Structure- and Motion-adaptive Regularization for High Accuracy Optic Flow

Andreas Wedel[1,2], Daniel Cremers[2], Thomas Pock[3], and Horst Bischof[3]

[1] Daimler Group Research [2] Bonn University* [3] Graz University of Technology†,‡

Abstract

The accurate estimation of motion in image sequences is of central importance to numerous computer vision applications. Most competitive algorithms compute flow fields by minimizing an energy made of a data and a regularity term. To date, the best performing methods rely on rather simple purely geometric regularizers favoring smooth motion. In this paper, we revisit regularization and show that appropriate adaptive regularization substantially improves the accuracy of estimated motion fields. In particular, we systematically evaluate regularizers which adaptively favor rigid body motion (if supported by the image data) and motion field discontinuities that coincide with discontinuities of the image structure. The proposed algorithm relies on sequential convex optimization, is real-time capable and outperforms all previously published algorithms by more than one average rank on the Middlebury optic flow benchmark.

1. Introduction

Estimating correspondences between pairs of points in either of two images remains one of the fundamental computational challenges in Computer Vision. Different variants of this problem arise in the estimation of motion in videos [9], the nonrigid registration of medical structures observed in different modalities [8], and the tracking of deformable objects [5]. Computationally the estimation of correspondences based on matching points of similar intensity is a classical ill-posed problem in the sense that merely imposing matching of similar intensities will typically not give rise to a unique solution.

To make the optic flow estimation well-posed researchers have reverted to *regularization*. In 1981 Horn and Schunck proposed what is typically considered the first variational method in Computer Vision. In order to compute a dense motion field $v : \Omega \to \mathbb{R}^2$ on the image plane $\Omega \subset \mathbb{R}^2$ for matching a pair of consecutive images from a gray value sequence $I : \Omega \times [0, T] \to \mathbb{R}$, they proposed to minimize the functional

$$E(v) = \int_\Omega \left(\nabla I^\top v + I_t\right)^2 + \lambda \left(|\nabla v_1|^2 + |\nabla v_2|^2\right) d^2x\,. \quad (1)$$

The data term aims at matching points of similar intensity by imposing the linearized brightness constancy constraint, while the regularity term (weighted by $\lambda > 0$) imposes spatial smoothness of the velocity field $v = [v_1, v_2]^\top$.

To date the work of Horn and Schunck has attracted 3900 citations, many of these dealing with applications of motion estimation in different scenarios, many suggesting alternative cost functionals, and many investigating alternative minimization strategies. A recently established optic flow benchmark covering a variety of different motion estimation problems allows to assess which variants of functionals and algorithms provide the most accurate motion estimates [3]. A glance at the current results indicates that the currently top performing methods typically minimize functionals of the form

$$E(v) = \int_\Omega \rho(v, x) + \lambda\, \psi(v, \nabla v, \dots)\, d^2x, \quad (2)$$

with robust L_1-penalized data fidelity term

$$\rho(v, x) = |I_1(x) - I_2(x + v(x))|, \quad (3)$$

and a discontinuity-preserving L_1-smoothness term

$$\psi(\nabla v) = |\nabla v_1| + |\nabla v_2|. \quad (4)$$

This regularity term is surprisingly simple, favoring flow fields which are spatially smooth. In contrast to the original L_2-regularity suggested by Horn and Schunck, the L_1-regularity is known to better preserve discontinuities [4, 7, 11, 14, 13]. In recent years, researchers have suggested far more sophisticated regularization techniques based on statistical learning [15]. So far these have not been able to outperform the more naive approaches. Of course it

*Supported by the German Research Foundation grant CR-250/2-1 and the Hausdorff Center for Mathematics.

†Supported by the Austrian Research Promotion Agency within the VM-GPU project (no. 813396) and the Austrian Science Fund (FWF) under the doctoral program Confluence of Vision and Graphics W1209.

‡We greatly acknowledge Nvidia for providing us a Tesla C1060 GPU.

is hard to say why this is the case, one reason may be that the challenge of learning "typical" flow patterns may not be feasible, given that different image structures, unknown object deformations, and camera motions may give rise to a multitude of motion patterns with little resemblance between motion fields from different videos. Nevertheless, appropriate regularization is of utmost importance for optic flow estimation, since it stabilizes the otherwise ill-posed optical flow problem and induces a filling-in in areas of low contrast or texture.

