An Unbiased Second-Order Prior for High-Accuracy Motion Estimation*

Werner Trobin[1], Thomas Pock[1,2], Daniel Cremers[2], and Horst Bischof[1]

[1] Institute for Computer Graphics and Vision, Graz University of Technology
[2] Department of Computer Science, University of Bonn

Abstract. Virtually all variational methods for motion estimation regularize the gradient of the flow field, which introduces a bias towards piecewise constant motions in weakly textured areas. We propose a novel regularization approach, based on decorrelated second-order derivatives, that does not suffer from this shortcoming. We then derive an efficient numerical scheme to solve the new model using projected gradient descent. A comparison to a TV regularized model shows that the proposed second-order prior exhibits superior performance, in particular in low-textured areas (where the prior becomes important). Finally, we show that the proposed model yields state-of-the-art results on the Middlebury optical flow database.

1 Introduction

Estimating the optical flow between two consecutive images of a scene requires finding corresponding points in the images. Since optical flow is a highly ill-posed, inverse problem, solely relying on the optical flow constraint, i.e. to assume that the intensities remained constant, does not provide enough information to infer meaningful flow fields. In particular, optical flow computation suffers from two problems: first, no information is available in untextured regions. Second, one can only compute the normal flow, i.e. the motion in direction of the image gradient; a restriction known as the *aperture problem*. Optical flow algorithms therefore have to implement an implicit or explicit prior model, to infer the flow in these problematic areas based on the flow field in a local neighborhood – this is typically referred to as smoothing or regularization.

Horn and Schunck were the first to formulate the optical flow problem in the framework of the calculus of variations [1]:

$$\min_{\boldsymbol{u}} \left\{ \int_{\Omega} |\nabla u_1|^2 + |\nabla u_2|^2 d\Omega + \lambda \int_{\Omega} \left(I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - I_0(\boldsymbol{x}) \right)^2 d\Omega \right\} \quad , \qquad (1)$$

where I_0 and I_1 are the images, $\boldsymbol{u}(\boldsymbol{x}) = (u_1(\boldsymbol{x}), u_2(\boldsymbol{x}))^T$ is the displacement field, and the parameter λ balances the influence of the regularization and the optical

* This work was supported by the Austrian Science Fund under grant P18110-B15 and the Austrian Research Promotion Agency within the VM-GPU project (no. 813396).

G. Rigoll (Ed.): DAGM 2008, LNCS 5096, pp. 396–405, 2008.

flow constraint. The choice of a quadratic error function in both the smoothness and the data terms leads to a well-understood system of linear partial differential equations. Yet, quadratic penalizers are sensitive to discontinuities in the flow field and to outliers in the data term.

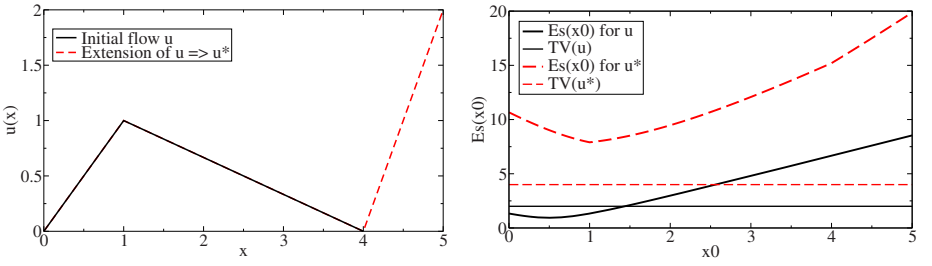
Numerous improvements to the Horn and Schunck model have been proposed to overcome its inherent limitations. Black and Anandan [2] introduced methods from robust statistics to avoid the quadratic penalization and allow for flow discontinuities and outliers in the data term. Motivated by the observation that the flow discontinuities at the boundaries of moving objects often coincide with large image gradients, Nagel et al. [3], Weickert et al. [4], and other authors proposed to replace the homogeneous regularization in the Horn-Schunck model with anisotropic diffusion approaches. For a more systematic survey on variational formulations of the optical flow problem, we refer to Aubert et al. [5].

