Over-parameterized Optical Flow using a Stereoscopic Constraint

Guy Rosman, Shachar Shem-Tov, David Bitton, Tal Nir,
Gilad Adiv, Ron Kimmel, Arie Feuer and Alfred M. Bruckstein *

Abstract. The success of variational methods for optical flow computation lies in their ability to regularize the problem at a differential (pixel) level and combine piecewise smoothness of the flow field with the brightness constancy assumptions. However, the piecewise smoothness assumption is often motivated by heuristic or algorithmic considerations. Lately, new priors were proposed to exploit the structural properties of the flow. Yet, most of them still utilize a generic regularization term.
In this paper we consider optical flow estimation in static scenes. We show that introducing a suitable motion model for the optical flow allows us to pose the regularization term as a geometrically meaningful one. The proposed method assumes that the visible surface can be approximated by a piecewise smooth planar manifold. Accordingly, the optical flow between two consecutive frames can be locally regarded as a homography consistent with the epipolar geometry and defined by only three parameters at each pixel. These parameters are directly related to the equation of the scene local tangent plane, so that their spatial variations should be relatively small, except for creases and depth discontinuities. This leads to a regularization term that measures the total variation of the model parameters and can be extended to a Mumford-Shah segmentation of the visible surface. This new technique yields significant improvements over state of the art optical flow computation methods for static scenes.

Key words: Optical Flow, Epipolar Geometry, Ambrosio-Tortorelli

1 Introduction

Optical flow is defined as the motion field between consecutive frames in a video sequence. Its computation often relies on the brightness constancy assumption, which states that pixel brightness corresponding to a given scene point is constant throughout the sequence. Optical flow computation is a notoriously ill-posed problem. Hence, additional assumptions on the motion are made in order to regularize the problem. Early methods assumed spatial smoothness of the optical flow [1, 2]. Parametric motion models [3, 4], and more recently machine learning [5] were introduced in order to take into account the specificity of naturally occurring video sequences. In parallel, the regularization process was made much more robust [6–9].

In this paper, we focus on optical flow computation in stereoscopic image pairs, given a reliable estimation of the fundamental matrix. This problem has already been

* This research was supported by the Israel Science foundation (ISF) grant no. 1551/09.

addressed in [10–13]. The papers [10, 11] expressed the optical flow as a one-dimensional problem. This was done either by working on a rectified image pair [10], or by solving for the displacement along the epipolar lines [11]. A different approach [12, 13] merely penalized deviation from the epipolar constraint. In addition, [12] proposed a joint estimation of the stereoscopic optical flow and the fundamental matrix. Finally, in order to treat the problem of occluded areas and object boundaries, Ben-Ari and Sochen [14] suggest to explicitly account for regions of discontinuities.

Yet, a third body of works turned to a complete modeling of the scene flow [15–17]. While this approach is the most general, we focus in this paper on static scenes, for which a more specific parameterization can be found.

While the reported experimental results in the aforementioned papers are very convincing, their regularization methods still rely on the traditional assumption that optical flow should be piecewise smooth. Here, motivated by the over-parameterization approach presented in [18], the optical flow is obtained by estimation of the space-time dependent parameters of a motion model, the regularization being applied to the model parameters. In [19], we used homogeneous coordinates to express a homography model, which allows to select a geometrically meaningful coordinate systems for this problem. Here we elaborate upon this model by adding an Ambrosio-Tortorelli scheme, which gives a physically meaningful interpretation for the minima obtained in the optimization process.

In the case of a static scene, the optical flow can be factored into a model determined by the camera motion and an over-parameterized representation of the scene. The scene motion is described locally as a homography satisfying the epipolar constraint and parameterized by the equation of a local planar approximation of the scene. Assuming that the scene can be approximated by a piecewise smooth manifold, enforcing piecewise spatial smoothness on the homography parameters becomes an axiomatically justified regularization criterion which favors piecewise smooth planar regions.

2 Background

2.1 The Variational Framework

In the variational framework for optical flow, brightness constancy and smoothness assumptions are integrated in an energy functional. Let $(u(x,y,t), v(x,y,t))$ denote the optical flow at pixel coordinates (x,y) and time t. Brightness constancy determines the data term of the energy functional

$$E_D(u,v) = \int \Psi\left(I_z^2\right), \tag{1}$$

where

$$I_z = I(x+u, y+v, t+1) - I(x,y,t) \tag{2}$$

and $\Psi(s^2) = \sqrt{s^2+\varepsilon^2}$ is a convex approximation of the L_1 norm for a small ε.