In this paper, we revisit regularization of optical flow fields and suggest adaptive regularization strategies which are not motivated by the paradigm of learning from examples but rather by a closer analysis of the inherent structure of motion in videos:

- **Adaptive rigid-body motion regularization** A large number of observed motions correspond to 3D rigid body motion. This is true for static scenes filmed by a moving camera. And it is also true for rigid objects moving in a scene. While such regularizers have recently been suggested in [16, 17], these techniques fail to improve the overall performance because they degrade motion estimates in areas which are not consistent with the rigid body motion. In this paper, we will therefore introduce adaptive regularization techniques which favor rigid body motion *only* if this is supported by the image data.

- **Structure-dependent Regularization** In many real-world image sequences, motion fields and intensity information are not independent. Commonly discontinuities in the motion fields arise at boundaries of objects where the brightness function is also likely to change. Although this is certainly not always true, making motion regularizers depend on local brightness variations can energetically favor motion discontinuities to coincide with brightness discontinuities and is likely to improve optic flow estimation. While such adaptive regularizers have been proposed before [1, 12, 19], the contribution of this paper is to quantitatively evaluate the improvement in performance that this adaptation brings about.

In the following we will introduce a novel optical flow functional which combines a robust data term with two regularity terms adaptively favoring rigid body motion and structure-dependent smoothness. Subsequently we will show how the proposed functional can be efficiently minimized by sequential convex optimization. In subsequent experiments we carefully assess the contribution of each regularizer and show the improvement in accuracy. The overall functional is fairly transparent, can be minimized quite efficiently and is the currently best performing algorithm for optic flow estimation.

2. Adaptive Optic Flow Regularization

Given a pair of consecutive input images $I \equiv (I_1, I_2)$ with $I_i : (\Omega \subset \mathbb{R}^2) \to \mathbb{R}$, we propose to compute a motion field $v : \Omega \to \mathbb{R}^2$ by minimizing the energy:

$$E(v) := E_{data}(v) + E_{rigid}(\mathbf{F}, v) + E_{struct}(v)\,. \qquad (5)$$

Here, $E_{data}(v) = \int_\Omega |I_1(x) - I_2(x+v)|\, d^2x$ is the well-known brightness constancy constraint. In order to allow for illumination changes between the images, we use the structure-texture decomposition approach presented in [18] which essentially removes low-frequency components. The two adaptive regularizers E_{rigid} and E_{struct} will be detailed in the following.

2.1. Adaptive Rigid Motion Regularization

Within the spectrum of conceivable optic flow patterns in Computer Vision, flow patterns that correspond to 3D rigid body motion play a central role. This is not surprising, since they invariably arise for static scenes filmed by a moving camera or for objects moving rigidly. It is well known that in the case of rigid body motion the two-dimensional optic flow estimation problem is reduced to a one-dimensional search along the epipolar lines which can actually be solved quite efficiently. The challenge is, that the epipolar lines are usually unknown and need to be estimated from established point correspondences themselves. This bootstrapping problem is usually solved iteratively and has one major drawback: If the scene is not stationary, both the epipolar lines and the optical flow suffer from the negative prior inflicted by the other.

In the following, we will therefore propose an adaptive regularization which favors rigid body motion only if this is supported by the image data. In particular we propose to adaptively favor rigid body motion using the regularizer

$$E_{rigid}(\mathbf{F}, v) = \gamma(v) \int_\Omega \rho_{\mathbf{F}}(v, x)\, d^2x \qquad (6)$$

where $\rho_{\mathbf{F}}(v, x)$ is the symmetric distance of the flow vector to the epipolar lines (see below). The adaptive weighting $\gamma(v)$ aims at engaging the rigid body energy based on the amount of independent motion found within the scene. For the computation of $\gamma(v)$, the motion field v is fixed and γ is a single global value. It is given by

$$\gamma(v) = \begin{cases} \lambda_{\mathbf{F}} & \text{if } \int_\Omega \rho_{\mathbf{F}}(v, x)/\|v\|\, d^2x < \delta_{\mathrm{F}} \\ 0 & \text{otherwise}\,. \end{cases} \qquad (7)$$

With the 3×3 fundamental matrix $\mathbf{F}$ the symmetric distance of the flow vector to the epipolar lines,

$$\tilde{x} = \mathbf{F}^\top \begin{bmatrix} x \\ 1 \end{bmatrix} \text{ and } \tilde{v} = \mathbf{F} \begin{bmatrix} x+v \\ 1 \end{bmatrix},$$

is defined as

$$\rho_{\mathbf{F}}(v,x) = \frac{1}{\tilde{x}_1^2 + \tilde{x}_2^2 + \tilde{v}_1^2 + \tilde{v}_2^2} \left| c + \tilde{x}_1 v_1 + \tilde{x}_2 v_2 \right| , \quad (8)$$

where a sub-index i of a vector denotes its i-th component and the constant c is computed as

$$c = \begin{bmatrix} x \\ 1 \end{bmatrix}^\top \mathbf{F} \begin{bmatrix} x \\ 1 \end{bmatrix} . \quad (9)$$

This formulation is symmetric w. r. t. the two input images and normalized in the sense that it does not depend on the scale factor used to compute $\mathbf{F}$ [10].