State-of-the-art variational approaches allow robust calculation of highly accurate flow fields [6], some of them even in real-time [7,8]. Virtually all of them use robust penalty functions, like total variation (TV) or the Lorentzian, to regularize the gradient of the flow field. Using the flow gradient, however, causes problems in untextured areas, where such a regularization tends to yield piecewise constant solutions. For stereo computation, a strongly related problem, Ishikawa recently showed that an affinity towards piecewise constant regions in a disparity map, i.e. fronto-parallel surfaces, is not consistent with human perception [9]. He proposed a novel prior model based on the total absolute Gaussian curvature, which unfortunately is hard to optimize. However, his results indicate that it is beneficial to penalize second-order derivatives. The main contribution of this paper is that we propose a novel flow regularization approach based on second-order derivatives, thereby inhibiting the implicit preference of piecewise constant flow fields. In contrast to Yuan et al. [10], who used higher-order derivatives for simultaneous flow estimation and decomposition, we do not aim to decompose the flow into physically meaningful parts, which significantly simplifies the problem.

Several authors proposed discrete methods to estimate the optical flow by fitting affine flow patches [11,12,13]. They typically split the domain into disjoint patches and estimate the parameters of the prevalent affine motion per patch. This leads to rather coarse estimates, especially near flow discontinuities. In the next section, we will briefly review the recent variational approach of Nir et al. [14], who proposed an over-parameterized, affine extension of the Horn and Schunck model. Motivated by the shortcomings of this approach, we will propose our second-order flow regularization approach in Sec. 3.

2 Affine Flow Via Over-Parameterization

Nir et al. [14] proposed to model the flow field using piecewise affine regions and penalize only the *deviations* from these models. To simplify the discussion, we will consider only a single horizontal line in a rectified stereo image pair, i.e. a one-dimensional flow problem. Even in this case, we can show that the approach



(a) (b)

Fig. 1. The given 1D flow field u and the extension yielding u^* is shown in (a). For both flow fields, (b) shows the smoothing energy E_S for the over-parameterized model as a function of x_0. The horizontal lines are TV(u) and TV(u^*), respectively.

of Nir et al. may not lead to the desired solution. In the 1D setting, the proposed affine model reduces to $u(x) = A_1 + A_2\hat{x}$, where $\hat{x} = \rho(x - x_0)/x_0$ and x_0 is fixed at the center of the domain. The weighting coefficients $A_i(x)$ are free to vary across the domain and the parameter ρ allows balancing their relative weight. Without loss of generality, we set $\rho = 1$ for the remainder of this section.

This over-parameterized representation allows multiple descriptions for any reasonable flow field, so instead of regularizing the flow field $u(x)$, Nir et al. suggest to directly regularize the coefficients A_i, using a convex approximation of the L^1 norm. For affine regions, the benefit of such a model seems to be clear: instead of directly representing the flow $\boldsymbol{u}$, incurring costs for its affine changes, the model allows representing affine motion fields with constant parameters, i.e. with zero cost. However, this benefit of constant parameters *within* affine regions is often outweighed by the costs of large discontinuities at the region borders.

In order to illustrate that, we will try to find exact representations for the piecewise affine flow field $u(x)$ given in Fig. 1(a), by following two alternative strategies. The first strategy is to set $A_2(x) = 0$ and $A_1(x) = u(x)$, i.e. to ignore the affine extension and fall back to a plain, TV-regularized flow model. The second strategy is to make use of this extension and select a constant set of coefficients for every affine region of u. For a fixed x_0, such a strategy yields one unique set $\{A_1^i, A_2^i\}$ per region i.

For the basic TV-regularized model, the smoothing cost is just TV(u), as depicted in Fig. 1(b). When using the affine extension, we get the energy depicted by the thick, solid line in Fig. 1(b). The two key observations are: first, the global minimum for the affine model is lower than TV(u) and second, for the ad-hoc choice of $x_0 = 2$ (in the center of the domain, as suggested by Nir et al.) it is actually higher than TV(u). Let us extend u with another linear region (the dashed line in Fig. 1(a)), yielding u^*. This innocuous looking extension leads to a significant advantage for the plain TV model (cf. the dashed lines in Fig. 1(b)). These surprisingly high smoothing costs for the affine model are caused by the step edges in the (by design) piecewise constant coefficient space; an effect amplified with increasing distance to the "origin" of the representation (x_0). Hence,

contrary to the expectations, the coefficients of the inferred solutions rarely are piecewise constant, even for perfectly affine motions. This limitation can only be resolved by allowing multiple "origins" x_0, which requires a simultaneous segmentation and estimation of the flow field, a problem that is inherently non-convex and therefore hard to optimize.

