terms

For overcoming the aperture problem and for regularising the estimated flow field, energy-based methods include a smoothness term (regulariser). It models the assumption of a smooth flow field. A quadratic smoothness term as proposed by Horn and Schunck [1] penalises the squared magnitude of the flow gradients:

$$V_1(\nabla_2 u, \nabla_2 v) = |\nabla_2 u|^2 + |\nabla_2 v|^2 . \quad (20)$$

In the corresponding Euler-Lagrange equations, this leads to a homogeneous diffusion term that tends to blur important flow edges. Since it is desirable that regularisers permit flow discontinuities, numerous discontinuity perserving smoothness terms have been proposed. They are classified as either *image-* or *flow-driven*, depending on whether the smoothing process is adapted to image edges or evolving flow edges. In addition, one can distinguish *isotropic* and *anisotropic* strategies. Whereas the first type makes use of a scalar valued diffusivity to reduce the smoothing at edges, the latter also takes into account the directional information by means of a diffusion tensor. For an extensive and in-depth survey on classical discontinuity preserving regularisers, we refer to [13].

Joint Image- and Flow-driven Regularisation (JIF). Recently, Sun et al. [14] presented an anisotropic smoothness term in a discrete setting. It is modelled by a Steerable Random Field [15] that uses directional flow derivatives steered by image structures. It thus combines the advantages of image- and flow-driven regularisers, a strategy that we will name *joint image- and flow-driven (JIF)* regularisation. To obtain directional information of image structures, the authors analyse the eigenvectors of the structure tensor [18] $S_\rho := K_\rho * [\nabla_2 f \nabla_2 f^\top]$, where K_ρ is a Gaussian of standard deviation ρ and $*$ denotes the convolution operator. The structure tensor is a symmetric, positive semidefinite 2×2 matrix that possesses two orthonormal eigenvectors \mathbf{s}_1 and \mathbf{s}_2 with corresponding eigenvalues $\mu_1 \geq \mu_2 \geq 0$. The vector \mathbf{s}_1 points across image structures, whereas the vector \mathbf{s}_2 points along them. With these notations, the regulariser from [14] can be written as

$$\begin{aligned} V_2(\nabla_2 u, \nabla_2 v) = & \Psi_V \left((\mathbf{s}_1^\top \nabla_2 u)^2 \right) + \Psi_V \left((\mathbf{s}_1^\top \nabla_2 v)^2 \right) \\ & + \Psi_V \left((\mathbf{s}_2^\top \nabla_2 u)^2 \right) + \Psi_V \left((\mathbf{s}_2^\top \nabla_2 v)^2 \right) . \end{aligned} \quad (21)$$

The corresponding Euler-Lagrange equations are

$$\partial_u M - \alpha \operatorname{div} (D_u(\mathbf{s}_1, \mathbf{s}_2, \nabla_2 u) \nabla_2 u) = 0, \quad (22)$$

$$\partial_v M - \alpha \operatorname{div} (D_v(\mathbf{s}_1, \mathbf{s}_2, \nabla_2 v) \nabla_2 v) = 0, \quad (23)$$

with the diffusion tensors

$$D_p(\mathbf{s}_1, \mathbf{s}_2, \nabla_2 p) := (\mathbf{s}_1 | \mathbf{s}_2) \begin{pmatrix} \Psi'_V((\mathbf{s}_1^\top \nabla_2 p)^2) & 0 \\ 0 & \Psi'_V((\mathbf{s}_2^\top \nabla_2 p)^2) \end{pmatrix} \begin{pmatrix} \mathbf{s}_1^\top \\ \mathbf{s}_2^\top \end{pmatrix}, \quad (24)$$

for $p \in \{u, v\}$. We observe that this regulariser indeed exhibits the desired image- and flow-driven behaviour: The regularisation direction is adapted to the image structure directions \mathbf{s}_1 and \mathbf{s}_2 , whereas the magnitude of the regularisation depends on the flow contrast encoded in $\nabla_2 p$. As a result, this regulariser yields the same sharp flow edges as image-driven methods but does not suffer from oversegmentation problems.

4.2 Our Novel Constraint Adaptive Regulariser (CAR)

In spite of its sophistication, the JIF model still suffers from a few shortcomings. As a remedy we will introduce three amendments that will be discussed now.