To this end, we start with computing an optic flow field and subsequently estimate the fundamental matrix by minimizing the non-linear criterion

$$\min_{\mathbf{F}} \left\{ \sum_{\Omega} \frac{1}{\tilde{x}_1^2 + \tilde{x}_2^2 + \tilde{v}_1^2 + \tilde{v}_2^2} \left(\begin{bmatrix} x \\ 1 \end{bmatrix}^\top \tilde{v} \right)^2 \right\} . \quad (10)$$

Then, the estimated fundamental matrix itself is used to drive the optical flow toward the epipolar lines.

The crucial part is how to weight this data term in order to maintain robustness in dynamic scenes and increase accuracy in static scenes. This is where the adaptive weighting $\gamma(v)$, which analyzes the amount of independent motion, becomes important. Essentially, a violation of the epipolar constraint denotes other moving objects in the scene while the counter-hypothesis does not generally hold (e. g. the motion of objects might coincide with the epipolar lines, see Figure 2). However, simply computing the average symmetric distance to the epipolar lines, $\int_\Omega \rho_{\mathbf{F}}(v,x)\, d^2x$, does not yield useful results as flow vectors in a dynamic scene might be relatively small in magnitude, yielding only small errors although the complete scene is dynamic. Therefore we propose to evaluate the relative symmetric distance, weighted by the inverse length of the computed optical flow. Such measure yields a rather robust estimate of the relative motion contained in the scene depicted by the two images.

In summary, the adaptive rigid body regularization has the following effect:

- If the average relative deviation $\int_\Omega \rho_{\mathbf{F}}(v,x)/\|v\|\, d^2x$ is above a predefined threshold δ_{F}, the fundamental matrix regularization will be switched off so it does not bias the estimation of motion in dynamic scenes.
- If on the other hand the relative deviation is smaller than δ_{F}, then the fundamental matrix regularization is imposed so as to favor the estimation of optic flow fields that are consistent with rigid body motion. Note that even in this case we do *not* enforce the fundamental matrix constraint to be exactly fulfilled. Instead, we allow for deviations of the optic flow from rigid body motion wherever this is supported by the image data.

2.2. Structure-Adaptive Regularization

For many real-world videos, discontinuities of the motion field tend to coincide with object boundaries and discontinuities of the brightness function. Although this is certainly not always true, our experiments on the optic flow benchmark [3] will demonstrate quantitatively that the introduction of brightness-adaptive smoothness constraints – which are not considered in the currently top-performing algorithms – leads to substantial improvements of optic flow estimates.

An elegant theoretical treatise of image-adaptive regularization of flow fields was presented in [19]. There, the authors introduce regularizers of the form

$$\Psi\left(\nabla v_1^\top D(\nabla I)\, \nabla v_1\right) + \Psi\left(\nabla v_2^\top D(\nabla I)\, \nabla v_2\right) , \quad (11)$$

corresponding to an inhomogeneous and potentially anisotropic regularization induced by a structure-dependent tensor $D(\nabla I)$. The central idea is that the smoothness of v along the two eigenvectors of D is weighted by the corresponding eigenvalues. In fact, anisotropic structure-dependent regularization was already proposed by Nagel in 1983 [12]. This is achieved by setting

$$D(\nabla I) = \frac{1}{|\nabla I|^2 + 2\lambda} \left(\nabla I^\perp \nabla I^{\top\perp} + \lambda^2\, \mathrm{Id} \right)$$

where Id denotes the unit matrix and $\nabla I^\perp$ is the vector perpendicular to ∇I. This leads to an anisotropic smoothing of v along the level lines of the image intensity while preserving discontinuities across level lines.