3 An Unbiased Second-Order Prior

Instead of modeling affine functions based on an explicit parameterization, our intention is to design an appropriate differential operator, having the intrinsic property to penalize only deviations from affine functions. First of all, we notice that second-order derivatives of affine functions are zero, i.e. a regularization based on second-order derivatives causes zero cost for affine functions. Unfortunately, second-order derivative operators are not orthogonal and the local information of orientation and shape are entangled. Therefore, penalizing each of them, e.g. using the Euclidean vector norm, would induce a bias in the sense that certain affine fields would be energetically favored over others.

In [15], Danielsson et al. defined a new operator set to resolve these problems. They used circular harmonic functions to map the the second-order derivative operators into an orthogonal space. In two spatial dimensions, the new operator is given by

$$\diamond = \sqrt{\frac{1}{3}} \left(\frac{\partial^2}{\partial_x^2} + \frac{\partial^2}{\partial_y^2},\ \sqrt{2} \left(\frac{\partial^2}{\partial_x^2} - \frac{\partial^2}{\partial_y^2} \right),\ \sqrt{8}\, \frac{\partial^2}{\partial_x \partial_y} \right)^T . \tag{2}$$

More interestingly, the magnitude can be directly calculated as the Euclidean vector norm:

$$\|\diamond u\| = \sqrt{\frac{1}{3}} \sqrt{\left(\frac{\partial^2 u}{\partial_x^2} + \frac{\partial^2 u}{\partial_y^2} \right)^2 + 2 \left(\frac{\partial^2 u}{\partial_x^2} - \frac{\partial^2 u}{\partial_y^2} \right)^2 + 8 \left(\frac{\partial^2 u}{\partial_x \partial_y} \right)^2} . \tag{3}$$

As intended, this norm exactly measures the local deviation of a function u from being affine. Hence, it is quite natural to use this norm as a new prior model for optical flow.

Based on [8], we propose the following variational model to compute the optical flow $\boldsymbol{u} = (u_1, u_2)^T$ of two images I_0 and I_1:

$$\min_{\boldsymbol{u}} \left\{ \sum_{d=1}^{2} \int_\Omega \|\diamond u_d\|\ d\Omega + \lambda \int_\Omega |\rho(\boldsymbol{u})|\ d\Omega \right\} . \tag{4}$$

The first term utilizes the new prior model to regularize the flow field in each direction u_d. Note that this term is non-smooth and therefore allows for discontinuities in the second-order derivatives. The second term penalizes violations of the optical flow constraint using the robust L^1 norm, where $\rho(\boldsymbol{u})$ denotes the locally linearized image residual given by $\rho(\boldsymbol{u}) = I_1(\boldsymbol{x} + \boldsymbol{u}_0) + \langle \boldsymbol{u} - \boldsymbol{u}_0, \nabla I_1 \rangle - I_0$

and $\boldsymbol{u}_0$ is a given disparity map [8]. Due to the linearization procedure, (4) becomes convex, but it has to be embedded into an iterative warping approach to compensate for image non-linearities [6].

Let us now describe an algorithm to compute the minimizer of (4). Since (4) is convex but non-strictly convex, we introduce an auxiliary variable $\boldsymbol{v} = (v_1, v_2)^T$ and propose to minimize the following convex approximation of (4):

$$\min_{\boldsymbol{u},\boldsymbol{v}} \left\{ \sum_{d=1}^{2} \int_\Omega \|\diamond u_d\| \, d\Omega + \frac{1}{2\theta} \sum_{d=1}^{2} \int_\Omega (u_d - v_d)^2 \, d\Omega + \lambda \int_\Omega |\rho(\boldsymbol{v})| \, d\Omega \right\} . \quad (5)$$

Unlike (4), (5) is a minimization problem in two variables $\boldsymbol{u}$ and $\boldsymbol{v}$. We therefore have to perform an alternating minimization procedure:

1. For $\boldsymbol{v}$ fixed, solve for every u_d:

$$\min_{u_d} \left\{ \int_\Omega \|\diamond u_d\| \, d\Omega + \frac{1}{2\theta} \int_\Omega (u_d - v_d)^2 \, d\Omega \right\} . \quad (6)$$

2. For $\boldsymbol{u}$ fixed, solve for $\boldsymbol{v}$:

$$\min_{\boldsymbol{v}} \left\{ \frac{1}{2\theta} \sum_{d=1}^{2} \int_\Omega (u_d - v_d)^2 \, d\Omega + \lambda \int_\Omega |\rho(\boldsymbol{v})| \, d\Omega \right\} . \quad (7)$$

The solutions of both steps are summarized in the following propositions.