$\mathcal{M}(\mathrm{a}, x, y, t)$ denotes a generic model of the optical flow at pixel (x,y) and time t, where $\mathrm{a} = (\mathrm{a_i}(x,y,t))_{i\in\{1,\ldots,n\}}$ is a family of functions parameterizing the model, i.e.,

$$\begin{pmatrix} u(x,y,t) \\ v(x,y,t) \end{pmatrix} = \mathcal{M}(\mathrm{a}, x, y, t). \tag{3}$$

We begin with the smoothness term proposed by Nir et al. in [18],

$$E_S(\mathrm{a}) = \int \Psi\left(\sum_{i=1}^{n} ||\nabla \mathrm{a_i}||^2\right). \tag{4}$$

In order to refine the discontinuities and obtain a physically meaningful regularization, we extend the smoothness prior using the Ambrosio-Tortorelli scheme [20, 21].

$$E_{S,AT}(\mathrm{a}) = \int v_{AT}^2 \Psi\left(\sum_{i=1}^{n} ||\nabla \mathrm{a_i}||^2\right) + \epsilon_1(1-v_{AT})^2 + \epsilon_2\|\nabla v_{AT}\|^2, \tag{5}$$

where v_{AT} is a diffusivity function, ideally serving as an indicator of the discontinuities set in the flow field. Choosing $\epsilon_1 = \frac{1}{\epsilon_2}$ and gradually decreasing ϵ_2 towards 0 can be used to approximate the Mumford-Shah [22] model via Γ-convergence process, but we do not pursue this direction in this paper.

While the Ambrosio-Tortorelli scheme has been used in the context of optical flow [23–25], in our case this seemingly arbitrary choice of regularization and segmentation has a physical meaning. The regularization of the flow becomes a segmentation process of the *visible surface* in the scene into planar patches, each with his own set of plane parameters. In addition, it helps us obtain accurate edges in the resulting flow.

Furthermore, the generalized Ambrosio-Tortorelli scheme allows us to explicitly reason about the places in the flow where the nonlinear nature of the data manifold manifests itself. Suppose we have a piecewise-planar, static, scene, and an ideal solution (a^*, v_{AT}^*) where a^* is piecewise constant, and the diffusivity function v_{AT}^* is 0 at planar region boundaries and 1 elsewhere. At such a solution, we expect two neighboring points which belong to different regions to have a very small diffusivity value v_{AT} connecting them, effectively nullifying the interaction between different planes' parameters. Furthermore the cost associated with this solution is directly attributed to the discontinuity set measure in the image. The proposed ideal solution therefore becomes a global minimizer of the functional, as determined by the measure of discontinuities in the $2\frac{1}{2}$*-D sketch* [26]. This is directly related to the question raised by Trobin et al. [27] regarding the over-parameterized affine flow model and its global minimizers.

The complete functional now becomes:

$$E(\mathrm{a}) = E_D(\mathcal{M}(\mathrm{a}, x, y, t)) + \alpha E_{S,AT}(\mathrm{a}). \tag{6}$$

In the remainder of this paper, we will propose a motion model enforcing the epipolar constraint and show how to minimize the proposed functional.

2.2 Epipolar Geometry

Let us introduce some background on epipolar geometry, so as to motivate the choice of the motion model. A complete overview can be found in [28, 29].

Given two views of a static scene, the optical flow is restricted by the epipolar constraint. Figure 1 shows that a pixel $\mathbf{m}$ in the left image is restricted to a line $\mathbf{l}'$ called an epipolar line in the right image. All the epipolar lines in the left (resp. right) image go through $\mathbf{e}$ (resp. $\mathbf{e}'$), which is called the left (resp. right) epipole.

In projective geometry, image points and lines are often represented by 3D homogeneous coordinates

$$\mathbf{m} = \left\{ \lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} | \lambda \in \mathbb{R}^{\star} \right\}. \tag{7}$$

Image points and their corresponding epipolar lines are related by the fundamental matrix $\mathcal{F}$

$$\mathbf{l}' = \mathcal{F}\mathbf{m}. \tag{8}$$

Consider a plane π, visible from both cameras, and the planar homography H_{π} which corresponds to the composition of the back-projection from the left view to a plane (π) and the projection from (π) to the right view (see Figure 1). The homography H_{π} gives rise to a useful decomposition of the fundamental matrix

$$\mathcal{F} = [\mathbf{e}']_{\times}\mathrm{H}_{\pi}, \tag{9}$$

where $[\mathbf{e}']_{\times}$ is a matrix representation of the cross product with $\mathbf{e}'$.