For the sake of simplicity and in order to facilitate fast implementations, we will in the following only consider inhomogeneous isotropic regularization. Following [1] we set $D(\nabla I) = g(|\nabla I|)\, \mathrm{Id}$ with a strictly decreasing positive function

$$g(|\nabla I|) = \exp\left(-\alpha |\nabla I|^\beta\right) \quad (12)$$

favoring discontinuities of the motion field to arise at locations of strong image gradient. Nevertheless, our framework can easily incorporate anisotropic regularity terms such as those discussed above. The structure-dependent regularizer in (5) is therefore given by

$$E_{struct}(v) = \lambda \int_\Omega \exp\left(-\alpha |\nabla I|^\beta\right) \left(|\nabla v_1| + |\nabla v_2|\right)\, d^2x\, . \quad (13)$$

3. Minimization by Quadratic Relaxation

The functional (5) introduced in the previous section is non-convex – as all state-of-the-art optic flow functionals. Moreover, its dependency on v is rather involved: the velocity v appears inside the arguments of the input images in the data term and in the regularity terms. In particular, the

weight $\gamma(v)$ of the rigid motion regularizer in (7). As a consequence, the quality of computed solutions will invariably depend on a careful choice of the minimization scheme.

Rather than reverting to the traditional scheme of linearization and fixed-point iteration, we follow a series of papers on quadratic relaxation [6, 2, 20] which leads to a decomposition of the original non-convex optimization problem into a sequence of convex optimization problems. More specifically, we reformulate the optimization of (5) with respect to v as a minimization of the functional

$$\begin{aligned} E(v,u) =& E_{data}(u) + E_{rigid}(\mathbf{F},u) \\ &+ \int_\Omega \frac{1}{2\theta}(u-v)^2\, d^2x + E_{struct}(v)\,. \end{aligned} \tag{14}$$

with respect to both the velocity field v and an auxiliary field $u : \Omega \to \mathbb{R}^2$. The central idea is that for $\theta \to 0$, the two functionals become identical. Yet, the optimization with respect to v and u can each be solved optimally.

Specifically, we initialize with $u = 0$, $v = 0$, $\tilde{v} = 0$, $\gamma(v) = 0$, $F = 0$, and iterate until convergence:

(A) Use the thresholding scheme for convex quadratic programming presented in [17] to solve

$$\min_u \left\{ E_{data}(u,I) + E_{rigid}(\mathbf{F},u) + \int_\Omega \frac{1}{2\theta}(u-v)^2\, d^2x \right\}.$$

(B) Solve the convex optimization problem

$$\min_v \left\{ \int_\Omega \frac{1}{2\theta}(u-v)^2\, d^2x + E_{struct}(v,I) \right\}$$

by means of a primal-dual algorithm as in [17, 20]. Essentially this step of the algorithm amounts the well-known weighted total variation noise removal.

(C) Estimate the fundamental matrix parameters, F, using the non-linear criterion as done in [10]. Recompute $\tilde{v}$ and $\gamma(v)$ and start again with (A).

The next section carefully investigates the accuracy improvement when using additional prior knowledge in terms of the adaptive rigid motion prior and the structure-dependent regularization for optic flow estimation.

4. Experiments

In this section, we will quantitatively evaluate the contribution of each of the two regularizers on the accuracy of estimated flow fields based on the Middlebury optic flow benchmark [3]. The benchmark provides a training data set where the ground truth optic flow is known and an evaluation set used to compare algorithms with each other.

4.1. Contribution of Each Regularizer

Table 1 demonstrates the increase in accuracy on the training data set. The table shows the average end-point error between the ground truth flow vectors and the estimated flow vectors for all data sets in the training set and for different choices of regularizers, the traditional smoothness constraint corresponding to the first line. Evidently the non-adaptive rigid motion prior increases accuracy in static scenes but worsens the results if the scene is dynamic. The structure-aware regularization does improve the optic flow accuracy on most test examples but only the combined adaptive approach yields top performing results.

4.2. Overall Performance

Figure 3 shows screen shots of the Middlebury evaluation homepage (taken on March 10th 2009). It demonstrates that the proposed combination of structure- and motion-adaptive regularizers gives rise to an optic flow algorithm outperforming all existing algorithms, both with respect to the angle error and with respect to the end-point error.

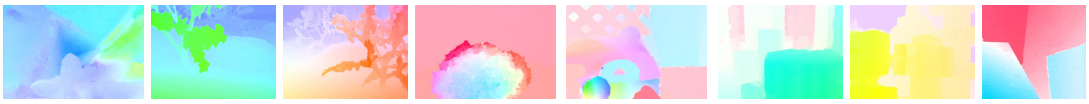
Table 2 shows the increase in computation time induced by each of the two regularizers. It indicates that using the proposed adaptive regularizer increases the computation time by about a factor of 2. Although currently only the baseline algorithm is implemented on the GPU, we believe that extending the proposed regularizers to the GPU will give rise to a similar speed-up factor of about 80.