Proposition 1. *The solution of (6) is given by*

$$u_d = v_d - \theta \diamond \cdot \boldsymbol{p}_d \ . \quad (8)$$

The dual variable $\boldsymbol{p}_d$ is obtained as the steady state of

$$\begin{aligned} \tilde{\boldsymbol{p}_d}^{k+1} &= \boldsymbol{p}^k + \frac{\tau}{\theta} \left[\diamond \left(v_d - \theta \diamond \cdot \boldsymbol{p}_d^k \right) \right] , \\ \boldsymbol{p}_d^{k+1} &= \frac{\tilde{\boldsymbol{p}_d}^{k+1}}{\max \left\{ 1, |\tilde{\boldsymbol{p}_d}^{k+1}| \right\}} , \end{aligned} \quad (9)$$

where k is the iteration number, $\boldsymbol{p}_d^0 = \boldsymbol{0}$, and $\tau \leq 3/112$.

Proof: Note that for convenience, we have dropped the subscript d out of (6). To overcome the non-differentiability of (3), we employ its dual formulation

$$\|\diamond u\| \equiv \max_{\|\boldsymbol{p}\| \leq 1} \{ \diamond u \cdot \boldsymbol{p} \} , \quad (10)$$

where $\boldsymbol{p} = (p^1, p^2, p^3)^T$ is the dual variable. Next, we follow the approach of [16] to derive the dual formulation of (6):

$$\min_{\|\boldsymbol{p}\| \leq 1} - \int_\Omega \diamond v \cdot \boldsymbol{p} + \frac{\theta}{2} (\diamond \cdot \boldsymbol{p})^2 \, d\Omega . \quad (11)$$

Hence, solving (11) is equivalent to solving the primal problem (6), which can in turn be recovered via

$$u = v - \theta \diamondsuit \cdot \boldsymbol{p} \; . \tag{12}$$

We proceed by describing a procedure to solve the dual problem (11). First we write down the Euler Lagrange equation of (11):

$$-\diamondsuit \cdot (v - \theta \diamondsuit \cdot \boldsymbol{p}) = 0 \quad \text{s.t.} \quad \|\boldsymbol{p}\| \leq 1 \; . \tag{13}$$

Being convex in $\boldsymbol{p}$, minimizing (11) is achieved by gradient descend of (13) and subsequent reprojection of $\boldsymbol{p}$ to ensure $\|\boldsymbol{p}\| \leq 1$ [17]. The gradient descend step reads as

$$\tilde{\boldsymbol{p}}^{k+1} = \boldsymbol{p}^k + \Delta t \left[\diamondsuit \left(v - \theta \diamondsuit \cdot \boldsymbol{p} \right) \right] \; , \tag{14}$$

and the subsequent reprojection is given by:

$$\boldsymbol{p}^{k+1} = \frac{\tilde{\boldsymbol{p}}^{k+1}}{\max \left\{ 1, \|\tilde{\boldsymbol{p}}^{k+1}\| \right\}} \; . \tag{15}$$

Finally, we have to determine the maximum value of Δt such that (14) remains stable. We employ the following standard finite differences approximation of the $\diamondsuit$ operator:

$$(\diamondsuit u)_{i,j} = \begin{pmatrix} \sqrt{\frac{1}{3}} \left(u_{i,j-1} + u_{i,j+1} + u_{i-1,j} + u_{i+1,j} - 4u_{i,j} \right) \\ \sqrt{\frac{2}{3}} \left(u_{i-1,j} + u_{i+1,j} - u_{i,j-1} - u_{i,j+1} \right) \\ \sqrt{\frac{8}{3}} \left(u_{i,j} + u_{i+1,j+1} - u_{i,j+1} - u_{i+1,j} \right) \end{pmatrix} \; , \tag{16}$$

and

$$\begin{aligned} (\diamondsuit \cdot \boldsymbol{p})_{i,j} = & \sqrt{1/3} \left(p^1_{i,j-1} + p^1_{i,j+1} + p^1_{i-1,j} + p^1_{i+1,j} - 4p^1_{i,j} \right) \\ & + \sqrt{2/3} \left(p^2_{i-1,j} + p^2_{i+1,j} - p^2_{i,j-1} - p^2_{i,j+1} \right) \\ & + \sqrt{8/3} \left(p^3_{i,j} + p^3_{i-1,j-1} - p^3_{i,j-1} - p^3_{i-1,j} \right) \; , \end{aligned} \tag{17}$$