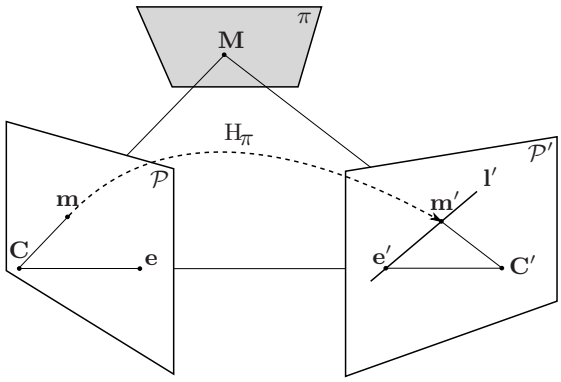


Fig. 1. Epipolar geometry

3 Estimation of the Fundamental Matrix

One of the main challenges in estimating optical flow using the epipolar geometry is to retrieve an accurate and robust estimation of the fundamental matrix. Mainberger et. al. [30] showed that robustness of the fundamental matrix estimation could be achieved by using dense optical flow instead of applying RANSAC or LMedS methods to a sparse set of matches. Hence, we use as initialization the Horn-Schunck with Charbonnier penalty function optical flow implementation provided by Sun et al. [31], modified to use color images. This represents a baseline nonlinear optical flow method, as in [31]. In addition to allowing the computation of the fundamental matrix, this initialization also serves as a starting point for our optical flow computation algorithm.

Many methods aimed at estimating the fundamental matrix can handle large numbers of correspondences. Among those, we choose a robust M-estimation method based on the symmetric epipolar distance, the implementation of which is made very efficient by the use of the Levenberg-Marquardt algorithm, as explained in [32].

4 A Flow Model Based on Local Homographies

We now proceed to develop the model and motivation for the flow equations. Suppose the camera is calibrated, with projection matrices

$$\mathrm{P}(t) = \mathrm{P}_0 = \left(\mathrm{I} \,|\mathbf{0}\right), \;\; \mathrm{P}(t+1) = \mathrm{P}_1 = \left(\mathrm{R} \,|\mathbf{t}\right). \tag{10}$$

where R is a rotation matrix and $\mathbf{t}$ is a translation vector expressing camera motion between the two consecutive frames at t and $t+1$. We assume that locally, the scene is well approximated by the plane

$$\mathbf{v}^{\mathrm{T}}\mathbf{x} + d = 0 \tag{11}$$

where $(\mathbf{x}^{\mathrm{T}}, d)^{\mathrm{T}} = (x, y, 1, d)^{\mathrm{T}}$ denotes the 3D scene point visible at pixel $\mathbf{x}$ in homogeneous coordinates. The corresponding point of $\mathbf{x}$ at time $t+1$ is

$$\mathbf{x}' = \mathrm{P}_1 \begin{pmatrix} \mathbf{x} \\ d \end{pmatrix} = \mathrm{R}\mathbf{x} + \mathbf{t}d = (\mathrm{R} - \mathbf{t}\mathbf{v}^{\mathrm{T}})\mathbf{x} \tag{12}$$

in homogeneous coordinates. $\mathbf{v}$ designates the normal of the local planar approximation of the scene, and $-(\mathbf{v}^{\mathrm{T}}\mathbf{x})^{-1}$ is the depth of the scene at time t. The planar homography expressed in (12) gives a geometrically meaningful motion model parameterized by $\mathbf{v}$. From now on, consider $\mathbf{v}$ as a function of the pixel coordinates. Under the assumption that the scene can be approximated by a piecewise smooth manifold, $\mathbf{v}$ must be piecewise smooth.