For all experiments we use fix parameter values of $\lambda = 30$, $\lambda_{\mathrm{F}} = 0.75$, and $\theta = 0.25$. We employ a multi-resolution pyramid with a factor of 0.5, initializing at each level with the solutions obtained on the lower resolution. The threshold for the relative fundamental matrix deviation is set to $\delta_{\mathbf{F}} = 0.05$. The same parameter settings are applied for the experiments on 12-bit image data.

4.3. Real-World Experiments

We applied the proposed optic flow algorithm to challenging real-world videos captured from a driving car. Figure 1 shows the first scene with an approaching car and a moving pedestrian. Although the scene is dynamic, the fundamental matrix data term is used for optic flow estimation. This is due to the fact that the approaching car moves along the epipolar lines, only slightly violating the fundamental matrix constraint. The optic flow for the approaching car is accurately estimated with flow discontinuities at image edges. Note that the walking person in the distance is detected as violating the epipolar constraint. Such violations are allowed because we do *not* require the fundamental matrix constraint to be exactly fulfilled.

Figure 2 depicts an example of a person running across the street. The fundamental matrix regularity is automatically deactivated due to the increased presence of independent motion in the scene. The optical flow result is quite



Measure	$\rho_{\mathbf{F}}$	$g(\lvert\nabla I\rvert)$	Dimetrodon	Grove2	Grove3	Hydrangea	RubberWhale	Urban2	Urban3	Venus
EPE	–	–	0.19	0.15	0.67	0.15	0.09	0.32	0.63	0.26
EPE	×	–	0.28	0.15	0.65	0.15	0.09	**0.29**	0.49	0.26
EPE	–	×	**0.19**	0.15	0.58	0.15	**0.08**	0.32	0.60	0.26
EPE	×	×	**0.19**	**0.14**	**0.56**	0.15	**0.08**	**0.29**	**0.45**	**0.25**
rel-$\rho_{\mathbf{F}}$	×	×	0.11	0.01	0.02	0.21	0.29	0.01	0.01	0.01

Table 1. Evaluation results on the Middlebury training data. The proposed regularizers, $\rho_{\mathbf{F}}$ and $g(|\nabla I|)$, systematically improve the optic flow estimates. In the table, EPE denotes the average end-point error of the obtained optic flow field and rel-$\rho_{\mathbf{F}} = \int_\Omega \rho_{\mathbf{F}}(v,x)/\|v\|\, d^2x$ is the average relative epipolar line distance. See Table 2 for execution times on the training data set.

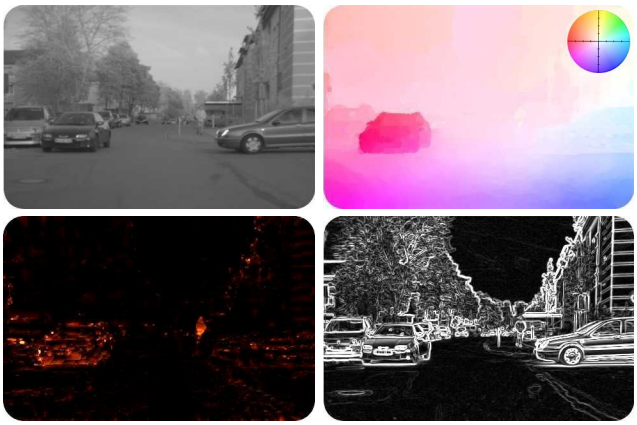


Figure 1. Optical flow estimation in a dynamic scene. The lower images show the $\rho_{\mathbf{F}}$ constraint deviation and the edge image $|\nabla I|$. Note the accurate flow estimation for the approaching car and the detection of independent motion for the distant pedestrian.

Figure 2. Optical flow estimation for a scene with a running person. Even the large displacement of the right foot is correctly matched. The lower images show the $\rho_{\mathbf{F}}$ constraint deviations which clearly identify the person and the edge image $|\nabla I|$.

convincing, preserving discontinuities along object boundaries. Though the proposed adaptive regularizers give rise to highly convincing flow fields, there is certainly still room for improvements: While the algorithm does capture the large motion of the right foot, it incorrectly estimates a false match to the right of this foot which is due to occlusions that are not explicitly modeled in our approach.