where (i, j) denote the indices of the discrete image domain. Since (14) is linear in $\boldsymbol{p}_{i,j}$, stability of (14) is guaranteed as long as

$$\Delta t \theta \rho(\diamondsuit^2 \boldsymbol{p})_{i,j} \leq 1 \; , \tag{18}$$

where $\rho(\diamondsuit^2 \boldsymbol{p})_{i,j}$ denotes the spectral radius and $\diamondsuit^2 = \diamondsuit \cdot \diamondsuit$. It therefore remains to give an upper bound of $\rho(\diamondsuit^2 \boldsymbol{p})_{i,j}$. In fact, $\rho(\diamondsuit^2 \boldsymbol{p})_{i,j} \leq \|\diamondsuit^2 \boldsymbol{p}\|_{i,j}$. Hence, we end up with an upper bound of $\Delta t \leq \frac{\tau}{\theta}$, where we explicitely computed $\tau = 3/112$. □

Proposition 2. *The solution of the minimization task in (7) is given by the following thresholding step:*

$$\boldsymbol{v} = \boldsymbol{u} + \begin{cases} \lambda \theta \nabla I_1 & \text{if} \quad \rho(\boldsymbol{u}) < -\lambda \theta |\nabla I_1|^2 \\ -\lambda \theta \nabla I_1 & \text{if} \quad \rho(\boldsymbol{u}) > \lambda \theta |\nabla I_1|^2 \\ -\rho(\boldsymbol{u}) \nabla I_1 / |\nabla I_1|^2 & \text{if} \quad |\rho(\boldsymbol{u})| \leq \lambda \theta |\nabla I_1|^2 \; . \end{cases} \tag{19}$$

Proof: This proof is a more technical variant of the proof presented in [8]. Let us first write down the Euler Lagrange equation of (7):

$$\boldsymbol{v} - \boldsymbol{u} + \lambda\theta \frac{\rho(\boldsymbol{v})}{|\rho(\boldsymbol{v})|} \nabla I_1 = 0 \ . \tag{20}$$

We proceed by analyzing the three possible cases $\rho(\boldsymbol{v}) > 0$, $\rho(\boldsymbol{v}) < 0$ and $\rho(\boldsymbol{v}) = 0$. The first two cases are straightforward. The third case requires some additional investigation. In fact, for $\rho(\boldsymbol{v}) = 0$ we have

$$I_1(\boldsymbol{x} + \boldsymbol{u}_0) - I_0(\boldsymbol{x}) + \langle \nabla I_1, \, \boldsymbol{v} - \boldsymbol{u}_0 \rangle = 0 \ . \tag{21}$$

It is easy to see that this system is undetermined, since one has only one equation for the unknown two-dimensional vector $\boldsymbol{v}$. However, from the first term of (7), we observe that an optimal $\boldsymbol{v}^*$ should also minimize the squared Euclidean distance to $\boldsymbol{u}$. Let us therefore consider the following Lagrangian optimization problem:

$$\boldsymbol{\mathcal{L}} = \min_{\boldsymbol{v}} \left\{ \sum_{d=1}^{2} \int_{\Omega} (u_d - v_d)^2 \, d\Omega + \int_{\Omega} \alpha \, \rho(\boldsymbol{v})^2 \, d\Omega \right\} \ , \tag{22}$$

where α are Lagrange multipliers. This optimization problem aims for finding an optimal $\boldsymbol{v}^*$, which minimizes the squared Euclidean distance to $\boldsymbol{u}$ with the constraint that $\rho(\boldsymbol{v}^*) = 0$. The Lagrange multipliers α are used to enforce this constraint. The Euler Lagrange equations with respect to $\boldsymbol{v}$ and α are given by