Regularisation Tensor. A first remark w.r.t. JIF regularisation is that the directional information from the structure tensor S_ρ is not consistent with the imposed constraints of our data term (16). It is more natural to take into account the directional information provided by the motion tensor (12) and to steer the anisotropic regularisation process w.r.t. “constraint edges” instead of image edges. We propose to analyse the eigenvectors \mathbf{r}_1 and \mathbf{r}_2 of the *regularisation tensor*

$$R_\rho := \sum_{i=1}^3 K_\rho * \left[\theta_0^i \nabla_2 f^i (\nabla_2 f^i)^\top + \gamma \left(\theta_x^i \nabla_2 f_x^i (\nabla_2 f_x^i)^\top + \theta_y^i \nabla_2 f_y^i (\nabla_2 f_y^i)^\top \right) \right]. \quad (25)$$

The regularisation tensor integrates neighbourhood information of the motion tensor entries for every colour channel. By exploiting the invariances of the HSV colour space, it is not prone to be misled by “phantom” edges, like shadow edges.

Rotational Invariance. Unfortunately the smoothness term V_2 lacks the desirable property of rotational invariance, because the projections of $\nabla_2 u$ and $\nabla_2 v$ onto the eigenvector directions are penalised separately. As a remedy we propose to jointly penalise the projections on the eigenvector directions of the regularisation tensor, yielding

$$V_3(\nabla_2 u, \nabla_2 v) = \Psi_V \left((\mathbf{r}_1^\top \nabla_2 u)^2 + (\mathbf{r}_1^\top \nabla_2 v)^2 \right) + \Psi_V \left((\mathbf{r}_2^\top \nabla_2 u)^2 + (\mathbf{r}_2^\top \nabla_2 v)^2 \right). \quad (26)$$

Single Robust Penalisation. The regulariser V_3 performs a *twofold robust penalisation* in both eigenvector directions. Because the data term mainly constrains the flow in the direction of the largest eigenvalue of the spatial motion tensor, we propose a *single robust penalisation* solely in \mathbf{r}_1 -direction. In the orthogonal \mathbf{r}_2 -direction, we opt for a strong quadratic penalisation. The advantages of this strong filling-in effect along constraint edges will be confirmed by our experiments in Sec. 6. Also incorporating the single robust penalisation yields the regulariser

$$V(\nabla_2 u, \nabla_2 v) = \Psi_V \left((\mathbf{r}_1^\top \nabla_2 u)^2 + (\mathbf{r}_1^\top \nabla_2 v)^2 \right) + (\mathbf{r}_2^\top \nabla_2 u)^2 + (\mathbf{r}_2^\top \nabla_2 v)^2, \quad (27)$$

where we use the Perona-Malik regulariser (Lorentzian) [19, 20] given by $\Psi_V(s^2) := \lambda^2 \log(1 + (s^2/\lambda^2))$ with a contrast parameter $\lambda > 0$. We call the regulariser V the *constraint adaptive regulariser (CAR)*. It complements the proposed robust data term M from (16) in an optimal fashion.

The corresponding Euler-Lagrange equations are

$$\partial_u M - \alpha \operatorname{div} (D(\mathbf{r}_1, \mathbf{r}_2, \nabla_2 u, \nabla_2 v) \nabla_2 u) = 0, \quad (28)$$

$$\partial_v M - \alpha \operatorname{div} (D(\mathbf{r}_1, \mathbf{r}_2, \nabla_2 u, \nabla_2 v) \nabla_2 v) = 0, \quad (29)$$

where the joint diffusion tensor is given by

$$D(\mathbf{r}_1, \mathbf{r}_2, \nabla_2 u, \nabla_2 v) := \begin{pmatrix} \mathbf{r}_1 & \mathbf{r}_2 \end{pmatrix} \begin{pmatrix} \Psi'_V \left((\mathbf{r}_1^\top \nabla_2 u)^2 + (\mathbf{r}_1^\top \nabla_2 v)^2 \right) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r}_1^\top \\ \mathbf{r}_2^\top \end{pmatrix}. \quad (30)$$

Comparing our joint diffusion tensor (30) to its JIF counterparts (24), the following innovations become apparent: (i) The smoothing direction is adapted to constraint edges instead of image edges. (ii) We achieve rotational invariance by coupling the two flow components in the argument of Ψ'_V . (iii) We only reduce the smoothing *across* constraint edges. Along them, always a strong diffusion with strength 1 is performed, resembling edge-enhancing anisotropic diffusion [21].