5. Conclusion

In this work, we revisited the aspect of regularization in optic flow estimation. Specifically, we presented motion- and structure-adaptive regularizers: The first one favors optic flow fields which are consistent with rigid body motion in scenes wherever this is supported by the data. The second one favors motion discontinuities to coincide with discontinuities in the intensity function. We experimentally evaluated the proposed algorithm in a three-fold manner: Firstly, we quantify the improvement in performance induced by each regularizer on a training set of Middlebury benchmark sequences. Secondly, we demonstrate on the Middlebury test set, that the proposed method outperforms all existing algorithms, both in angle error and in end-point error. And thirdly, we demonstrate that the algorithm provides robust performance on various real-world sequences filmed from a moving vehicle.

References

[1] L. Alvarez, J. Esclarín, M. Lefébure, and J. Sánchez. A PDE model for computing the optical flow. In *Proc. XVI Congreso de Ecuaciones Diferenciales y Aplicaciones*, pages 1349–1356, Gran Canaria, Spain, September 1999. 2, 3

[2] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition - modeling, algorithms, and parameter selection. *Int. J. Comp. Vis.*, 67(1):111–136, 2006. 4

[3] S. Baker, S. Roth, D. Scharstein, M. Black, J. Lewis, and R. Szeliski. A database and evaluation methodology for optical flow. In *Int. Conf. Comp. Vis.*, pages 1–8, 2007. 1, 3, 4

[4] M. J. Black and P. Anandan. A framework for the robust estimation of optical flow. In *Int. Conf. Comp. Vis.*, pages 231–236, 1993. 1

[5] T. Brox, B. Rosenhahn, D. Cremers, and H.-P. Seidel. High accuracy optical flow serves 3-D pose tracking: exploiting

Processor Type	Warps × Iter.	$\rho_{\mathbf{F}}$	$g(\|\nabla I\|)$	[420 × 380] Venus	[584 × 380] Dimetrodon Hydrangea RubberWhale	[640 × 480] Grove{2,3} Urban{2,3}
NVidia® GeForce® GTX 285	35 × 5	–	–	0.0549	0.068	0.077
Intel® Core™2 Extreme 2.8GHz	35 × 5	–	–	4.24	5.29	8.15
Intel® Core™2 Extreme 2.8GHz	35 × 5	×	×	8.02	10.9	16.1

Table 2. Execution time in seconds for optic flow computation on the Middlebury training data (per frame). The table shows that the computational time roughly doubles when employing a fundamental matrix data term and a gradient-weighted regularization. For the baseline algorithm we included the computational time on a modern GPU, yielding a speedup factor of 80.

Average angle error	avg. rank	**Army** (Hidden texture) all	disc	untext	**Mequon** (Hidden texture) all	disc	untext	**Schefflera** (Hidden texture) all	disc	untext	**Wooden** (Hidden texture) all	disc	untext	**Grove** (Synthetic) all	disc	untext	**Urban** (Synthetic) all	disc	untext	**Yosemite** (Synthetic) all	disc	untext	**Teddy** (Stereo) all	disc	untext
Adaptive [24]	3.5	**3.29** 1	**9.43** 1	**2.28** 1	3.10 4	11.4 5	2.46 5	6.58 7	15.7 6	2.52 3	**3.14** 1	**15.6** 1	**1.56** 1	3.67 4	4.46 3	3.48 5	**3.32** 1	**13.0** 1	**2.38** 1	2.76 10	4.39 9	1.93 7	3.58 2	8.18 2	2.88 2
Spatially variant [22]	4.8	3.73 3	10.2 4	3.33 4	3.02 3	11.0 4	2.67 6	5.36 3	13.8 4	**2.35** 1	3.67 2	19.3 5	1.84 3	3.81 6	4.81 11	3.69 8	4.48 4	16.0 5	3.90 4	2.11 4	3.26 2	2.12 9	4.66 6	9.41 6	4.35 8
TV-L1-improved [20]	5.5	3.36 2	9.63 2	2.62 2	2.82 2	10.7 3	2.23 2	6.50 6	15.8 7	2.73 4	3.80 5	21.3 9	1.76 2	**3.34** 1	4.38 2	**2.39** 1	5.97 6	18.1 11	5.67 9	3.57 13	4.92 14	3.43 15	4.01 5	9.84 7	3.44 3
F-TV-L1 [18]	7.0	5.44 10	12.5 9	5.69 14	5.46 11	15.0 11	4.03 11	7.48 11	16.3 9	3.42 9	5.08 10	23.3 13	2.81 8	3.42 2	**4.34** 1	3.03 2	4.05 3	15.1 3	3.18 2	2.43 7	3.92 7	1.87 6	3.90 4	9.35 5	**2.61** 1
Brox et al. [8]	8.6	4.80 9	14.4 13	4.29 11	4.05 8	13.5 8	3.71 9	6.63 8	16.0 8	7.26 11	5.22 11	22.7 12	3.22 11	4.56 15	6.09 20	3.40 3	3.97 2	17.9 9	3.41 3	2.07 3	3.76 5	1.18 2	5.14 8	11.9 10	4.28 7