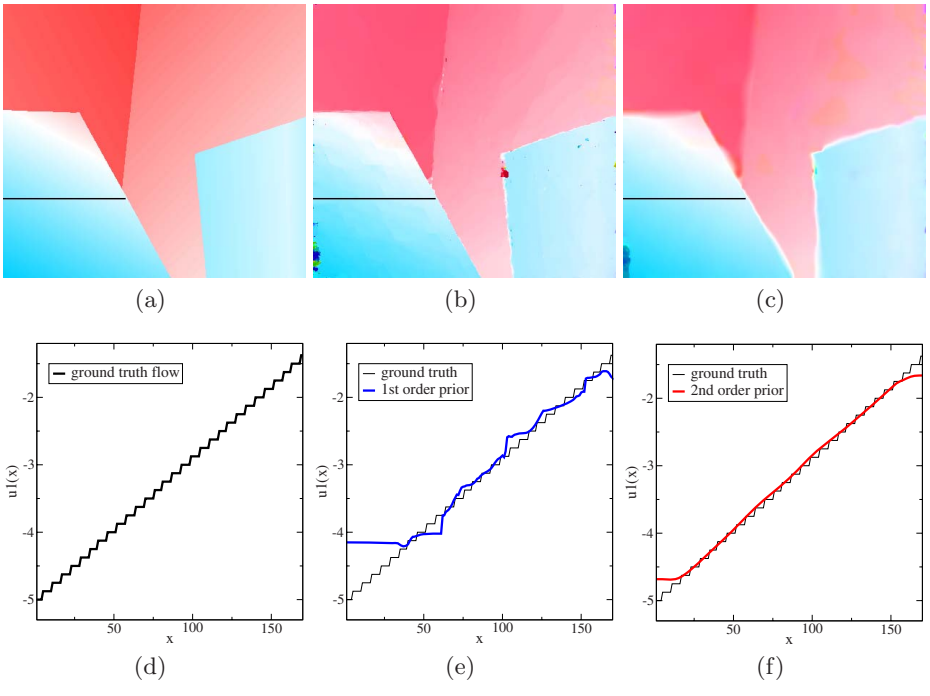
$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{v}} = \boldsymbol{v} - \boldsymbol{u} + \alpha \, \rho(\boldsymbol{v}) \, \nabla I_1 = 0 \qquad \text{and} \qquad \frac{\partial \mathcal{L}}{\partial \alpha} = \rho(\boldsymbol{v})^2 = 0 \ . \tag{23}$$

It is easy to see that this system of equations is well posed. We can therefore solve this system with respect to $\boldsymbol{v}$, resulting in the third case of the thresholding scheme presented in (19). □

4 Results

We evaluated our `C++/CUDA` 1.0 implementation of the proposed model on the Middlebury optical flow database [18], a collection of challenging synthetic and real world image sequences. The test setup was a workstation with an Intel Core 2 CPU at 2.66 GHz (the software is single-threaded, so only one core was used) with an NVidia GeForce 8800 GTX graphics card, running a 32 bit Linux operating system and recent NVidia display drivers.

First, we demonstrate the effects of the new second-order prior by comparing the proposed model to an implementation of the TV-regularized model [8] on the Middlebury "Venus" sequence. This sequence lends itself to such a comparison for four reasons: 1. the ground truth flow is available, 2. large areas of the flow field are perfectly affine, 3. there are several weakly textured areas (emphasizing the prior's influence), and 4. the vertical flow component u_2 is zero (because it is a rectified stereo pair), which simplifies the comparison. Figure 2(a) shows the