We now derive the motion parameterization. In general, the camera parameters are not known, but we can re-parameterize the planar homography using $\mathbf{e}'$ and $\mathcal{F}$. In the following derivation we assume a calibrated view for simplicity's sake. Let $\mathrm{H}(x, y, t)$ denote the planar homography motion model. We have

$$\mathrm{H} \propto \mathrm{R} - \mathbf{t}\mathbf{v}^{\mathrm{T}}. \tag{13}$$

For any compatible planar homography H_0 (cf. [29], 13.1.1.1, we will provide a specific choice later on),

$$\exists(\mathbf{v_0}, \mu): \quad \mathrm{H}_0 = \mu(\mathrm{R} - \mathbf{t}\mathbf{v_0}^{\mathrm{T}}) \tag{14}$$

$$\mathrm{H} = \mathrm{H}_0 - \mu\mathbf{t}(\mathbf{v} - \mathbf{v_0})^{\mathrm{T}}. \tag{15}$$

As $\mathbf{t}$ and $\mathbf{e}'$ are parallel, we can also write

$$\mathrm{H} = \mathrm{H}_0 + \mathbf{e}'\frac{-\mu\mathbf{e}'^{\mathrm{T}}\mathbf{t}}{||\mathbf{e}'||^2}(\mathbf{v} - \mathbf{v_0})^{\mathrm{T}}. \tag{16}$$

Hence, $\mathrm{H}(x,y,t)$ can be parameterized by the function

$$\mathbf{a}(x,y,t) = \frac{-\mu \mathbf{e'}^{\mathrm{T}}\mathbf{t}}{||\mathbf{e'}||^2}(\mathbf{v}(x,y,t) - \mathbf{v_0}), \tag{17}$$

so that

$$\mathrm{H}(x,y,t) = \mathrm{H}_0 + \mathbf{e'}\mathbf{a}(x,y,t)^{\mathrm{T}}. \tag{18}$$

The parameterization $\mathbf{a}$ is the unknown field we want to compute in order to model and estimate the optical flow. The piecewise smoothness of $\mathbf{a}$ is a direct consequence of the piecewise smoothness of $\mathbf{v}$, as testified by (17). More precisely, minimization of the Ambrosio-Tortorelli regularization term favors segmentation of the visible surface into planar patches where the data evidence permits it.

When the cameras are not calibrated, the relationship between the parameterization $\mathbf{a}$ and $\mathbf{v}$ is still linear. In fact, the calibration matrices mainly affect the relative weighting of the model parameters smoothness. Our experiments show that even without controlling the relative smoothness of the model parameters, the optical flow can be estimated accurately.

Note that the parameterization $\mathbf{a}$ can also be derived directly from the fundamental matrix decomposition (9).

For H_0, we can choose the special matrix

$$H_0 = \mathrm{S} = [\mathbf{e'}]_{\times}\mathcal{F}. \tag{19}$$

Each column of S with the corresponding column of $\mathcal{F}$ and $\mathbf{e'}$ form an orthogonal basis of $\mathbb{R}^3$ so that (9) is satisfied. S is a degenerate homography which projects points in the left image to points of the line represented by $\mathbf{e'}$ in the right image. Next, we use the notations

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \mathbf{e'} = \begin{pmatrix} x_{e'} \\ y_{e'} \\ z_{e'} \end{pmatrix}, \quad \mathrm{H}_0 = \begin{pmatrix} \mathbf{h_1}^{\mathrm{T}} \\ \mathbf{h_2}^{\mathrm{T}} \\ \mathbf{h_3}^{\mathrm{T}} \end{pmatrix}, \tag{20}$$

to signify the 3D point coordinates, the epipole's 2D homogeneous coordinates, and the homography matrix rows, respectively. The parameterization of H is introduced into the expression of the optical flow

$$\mathcal{M}(\mathrm{a},x,y,t) = \begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{h_1}^{\mathrm{T}}\mathbf{x} + x_{e'}\mathbf{a}^{\mathrm{T}}\mathbf{x} \\ \mathbf{h_2}^{\mathrm{T}}\mathbf{x} + y_{e'}\mathbf{a}^{\mathrm{T}}\mathbf{x} \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix}, \quad \lambda = \frac{1}{\mathbf{h_3}^{\mathrm{T}}\mathbf{x} + z_{e'}\mathbf{a}^{\mathrm{T}}\mathbf{x}}. \tag{21}$$

where $\begin{pmatrix} x \\ y \end{pmatrix}$ are the corresponding pixels in the left image.

4.1 Euler-Lagrange Equations

By interchangeably fixing $a_i, i = 1...n$ and v_{AT}, we obtain the Euler-Lagrange equations which minimize the functional.