When using $\partial_u M$ and $\partial_v M$ as given in (18) and (19) in the Euler-Lagrange equations (28) and (29), we obtain the Euler-Lagrange equations for our proposed *complementarity optic flow (COF)* method.

5 Implementation

To solve the Euler-Lagrange equations we use a coarse-to-fine multiscale warping approach [4]. On each warping level, small flow increments are computed via a linearised approach, allowing to handle large displacements correctly. The computations are speeded up by a nonlinear multigrid scheme [16] that solves the problem at each warping level based on a Gauß-Seidel type solver with alternating line relaxation [16].

Spatial image and flow derivatives are discretised via central finite differences of fourth and second order, respectively [16]. For the motion tensor, these derivatives are averaged from the two frames $f(x, y, t)$ and $f(x, y, t + 1)$, whereas for the regularisation tensor, they are solely computed at the first frame.

6 Experiments

Our first experiment shows the importance of different constituents of our model. We compare our method to four modified versions of it where we have changed one distinct feature: (i) No data term normalisation. (ii) Using a regulariser with twofold instead of single robust penalisation as in V_3 . (iii) Using a regulariser with single robust penalisation as in V , but based on the structure tensor instead of the regularisation tensor. (iv) Using the RGB colour space. For the latter version, we only separately robustify the BCA and the GCA, as a separate robustification of the RGB channels makes no sense. In Fig. 1, we show results for the *Urban3* sequence from the recent optic flow database [22] of the Middlebury University⁵. To visualise flow fields we use a colour code where colour encodes the flow direction and brightness the magnitude, as is shown in Fig. 1 (c). Throughout our experiments we use the parameters $\zeta = 0.1$, $\varepsilon = 0.001$, $\lambda = 0.1$. Specifically for the *Urban3* sequence, we fixed the parameters $\sigma = 0.7$, $\gamma = 1.0$, $\rho = 1.5$ and only tuned the value of α , as is given in the caption of Fig. 1. There, we also state the corresponding average angular error (AAE) measures [23] in order to compare the quality of estimated flow fields to the ground truth. Note that the errors were computed for the whole image, whereas for visualisation purposes, the flow fields in Fig. 1 (d)–(i) show details. We notice a lot of artifacts for the method without data term normalisation (Fig. 1 (e)) that severely deteriorate the flow estimate. With a twofold robust penalisation (Fig. 1 (f)), artifacts at flow edges emerge due to the inhibited smoothing along edges. The results for the RGB version (Fig. 1 (g)) and our approach (Fig. 1 (i)) look rather similar due to the uncritical illumination conditions in this synthetic sequence. However, at the connection between the two buildings in the middle of the image, our approach performs better. When using the directional information from the structure tensor (Fig. 1 (h)), the results look promising, but artifacts at flow corners appear.

For a second experiment we created a real world test sequence with difficult illumination conditions caused by pronounced shadows, see Fig. 2 (a)–(b). Using this test sequence, we compare our method to the RGB version, the method of Brox et al. [4] and a version of our approach with a rotationally invariant JIF regulariser. This regulariser is similar to V_3 but uses the eigenvectors \mathbf{s}_1 and \mathbf{s}_2 instead of \mathbf{r}_1 and \mathbf{r}_2 . As fixed parameters we set $\sigma = 0.5$, $\gamma = 20.0$ and $\rho = 2.5$. We see that the RGB method (Fig. 2 (c)) suffers from artifacts due to the shadows in the marked regions. When using the JIF regulariser (Fig. 2 (d)), the flow edges are dislocated and the shadow edges in the marked regions yield unpleasant artifacts. This demonstrates the drawbacks of the use of the structure tensor instead of the regularisation tensor, and of the twofold robust penalisation. Because of the latter, perturbing staircasing artifacts arise. The method of Brox et al. [4] (Fig. 2 (e)), that is considered to be accurate and robust, gives poor results for this sequence. Solely our method (Fig. 2 (f)) produces an agreeable flow field