Average end-point error	avg. rank	**Army** (Hidden texture) all	disc	untext	**Mequon** (Hidden texture) all	disc	untext	**Schefflera** (Hidden texture) all	disc	untext	**Wooden** (Hidden texture) all	disc	untext	**Grove** (Synthetic) all	disc	untext	**Urban** (Synthetic) all	disc	untext	**Yosemite** (Synthetic) all	disc	untext	**Teddy** (Stereo) all	disc	untext
Adaptive [24]	3.5	**0.09** 1	**0.26** 1	**0.06** 1	0.23 6	0.78 4	0.18 5	0.54 8	1.19 9	0.21 3	**0.18** 1	0.91 2	**0.10** 1	0.88 2	1.25 2	0.73 3	**0.50** 1	**1.28** 1	**0.31** 1	0.14 9	0.16 9	0.22 9	**0.65** 1	1.37 2	0.79 3
Spatially variant [22]	5.0	0.10 3	0.27 3	0.08 3	0.22 3	0.75 3	0.19 6	0.43 3	1.00 4	**0.18** 1	0.19 2	1.05 4	**0.10** 1	1.05 7	1.41 8	1.16 11	0.59 4	1.61 4	0.43 3	0.13 5	**0.11** 1	0.28 14	0.96 8	1.72 7	1.28 11
TV-L1-improved [20]	5.6	**0.09** 1	**0.26** 1	0.07 2	0.20 2	0.71 2	0.16 2	0.53 7	1.18 8	0.22 4	0.21 5	1.24 10	0.11 3	0.90 3	1.31 4	0.72 2	1.51 10	1.93 8	0.84 8	0.18 14	0.17 10	0.31 15	0.73 5	1.62 5	0.87 4
F-TV-L1 [18]	7.1	0.14 10	0.35 10	0.14 13	0.34 10	0.98 9	0.26 9	0.59 12	1.19 9	0.26 8	0.27 10	1.36 13	0.16 8	0.90 3	1.30 3	0.76 5	0.54 2	1.62 5	0.36 2	0.13 5	0.15 8	0.20 8	0.68 3	1.56 4	**0.66** 1
DPOF [21]	7.2	0.15 11	0.30 5	0.11 9	0.34 10	1.01 11	0.25 8	**0.29** 1	**0.59** 1	0.26 8	0.26 8	0.94 3	0.20 11	**0.80** 1	**1.13** 1	**0.63** 1	0.90 5	1.85 7	0.66 6	0.27 20	0.22 19	0.54 19	**0.65** 1	**1.20** 1	0.99 5
Fusion [9]	7.6	0.11 6	0.34 8	0.10 7	**0.19** 1	**0.69** 1	0.16 2	**0.29** 1	0.66 2	0.23 5	0.20 3	1.19 8	0.14 6	1.07 8	1.42 9	1.22 12	1.35 7	1.49 3	0.86 10	0.20 15	0.20 17	0.26 11	1.07 12	2.07 14	1.39 14
Brox et al. [8]	8.4	0.12 8	0.37 12	0.11 9	0.31 9	0.97 8	0.28 10	0.48 6	1.11 6	0.48 11	0.28 11	1.28 11	0.18 10	1.13 12	1.57 16	1.11 9	1.02 6	2.02 9	0.60 5	0.10 2	0.13 4	**0.11** 1	0.93 7	2.00 12	1.07 7
Dynamic MRF [10]	9.2	0.12 8	0.34 8	0.11 9	0.22 3	0.89 7	0.16 2	0.44 4	1.13 7	0.20 2	0.24 7	1.29 12	0.14 6	1.11 10	1.52 13	1.13 10	1.54 11	2.37 19	0.93 11	0.13 5	0.12 3	0.31 15	1.27 16	2.33 18	1.66 15
CBF [15]	9.4	0.10 3	0.28 4	0.09 5	0.29 7	0.79 5	0.29 11	0.45 5	0.98 3	0.24 7	0.21 5	1.22 9	0.13 5	0.96 5	1.39 7	0.74 4	2.32 21	2.19 13	1.29 19	0.34 22	0.27 23	0.85 22	0.86 6	1.78 8	1.06 6
SegOF [13]	9.6	0.15 11	0.36 11	0.10 7	0.57 13	1.16 13	0.59 18	0.68 13	1.24 11	0.64 12	0.32 13	**0.86** 1	0.26 13	1.18 15	1.50 12	1.47 16	1.63 14	2.09 10	0.96 12	**0.08** 1	0.13 4	0.12 2	0.70 4	1.50 3	0.69 2
Second-order prior [11]	10.5	0.10 3	0.30 5	0.08 3	0.22 3	0.85 6	**0.15** 1	0.57 11	1.28 13	0.23 5	0.20 3	1.14 6	0.11 3	1.13 12	1.55 15	1.03 8	2.52 22	2.45 20	1.25 18	0.42 23	0.25 22	1.09 24	0.98 10	1.92 9	1.07 7
Learning Flow [14]	11.5	0.11 6	0.32 7	0.09 5	0.29 7	0.99 10	0.23 7	0.55 9	1.24 11	0.29 10	0.36 14	1.56 15	0.25 12	1.25 17	1.64 18	1.41 15	1.55 13	2.32 18	0.85 9	0.14 9	0.18 14	0.24 10	1.09 13	2.09 16	1.27 10
Filter Flow [23]	11.8	0.17 14	0.39 14	0.13 12	0.43 12	1.09 12	0.38 12	0.75 14	1.34 14	0.78 17	0.70 17	1.54 14	0.68 17	1.13 12	1.38 6	1.51 17	0.57 3	1.32 2	0.44 4	0.22 17	0.23 20	0.26 11	0.96 8	1.66 6	1.12 9