Minimization with respect to a_i. Fixing v_{AT}, we obtain

$$\forall i, \quad \nabla_{\mathrm{a_i}}(E_D + \alpha v_{AT}^2 E_S) = 0. \tag{22}$$

the variation of the data term with respect to the model parameter function $\mathrm{a_i}$ is given by

$$\nabla_{\mathrm{a_i}} E_D(u,v) = 2\Psi'\left(I_z^2\right) I_z \nabla_{\mathrm{a_i}} I_z, \tag{23}$$

where

$$\nabla_{\mathrm{a_i}} I_z \;=\; \lambda^2 x_i (x_{e'} \mathbf{h_3}^{\mathrm{T}}\mathbf{x} - z_{e'}\mathbf{h_1}^{\mathrm{T}}\mathbf{x}) I_x^+ + \lambda^2 x_i (y_{e'}\mathbf{h_3}^{\mathrm{T}}\mathbf{x} - z_{e'}\mathbf{h_2}^{\mathrm{T}}\mathbf{x}) I_y^+, \tag{24}$$

and

$$I_x^+ = I_x(x+u, y+v, t+1) \tag{25}$$

$$I_y^+ = I_y(x+u, y+v, t+1). \tag{26}$$

For the smoothness term, the Euler-Lagrange equations are

$$\nabla a_i E_s = 2 v_{AT} \Psi\left(\sum_{i=1}^{n} ||\nabla a_i||^2\right) + 2 v_{AT}^2 div\left(\Psi'\left(\sum_j ||\nabla a_j||^2\right)\nabla a_i\right) \tag{27}$$

thus, the energy is minimized by solving the nonlinear system of equations

$$\Psi'\left(I_z^2\right) I_z \nabla_{a_i} I_z - \alpha \nabla\left(v_{AT}^2 \Psi'\left(\sum_{i=1}^{n} ||\nabla a_i||^2\right)\right)^T \nabla a_i - \alpha v_{AT}^2 div\left(\Psi'\left(\sum_j ||\nabla a_j||^2\right)\nabla a_i\right) = 0. \tag{28}$$

Minimization with respect to v_{AT}. Fixing a_i, we obtain

$$2\alpha v_{AT}\Psi\left(\sum_{i=1}^{n} ||\nabla a_i||^2\right) + 2\epsilon_1(v_{AT} - 1) - \epsilon_2 \Delta v_{AT} = 0 \tag{29}$$

4.2 Implementation

Minimization with respect to v_{AT} is straightforward, as the equations are linear with respect to v_{AT}, therefore we will only elaborate on the minimization with respect to a_i

The nonlinear Euler-Lagrange equation minimizing a_i, are linearized by adopting three embedded loops, similarly to [18]. First, the warped image gradient (I_x^+, I_y^+) is frozen, and so is λ. At each iteration k, we have

$$(\nabla_{\mathrm{a_i}} I_z)^k = x_i d^k \tag{30}$$

where

$$d^k = (\lambda^k)^2(x_{e'}\mathbf{h_3}^{\mathrm{T}}\mathbf{x} - z_{e'}\mathbf{h_1}^{\mathrm{T}}\mathbf{x})(I_x^+)^k + (\lambda^k)^2(y_{e'}\mathbf{h_3}^{\mathrm{T}}\mathbf{x} - z_{e'}\mathbf{h_2}^{\mathrm{T}}\mathbf{x})(I_y^+)^k,$$

and the following approximation is made using first order Taylor expansions

$$I_z^{k+1} \approx I_z^k + d^k \sum_{i=1}^{3} x_i d\mathrm{a_i}^k \tag{31}$$

where

$$d\mathbf{a}^k = \mathbf{a}^{k+1} - \mathbf{a}^k. \tag{32}$$

The system of equations (28) becomes

$$\Psi'\left((I_z^{k+1})^2\right)\left(I_z^k + d^k \sum_{j=1}^{3} x_j d\mathrm{a_j}^k\right) x_i d^k - \alpha \operatorname{div}\left(\Psi'\left(\sum_j ||\nabla \mathrm{a_j}^{k+1}||^2\right)\nabla \mathrm{a_i}^{k+1}\right) = 0.$$

A second loop with superscript l is added to cope with the nonlinearity of Ψ'.

$$(\Psi')^{k,l}_{\mathrm{Data}}\left(I_z^k + d^k \sum_{j=1}^{3} x_j d\mathrm{a_j}^{k,l+1}\right) x_i d^k - \alpha \operatorname{div}\left((\Psi')^{k,l}_{\mathrm{Smooth}} \nabla \mathrm{a_i}^{k,l+1}\right) = 0$$

where

$$(\Psi')^{k,l}_{\mathrm{Data}} = \Psi'\left(\left(I_z^k + d^k \sum_{i=1}^{3} x_i d\mathrm{a_i}^{k,l}\right)^2\right), \quad (\Psi')^{k,l}_{\mathrm{Smooth}} = \Psi'\left(\sum_j ||\nabla \mathrm{a_j}^{k,l}||^2\right).$$

At this point, the system of equations is linear and sparse in the spatial domain. The solution $\mathbf{a}$, as well as the diffusivity term v_{AT} are obtained through Gauss-Seidel iterations. In the case of the Ambrosio-Tortorelli regularization term, the diffusion term of the equation is modulated by v_{AT}.