⁵ available under <http://vision.middlebury.edu/flow/data/>

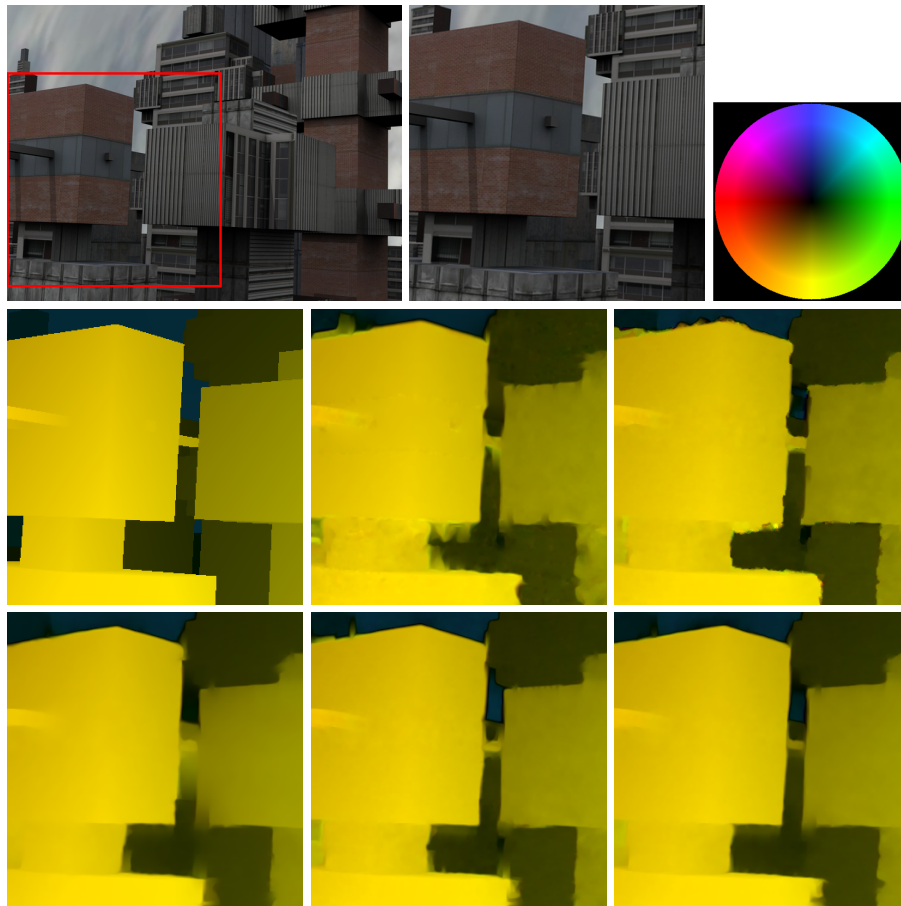


Fig. 1. The *Urban3* sequence. **First row:** (a) Frame 10. (b) Zoom in marked region. (c) Used colour code. **Second row:** (d) Ground truth flow field in marked region. (e) Corresponding flow field without normalisation ($\alpha = 300.0$, AAE=4.57). (f) Twofold robust penalisation ($\alpha = 50.0$, AAE=3.56). **Third row:** (g) RGB colour space ($\alpha = 100.0$, AAE=3.09). (h) Structure tensor ($\alpha = 50.0$, AAE=2.99). (i) Our approach ($\alpha = 75.0$, AAE=2.95).

Table 1. The Top 8 of the Middlebury ranking for the AAE (as of June 12, 2009).

Method	Our Method	Adaptive	Aniso. Huber-L1	Spatially variant	TV-L1-improved	Occlusion bounds	Brox et al.	Multicue MRF
Avg. rank	4.8	5.0	6.5	7.1	7.9	8.6	9.2	9.2