Average angle error	avg. rank	Army (Hidden texture) GT	Army im0	Army im1	Mequon (Hidden texture) GT	Mequon im0	Mequon im1	Schefflera (Hidden texture) GT	Schefflera im0	Schefflera im1	Wooden (Hidden texture) GT	Wooden im0	Wooden im1	Grove (Synthetic) GT	Grove im0	Grove im1	Urban (Synthetic) GT	Urban im0	Urban im1	Yosemite (Synthetic) GT	Yosemite im0	Yosemite im1	Teddy (Stereo) GT	Teddy im0	Teddy im1
		all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext
Fusion [8]	2.6	4.43 2	13.7 2	4.08 2	**2.47** 1	**8.91** 1	2.24 2	**3.70** 1	**9.68** 1	3.12 2	**3.68** 1	**19.8** 1	2.54 2	4.06 3	4.82 4	4.14 2	6.30 2	**17.3** 1	5.96 2	4.55 8	5.78 8	3.10 4	7.12 8	13.6 8	7.86 4
Second-order Prior	3.8	**3.84** 1	**11.2** 1	**3.11** 1	3.14 2	13.1 2	**2.18** 1	7.20 8	17.7 2	**2.96** 1	3.86 2	20.9 2	**2.10** 1	4.47 7	5.27 6	**3.83** 1	15.0 9	22.0 7	11.4 9	7.74 10	6.88 9	11.7 10	6.74 2	13.4 2	**5.80** 1
Dynamic MRF [9]	4.0	4.95 9	15.0 3	4.20 3	3.71 8	15.3 4	2.58 3	6.77 2	17.8 3	3.26 3	4.50 3	23.7 3	2.62 3	3.95 2	4.63 3	4.65 6	7.67 4	19.1 5	7.20 5	3.63 5	5.29 5	4.62 7	8.32 5	17.8 7	8.16 5
LP Registration [7]	4.8	7.36 4	16.8 4	6.30 4	3.94 4	13.8 3	3.00 4	7.33 4	17.8 3	4.43 4	5.54 4	24.5 4	3.57 4	4.24 4	4.56 2	5.63 8	10.5 8	21.4 6	9.57 8	4.54 7	5.48 7	3.95 5	8.15 4	17.9 8	7.82 3
Black & Anandan 2 [2]	5.0	7.83 5	18.7 6	6.41 5	9.70 6	21.9 7	8.60 6	13.7 5	23.7 5	18.1 7	10.9 5	30.0 6	9.44 5	4.43 5	5.23 5	4.94 7	7.95 5	18.2 9	6.51 8	2.61 2	4.44 9	2.15 8	8.58 6	14.3 4	8.54 6
2D-CLG [3]	5.0	10.1 8	22.6 10	7.59 6	9.84 7	16.9 5	11.1 7	16.9 7	28.2 9	18.8 8	14.1 9	31.1 8	13.1 8	**3.66** 1	**4.25** 1	4.41 4	6.69 8	22.2 8	6.96 4	**1.76** 1	**3.14** 1	**1.46** 1	**6.29** 1	**12.9** 1	5.81 2

(k)

Fig. 2. First row: color-coded flow fields for the Middlebury "Venus" sequence; (a) ground truth, (b) calculated with TV-L^1 method, and (c) with the proposed method. Second row: (d–f) shows the u_1 component of the flow at the start of line 270 of (a–c), respectively. The corresponding interval in line 270 is marked black in (a–c). Third row: the ground truth flows (g), (i) and the color-coded flows (h), (j) for the Middlebury "Army" (AAE = 3.84°) and "Mequon" (AAE = 3.14°) sequences, respectively. Last row: Average angular error results for the Middlebury benchmark dataset; the proposed method, labeled "Second-order Prior", was ranked 2^{nd} at the time of submission.

color-coded ground truth flow, where an interval at the start of line 270 has been marked black. The u_1 component of the flow in this interval can be seen as a profile plot in Fig. 2(d). Note that the staircasing effect visible in this profile seems to be caused by a coarse quantization – based on the scene, we would expect a straight line here. Figures 2(b) and 2(c) show the color-coded flows for the TV-regularized ($\lambda = 76.5$, $\theta = 0.25$) and the proposed model ($\lambda = 45$, $\theta = 0.25$), respectively. Directly below each solution, Figs. 2(e) and 2(f) allow a close comparison to the ground truth in the selected interval. As intended, we got rid of the bias towards piecewise-constant motion by using the second-order prior.

We also evaluated the proposed model on the Middlebury optical flow benchmark, using two grayscale frames per sequence. To reduce the influence of illumination changes and shadows, we performed a structure-texture decomposition of the images using the well known ROF model [19] ($\lambda_{\text{ROF}} = 10$) and used only the textural part for the motion estimation. The model parameters $\lambda = 55$ and $\theta = 0.1$ were held constant across the entire dataset. Estimating the flow fields for all eight sequences, including preprocessing, took 44 s (5.6 s on a single 584×388 image pair). Figures 2(h) and 2(j) show the resulting flow fields for the "Army" and the "Mequon" sequences, while Fig. 2(g) and 2(i) depict the corresponding ground truth flows. At the time of submission, the method was ranked second out of ten submissions at the Middlebury test site for both the average angular error and the average end-point error; see Fig. 2(k) for the angular error table of the six leading entries. For further images and a comparison on other statistical measures, we encourage the reader to visit `http://vision.middlebury.edu/flow/eval/`. A closer inspection of the table in Fig. 2(k) leads to three observations: first, the proposed method is the top performer in untextured areas for six out of eight sequences, indicating an advantage of the new prior. Second, the subpar performance on the "Grove" and the "Urban" sequences, which both contain large flow magnitudes, can be explained by deficiencies of the optimization approach (multiple scales, repeated warping, and locally linear approximation) and are not an inherent limitation of our new model. Finally, it is clearly apparent that the parameters for the baseline results provided by [18] have been optimized for the synthetic "Yosemite" sequence. This causes overly smooth results on real-world sequences, which considerably skews the comparison. On the other hand, this also explains the bad "Yosemite" results of our implementation.