5 Experimental results

We now demonstrate motion estimation results using our algorithm, both visually and in terms of the average angular error (AAE). No post-processing was applied to the optical flow field obtained after energy minimization. The algorithm was tested on image pairs from the Middlebury optical flow test set [33], as well as all images with a static scene and publicly available ground truth optical flow from the training set. Results from the training set are presented in Table 1.

The flow, parameters, and diffusivity field resulting from our method are presented in Figure 3. The optical flow is shown with color encoding and a disparity map.

Results from the test set are shown in Figure 2. A smoothness parameter α of 400 was used in all experiments, and the Ambrosio-Tortorelli coefficients were set to $\epsilon_1 =$

$20, \epsilon_2 = 5 \times 10^{-5}$. The proposed method produced the best results to date on the static Yosemite and Urban scenes. The algorithm is not designed, however, for non-static scenes, where the computed epipolar lines have no meaning. One possible solution to this shortcoming is to return to a 2D search [13]. Such a combined approach is left for future work.

In the Teddy and Grove test images, the initialization of our algorithm introduced errors in significant parts of the image, which our method could not overcome. This behavior is related to the problem of finding a global minimum for the optical flow, which is known to have several local minima. Improving the global convergence using discrete graph-based techniques, has been the focus of several papers (see [34–36], for example), and is beyond the scope of this work. We expect better initialization to improve the accuracy to that of the Yosemite and Urban image pairs.

Our optical flow estimation for the Yosemite and Urban sequences gives the best results to date, achieving an AAE of 1.25 for the Yosemite sequence test pair and 2.38 for the Urban sequence, as shown in Figure 2. When the fundamental matrix estimate was improved (by estimating from the ground truth optical flow), we reduced the AAE to 0.66 for Yosemite!

	AAE	STD
Grove2	2.41	7.16
Grove3	5.53	15.76
Urban2	2.15	9.22
Urban3	3.84	16.88
Venus	4.29	12.01
Yosemite	0.85	1.24

(a) Middlebury training set

Method	AAE	Method	AAE
Brox et al. [7]	1.59	Roth/Black [5]	1.43
Mémin/Pérez [4]	1.58	Valgaerts et al. [12]	1.17
Bruhn et al. [8]	1.46	Nir et al. [18]	1.15
Amiaz et al. [37]	1.44	Our method	0.85

(b) Yosemite sequence

Table 1. AAE comparison for static scenes of the Middlebury training set and for the Yosemite sequence

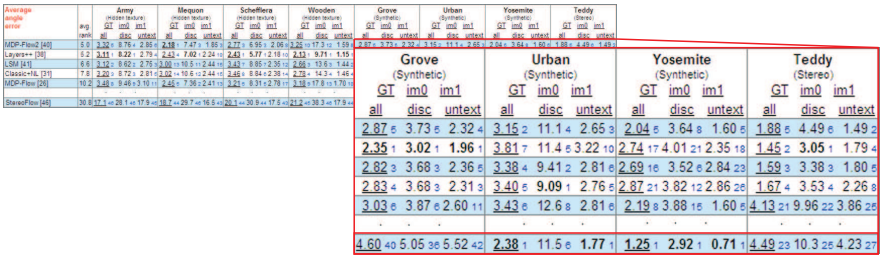
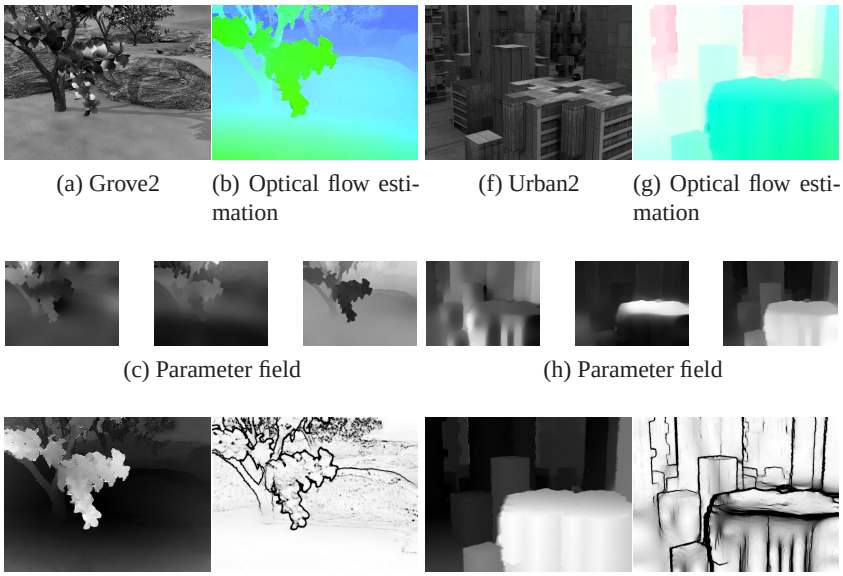