Figure 3. Screen shot from the Middlebury optic flow benchmark (as on April 5th 2009). The proposed method, *Adaptive*, is ranked first place in both accuracy measures, the *angle error* and the *end-point error* category.

contour and flow based constraints. In *Europ. Conf. on Comp. Vis.*, LNCS, pages 98–111. Springer, 2006. 1

[6] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Im. Vis.*, 20(1-2):89–97, 2004. 4

[7] R. Deriche, P. Kornprobst, and G. Aubert. Optical flow estimation while preserving its discontinuities: A variational approach. In *Asian Conf. on Computer Vision*, volume 2, pages 290–295, Singapore, 1995. 1

[8] O. F. G. Hermosillo, C. Chef d'Hotel. Variational methods for multimodal image matching. *Int. J. Comp. Vis.*, 50(3):329–343, 2002. 1

[9] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981. 1

[10] Q. Luong and O. Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *Int. J. Comp. Vis.*, 17(1):43–75, 1996. 3, 4

[11] H. Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Trans. on Patt. Anal. and Mach. Intell.*, 8(5):565–593, 1986. 1

[12] H. H. Nagel. Constraints for the estimation of displacement vector fields from image sequences. In *Proc. Eighth Int. Conf. Artif. Intell.*, pages 945–951, Karlsruhe, Germany, 1983. 2, 3

[13] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert. Highly accurate optic flow computation with theoretically justified warping. *International Journal of Computer Vision*, 67(2):141–158, April 2006. 1

[14] C. Schnörr. Segmentation of visual motion by minimizing convex non-quadratic functionals. In *12th Int. Conf. on Pattern Recognition*, Jerusalem, Israel, October 1994. 1

[15] D. Sun, S. Roth, J. P. Lewis, and M. J. Black. Learning optical flow. In *Europ. Conf. on Comp. Vis.*, volume 3, pages 83–91, 2008. 1

[16] L. Valgaerts, A. Bruhn, and J. Weickert. A variational approach for the joint recovery of the optical flow and the fundamental matrix. In *Pattern Recognition (Proc. DAGM)*, pages 314–324, Munich, Germany, June 2008. 2

[17] A. Wedel, T. Pock, J. Braun, U. Franke, and D. Cremers. Duality TV-L1 flow with fundamental matrix prior. In *Proc. Image and Vision Computing New Zealand*, Christchurch, New Zealand, November 2008. 2, 4

[18] A. Wedel, T. Pock, C. Zach, D. Cremers, and H. Bischof. An improved algorithm for TV-L1 optical flow. In *Proc. of the Dagstuhl Motion Workshop*, LNCS, pages 23–45, 2008. 2

[19] J. Weickert and C. Schnörr. A theoretical framework for convex regularizers in PDE–based computation of image motion. *Int. J. Comp. Vis.*, 45(3):245–264, 2001. 2, 3

[20] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime TV-L1 optical flow. In *Pattern Recognition (Proc. DAGM)*, LNCS, pages 214–223, Heidelberg, Germany, 2007. Springer. 4