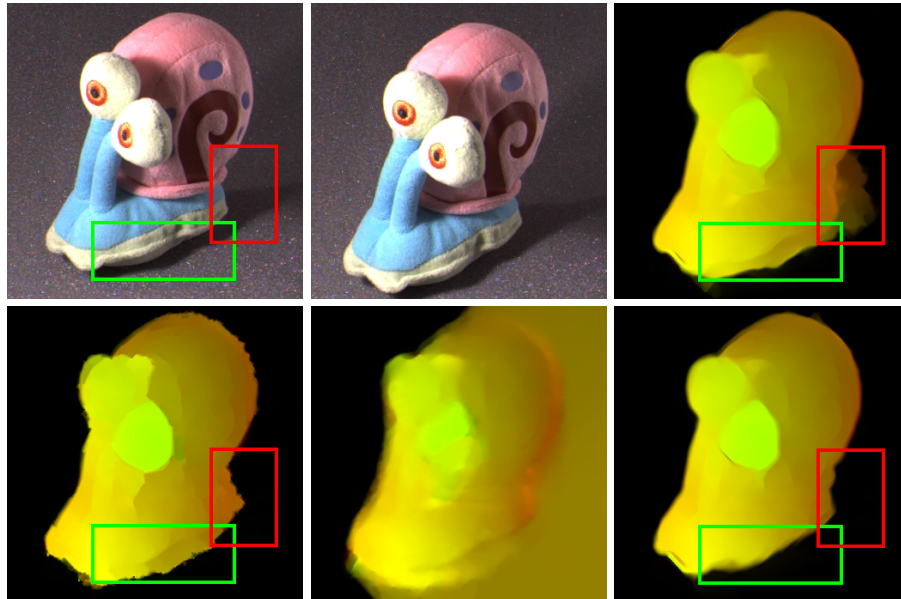


Fig. 2. The *Snail* sequence. **First row:** (a) Frame 1. (b) Frame 2. (c) Flow field with RGB colour space ($\alpha = 800.0$). **Second row:** (d) JIF ($\alpha = 1500.0$) (e) Brox et al. [4] ($\alpha = 75.0$). (f) Our approach ($\alpha = 1500.0$).

in spite of the difficult illumination conditions and the large displacements (up to 25 pixels) in this sequence.

For a final comparison of our method to state-of-the-art approaches, we submitted our results to the Middlebury benchmark page⁶. In accordance to their guidelines, we used a fixed set of parameters for all sequences: $\alpha = 600.0$, $\sigma = 0.5$, $\gamma = 20.0$ and $\rho = 2.5$. In Tab. 1, we show the average rank of the Top 8 methods for the AAE. With the proposed method we are able to achieve the first rank. This shows that a sophisticated and transparent modelling allows to outperform other well-engineered methods that incorporate many more processing steps. The running time for the *Urban* sequence was 44.3 s on a standard PC (3.2 GHz Intel Pentium 4, 256 MB RAM). This proves that the used multigrid scheme [16] allows to obtain moderate runtimes for standard test sequences.

7 Conclusions and Outlook

We have presented a novel variational optic flow technique based on the concept of complementarity between data and smoothness term. By refraining from the traditional viewpoint that such terms are natural competitors within a joint

⁶ available under <http://vision.middlebury.edu/flow/eval/results/>

energy-based framework, we succeeded to unify their advantages and achieve the currently most accurate results in the Middlebury benchmark.

Our data term integrates sophisticated components such as embedding higher order constancy assumptions in an HSV colour representation with a separate robust penalisation of each channel, renouncement of linearisations, and constraint normalisation. The directional information that results from these constraints is used in our complementary anisotropic smoothness term. This smoothness term combines the advantages of image- and flow-driven regularisation. However, compared to the approach of Sun et al. [14], it is rotationally invariant, respects constraint edges instead of image edges, and it restricts robust penalisation to the constraint direction. We have given detailed motivations showing that these model refinements arise in a natural and systematic way. Moreover, in the experiments we have proven that each of our amendments in the data and smoothness term is beneficial and contributes to the favourable accuracy of our complementary optic flow (COF) approach.

We hope that our research triggers further investigations on incorporating complementarity concepts in image processing and computer vision. This may allow to exploit similar synergies also in the context of other tasks that are currently dominated by energy-based strategies.

Acknowledgements. Our research was partly funded by the International Max-Planck Research School (IMPRS), the Deutsche Forschungsgemeinschaft (DFG) under the project *WE 2602/6-1*, and the German Academic Exchange Service (DAAD). This is gratefully acknowledged.