5 Conclusion

We proposed a novel regularization approach, based on decorrelated second-order derivatives, which does not penalize locally affine motions and therefore avoids a bias towards fronto-parallel flow fields. Subsequently, we derived an efficient numerical scheme for solving this model. By achieving state-of-the-art results on the challenging Middlebury optical flow database, especially in untextured areas, we have clearly shown the applicability of the presented method.

Future work will mainly focus on finding better optimization approaches to improve the performance on image pairs with large flow magnitudes (e.g. the

Middlebury "Urban" sequence). We also plan to investigate more complex data terms and an extension to color images.

References

1. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artificial Intelligence 17, 185–203 (1981)
2. Black, M.J., Anandan, P.: A framework for the robust estimation of optical flow. In: Proc. of the ICCV, May 1993, pp. 231–236 (1993)
3. Nagel, H.H., Enkelmann, W.: An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. IEEE Trans. Pattern Anal. Mach. Intell. 8(5), 565–593 (1986)
4. Weickert, J., Brox, T.: Diffusion and regularization of vector- and matrix-valued images. Inverse Problems, Image Analysis, and Medical Imaging. Contemporary Mathematics 313, 251–268 (2002)
5. Aubert, G., Deriche, R., Kornprobst, P.: Computing optical flow via variational techniques. SIAM Journal on Applied Mathematics 60(1), 156–182 (2000)
6. Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J.: Highly accurate optic flow computation with theoretically justified warping. International Journal of Computer Vision 67(2), 141–158 (2006)
7. 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), 257–277 (2006)
8. Zach, C., Pock, T., Bischof, H.: A duality based approach for realtime TV-L1 optical flow. In: Hamprecht, F.A., Schnörr, C., Jähne, B. (eds.) DAGM 2007. LNCS, vol. 4713, pp. 214–223. Springer, Heidelberg (2007)
9. Ishikawa, H.: Total absolute gaussian curvature for stereo prior. In: Yagi, Y., Kang, S.B., Kweon, I.S., Zha, H. (eds.) ACCV 2007, Part I. LNCS, vol. 4843, pp. 537–548. Springer, Heidelberg (2007)
10. Yuan, J., Schörr, C., Steidl, G.: Simultaneous higher-order optical flow estimation and decomposition. SIAM J. on Scientific Computing 29(6), 2283–2304 (2007)
11. Ju, S., Black, M.J., Jepson, A.: Skin and bones: Multi-layer, locally affine, optical flow and regularization with transparency. In: Proc. of the CVPR, pp. 307–314 (1996)
12. Bereziat, D., Herlin, I.: Object based optical flow estimation with an affine prior model. In: Proc. of the CVPR, vol. 3, pp. 1048–1051 (2000)
13. Le, H.V., Seetharaman, G., Zavidovique, B.: A multiscale technique for optical flow computation using piecewise affine approximation. In: Proc. of the Int. Workshop on Computer Architectures for Machine Perception, pp. 221–230 (2003)
14. Nir, T., Bruckstein, A.M., Kimmel, R.: Over-parameterized variational optical flow. International Journal of Computer Vision 76, 205–216 (2008)
15. Danielsson, P.E., Lin, Q.: Efficient detection of second-degree variations in 2D and 3D images. Journal of Visual Comm. and Image Representation 12, 255–305 (2001)
16. Carter, J.: Dual Methods for Total Variation-based Image Restoration. PhD thesis, UCLA, Los Angeles, CA (2001)
17. Chambolle, A.: Total variation minimization and a class of binary MRF models. In: Energy Minimization Methods in Comp. Vision and Pattern Rec., pp. 136–152 (2005)
18. Baker, S., Scharstein, D., Lewis, J.P., Roth, S., Black, M., Szeliski, R.: A database and evaluation methodology for optical flow. In: Proc. of the ICCV (2007)
19. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D 60, 259–268 (1992)