Fig. 2. Average angular error values of our algorithm, compared on the middlebury test set. The smoothness coefficient was set to $\alpha = 400$ in all experiments. Red marks the row of the suggested algorithm.



(a) Grove2 (b) Optical flow estimation (f) Urban2 (g) Optical flow estimation

(c) Parameter field (h) Parameter field

(d) Disparity estimation (e) Diffusivity function (i) Disparity estimation (j) Diffusivity function

Fig. 3. Grove2 and Urban2 sequence results

It is interesting to look at the results obtained for scenes with planar regions, such as the Urban2 (Figure 3) image pair. In Urban2, the scene is composed of many planar patches, modeled by constant patches in the model parameters. In both these scenes, as well as others, the resulting diffusivity field clearly marks the contours of planar regions in the image such as the buildings in Urban2 and the tree and soil ridges in Grove2.

6 Conclusions

A new method for optical flow computation was presented, which hinges on a guiding principle that optic flow regularization should have a strong theoretical foundation. The method is applicable to static scenes and retrieves meaningful local motion parameters related to the scene geometry. At each pixel, the parameters provide an estimation of the plane tangent to the scene manifold, up to a fixed shift and scale. To that extent, they can be seen as a higher level output than optical flow in the computer vision hierarchy.

An interesting aspect of our energy functional, which was already mentioned in [18], is that given a carefully selected over-complete parameter field, the different parameters support each other to find a smooth piecewise constant parameter patches, while the incorporated Ambrosio-Tortorelli scheme prevents diffusion across discontinuities. Furthermore, the Ambrosio-Tortorelli scheme allows us to combine regular-

ization and segmentation, resulting in a physically meaningful regularization process, while minimizing the dependency on the relative scaling of the coefficients.

Finally, although the performance demonstrated already goes beyond the latest published results, there is still much gain to be expected from better fundamental matrix estimation and algorithm initialization. In addition, when more than two frames are available and the camera pose is known, augmenting the model with time-smoothness is expected to systematically improve the results.

References

1. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artificial Intelligence **17** (1981) 185–203
2. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: International Joint Conference on Artificial Intelligence. (1981) 674–679
3. Bergen, J., Anandan, P., Hanna, K., Hingorani, R.: Hierarchical model-based motion estimation. In: ECCV. Volume 2., Springer (1992) 237–252
4. Mémin, E., Pérez, P.: Hierarchical estimation and segmentation of dense motion fields. International Journal of Computer Vision **46** (2002) 129–155
5. Roth, S., Black, M.J.: On the spatial statistics of optical flow. International Journal of Computer Vision **74** (2007) 33–50
6. Black, M.J., Anandan, P.: A framework for the robust estimation of optical flow. In: International Conference on Computer Vision. (1993) 231–236
7. Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: High accuracy optical flow estimation based on a theory for warping. In: European Conference on Computer Vision. Volume 3024., Springer (2004) 25–36
8. Bruhn, A., Weickert, J., Schnörr, C.: Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. International Journal of Computer Vision **61** (2005) 211–231
9. Cohen, I.: Nonlinear variational method for optical flow computation. In: Proc. Eighth Scandinavian Conference on Image Analysis. Volume 1. (1993) 523–530
10. Birchfield, S., Tomasi, C.: Depth discontinuities by pixel-to-pixel stereo. International Journal of Computer Vision **35** (1999) 269–293
11. Slesareva, N., Bruhn, A., Weickert, J.: Optic flow goes stereo: A variational method for estimating discontinuity-preserving dense disparity maps. In: DAGM Symposium. Volume 3663., Springer (2005) 33–40
12. Valgaerts, L., Bruhn, A., Weickert, J.: A variational model for the joint recovery of the fundamental matrix and the optical flow. In: DAGM Symposium, Springer-Verlag (2008) 314–324
13. Wedel, A., Pock, T., Braun, J., Franke, U., Cremers, D.: Duality TV-L1 flow with fundamental matrix prior. Image Vision and Computing New Zealand (2008) 1–6
14. Ben-Ari, R., Sochen, N.: Variational stereo vision with sharp discontinuities and occlusion handling. In: International Conference on Computer Vision, IEEE Computer Society (2007) 1–7
15. Pons, J.P., Keriven, R., Faugeras, O., Hermosillo, G.: Variational stereovision and 3d scene flow estimation with statistical similarity measures. International Conference on Computer Vision **1** (2003) 597
16. Huguet, F., Devernay, F.: A variational method for scene flow estimation from stereo sequences. Computer Vision and Pattern Recognition **0** (2007) 1–7
17. Basha, T., Moses, Y., Kiryati, N.: Multi-view scene flow estimation: A view centered variational approach. In: Computer Vision and Pattern Recognition. (2010) 1506–1513