References

1. Horn, B., Schunck, B.: Determining optical flow. *Artificial Intelligence* **17** (1981) 185–203
2. Nagel, H.H., Enkelmann, W.: An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8** (1986) 565–593
3. Black, M.J., Anandan, P.: Robust dynamic motion estimation over time. In: *Proc. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Maui, HI, IEEE Computer Society Press (June 1991) 292–302
4. Brox, T., Bruhn, A., Papenber, N., Weickert, J.: High accuracy optical flow estimation based on a theory for warping. In Pajdla, T., Matas, J., eds.: *Computer Vision – ECCV 2004, Part IV*. Volume 3024 of *Lecture Notes in Computer Science*. Springer, Berlin (2004) 25–36
5. Schnörr, C.: Segmentation of visual motion by minimizing convex non-quadratic functionals. In: *Proc. Twelfth International Conference on Pattern Recognition*. Volume A., Jerusalem, Israel, IEEE Computer Society Press (October 1994) 661–663
6. Tretiak, O., Pastor, L.: Velocity estimation from image sequences with second order differential operators. In: *Proc. Seventh International Conference on Pattern Recognition*, Montreal, Canada (July 1984) 16–19

7. Bruhn, A., Weickert, J.: Towards ultimate motion estimation: Combining highest accuracy with real-time performance. In: Proc. Tenth International Conference on Computer Vision. Volume 1., Beijing, China, IEEE Computer Society Press (October 2005) 749–755
8. van de Weijer, J., Gevers, T.: Robust optical flow from photometric invariants. In: Proc. 2004 IEEE International Conference on Image Processing. Volume 3., Singapore, IEEE Signal Processing Society (October 2004) 1835–1838
9. Mileva, Y., Bruhn, A., Weickert, J.: Illumination-robust variational optical flow with photometric invariants. In Hamprecht, F., Schnörr, C., Jähne, B., eds.: Pattern Recognition. Volume 4713 of Lecture Notes in Computer Science. Springer, Berlin (2007) 152–162
10. Simoncelli, E.P., Adelson, E.H., Heeger, D.J.: Probability distributions of optical flow. In: Proc. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Maui, HI, IEEE Computer Society Press (June 1991) 310–315
11. Lai, S.H., Vemuri, B.C.: Reliable and efficient computation of optical flow. *International Journal of Computer Vision* **29**(2) (October 1998) 87–105
12. Alvarez, L., Esclarín, J., Lefébure, M., Sánchez, J.: A PDE model for computing the optical flow. In: Proc. XVI Congreso de Ecuaciones Diferenciales y Aplicaciones, Las Palmas de Gran Canaria, Spain (September 1999) 1349–1356
13. Weickert, J., Schnörr, C.: A theoretical framework for convex regularizers in PDE-based computation of image motion. *International Journal of Computer Vision* **45**(3) (December 2001) 245–264
14. Sun, D., Roth, S., Lewis, J., Black, M.: Learning optical flow. In Forsyth, D., Torr, P., Zisserman, A., eds.: Computer Vision – ECCV 2008, Part III. Volume 5304 of Lecture Notes in Computer Science. Springer, Berlin (2008) 83–97
15. Roth, S., Black, M.: Steerable random fields. In: Proc. 2007 IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, IEEE Computer Society Press (October 2007)
16. Bruhn, A., Weickert, J., Kohlberger, T., Schnörr, C.: A multigrid platform for real-time motion computation with discontinuity-preserving variational methods. *International Journal of Computer Vision* **70**(3) (December 2006) 257–277
17. Golland, P., Bruckstein, A.M.: Motion from color. *Computer Vision and Image Understanding* **68**(3) (December 1997) 346–362
18. Förstner, W., Gülch, E.: A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland (June 1987) 281–305
19. Black, M.J., Anandan, P.: The robust estimation of multiple motions: parametric and piecewise smooth flow fields. *Computer Vision and Image Understanding* **63**(1) (January 1996) 75–104
20. Perona, P., Malik, J.: Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12** (1990) 629–639
21. Weickert, J.: Theoretical foundations of anisotropic diffusion in image processing. *Computing Supplement* **11** (1996) 221–236
22. Baker, S., Roth, S., Scharstein, D., Black, M., Lewis, J., Szeliski, R.: A database and evaluation methodology for optical flow. In: Proc. 2007 IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, IEEE Computer Society Press (October 2007)
23. Barron, J.L., Fleet, D.J., Beauchemin, S.S.: Performance of optical flow techniques. *International Journal of Computer Vision* **12**(1) (February 1994) 43–77