18. Nir, T., Bruckstein, A.M., Kimmel, R.: Over-parameterized variational optical flow. International Journal of Computer Vision **76** (2008) 205–216
19. Anonymous: Over-parameterized optical flow using a stereoscopic constraint. (Technical report)
20. Ambrosio, L., Tortorelli, V.M.: Approximation of functional depending on jumps by elliptic functional via Γ-convergence. Communications on Pure and Applied Mathematics **43** (1990) 999–1036
21. Shah, J.: A common framework for curve evolution, segmentation and anisotropic diffusion. In: Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition (CVPR '96). Computer Vision and Pattern Recognition, Washington, DC, USA, IEEE Computer Society (1996) 136–
22. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics **42** (1989) 577–685
23. Amiaz, T., Kiryati, N.: Piecewise-smooth dense optical flow via level sets. International Journal of Computer Vision **68** (2006) 111–124
24. Brune, C., Maurer, H., Wagner, M.: Detection of intensity and motion edges within optical flow via multidimensional control. **2** (2009) 1190–1210
25. Ben-Ari, R., Sochen, N.A.: Stereo matching with mumford-shah regularization and occlusion handling. IEEE Trans. Pattern Anal. Mach. Intell. **32** (2010) 2071–2084
26. Marr, D.: Vision: a computational investigation into the human representation and processing of visual information. W. H. Freeman, San Francisco (1982)
27. Trobin, W., Pock, T., Cremers, D., Bischof, H.: An unbiased second-order prior for high-accuracy motion estimation. In: DAGM Symposium, Springer-Verlag (2008) 396–405
28. Faugeras, O., Luong, Q.T.: The Geometry of Multiple Images. The MIT Press, ISBN: 0262062208 (2001)
29. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press, ISBN: 0521540518 (2004)
30. Mainberger, M., Bruhn, A., Weickert, J.: Is dense optic flow useful to compute the fundamental matrix? In: International Conference on Image Analysis and Recognition, Springer-Verlag (2008) 630–639
31. Sun, D., Roth, S., Black, M.J.: Secrets of optical flow estimation and their principles. In: CVPR. (2010) 2432–2439
32. Klappstein, J.: Optical-Flow Based Detection of Moving Objects in Traffic Scenes. PhD thesis, Ruprecht-Karls-Universität, Heidelberg (2008)
33. Baker, S., Roth, S., Scharstein, D., Black, M., Lewis, J., Szeliski, R.: A Database and Evaluation Methodology for Optical Flow. In: International Conference on Computer Vision. (2007) 1–8
34. Lempitsky, V., Rother, C., Roth, S., Blake, A.: Fusion moves for markov random field optimization. IEEE Trans. Pattern Anal. Mach. Intell. **99** (2009)
35. Liu, Y., Cao, X., Dai, Q., Xu, W.: Continuous depth estimation for multi-view stereo. In: Computer Vision and Pattern Recognition. (2009) 2121–2128
36. Kim, W., Park, J., Lee, K.: Stereo matching using population-based mcmc. International Journal of Computer Vision **83** (2009) 195–209
37. Amiaz, T., Lubetzky, E., Kiryati, N.: Coarse to over-fine optical flow estimation. Pattern Recognition **40** (2007) 2496–2503