Over-Parameterized Variational Optical Flow

Tal Nir · Alfred M. Bruckstein · Ron Kimmel

Abstract A novel optical flow estimation process based on a spatio-temporal model with varying coefficients multiplying a set of basis functions at each pixel is introduced. Previous optical flow estimation methodologies did not use such an over parameterized representation of the flow field as the problem is ill-posed even without introducing any additional parameters: Neighborhood based methods of the Lucas–Kanade type determine the flow at each pixel by constraining the flow to be described by a few parameters in small neighborhoods. Modern variational methods represent the optic flow directly via the flow field components at each pixel. The benefit of over-parametrization becomes evident in the smoothness term, which instead of directly penalizing for changes in the optic flow, accumulates a cost of deviating from the assumed optic flow model. Our proposed method is very general and the classical variational optical flow techniques are special cases of it, when used in conjunction with constant basis functions. Experimental results with the novel flow estimation process yield significant improvements with respect to the best results published so far.

Keywords Optical flow · Variational methods · L_1 regularization · Over-parametrization

T. Nir (✉) · A.M. Bruckstein · R. Kimmel
Department of Computer Science, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel
e-mail: taln@cs.technion.ac.il

A.M. Bruckstein
e-mail: freddy@cs.technion.ac.il

R. Kimmel
e-mail: ron@cs.technion.ac.il

1 Introduction

Despite much research effort invested in addressing optical flow computation it remains a challenging task in the field of computer vision. It is a necessary step in various applications like stereo matching, video compression, object tracking, depth reconstruction and motion based segmentation. Hence, many approaches have been proposed for optical flow computation. Most methods assume brightness constancy and introduce additional assumptions on the optical flow in order to deal with the inherent aperture problem. Lucas and Kanade (1981) tackled the aperture problem by solving for the parameters of a constant motion model over image patches. Subsequently, Irani et al. (1993, 1997) used motion models in a region in conjunction with Lucas–Kanade in order to recover the camera ego-motion. Spline based motion models were suggested by Szeliski and Coughlan (1997).

Horn and Schunck (1981) sought to recover smooth flow fields and were the first to use functional minimization for solving optical flow problems employing mathematical tools from calculus of variations. Their pioneering work put forth the basic idea for solving dense optical flow fields over the whole image by introducing a quality functional with two terms: a data term penalizing for deviations from the brightness constancy equation, and a smoothness term penalizing for variations in the flow field. Several important improvements have been proposed following their work. Nagel (1990, 1986) proposed an oriented smoothness term that penalizes anisotropically for variations in the flow field according to the direction of the intensity gradients. Ari and Sochen (2006) recently used a functional with two alignment terms composed of the flow and image gradients. Replacing quadratic penalty terms by robust statistics integral measures was proposed in (Black and Anandan 1996;

Deriche et al. 1995) in order to allow sharp discontinuities in the optical flow solution along motion boundaries. Extensions to multi-frame formulations of the initial two-frames formulation allowed the consideration of spatio-temporal smoothness to replace the original spatial smoothness term (Black and Anandan 1991; Farnebäck 2001; Nagel 1990; Weickert and Schnörr 2001). Brox et al. (2004, 2006) demonstrated the importance of using the exact brightness constancy equation instead of its linearized version and added a gradient constancy to the data term which may be important if the scene illumination changes in time. Cremers and Soatto (2005) proposed a motion competition algorithm for variational motion segmentation and parametric motion estimation. Amiaz and Kiryati (2005) followed by Brox et al. (2006) introduced a variational approach for joint optical flow computation and motion segmentation. In Farneback (2000, 2001), a constant and affine motion model is employed. The motion model is assumed to act on a region, and optic flow based segmentation is performed by a region growing algorithm. In a classical contribution to structure from motion Adiv (1985) used optical flow in order to determine motion and structure of several rigid objects moving in the scene. Sekkati and Mitiche (2006) used joint segmentation and optical flow estimation in conjunction with a single rigid motion in each segmented region. Vázquez et al. (2006) used joint multi-region segmentation with high order DCT basis functions representing the optical flow in each segmented region.

In this paper, we propose to represent the optical flow vector at each pixel by different coefficients of the same motion model in a variational framework. Such a grossly over-parameterized representation has the advantage that the smoothness term may now penalize deviations from the motion model instead of directly penalizing the change of the flow. For example, in an affine motion model, if the flow in a region can be accurately represented by an affine model, then in this region there will be no flow regularization penalty, while in the usual setting there is a cost resulting from the changes in the flow induced by the affine model. This over-parameterized model thereby offers a richer means for optical flow representation. For segmentation purposes, the over-parametrization has the benefit of making segmentation decisions in a more appropriate space (e.g. the parameters of the affine flow) rather than in a simple constant motion model space. The work of Ju et al. (1996) is related to our methodology, they used local affine models to describe the motion in image regions imposing spatial smoothness on the affine parameters between neighboring patches. The key and conceptually very important difference is that, in our approach, the model is represented at the pixel level which makes the problem over-parameterized while the patch size chosen in (Ju et al. 1996) makes it under-parameterized and requires the choice of a neighborhood size.

The paper is organized as follows: Sect. 2 introduces the over-parameterized optical flow representation model and the corresponding functional together with the resulting Euler–Lagrange equations, examples of specific optical flow models are described. Section 3 discusses numerical solution considerations. Section 4 describes the parameter settings and the experiments conducted to evaluate our method. Finally, Sect. 5 concludes the paper.

2 Over-Parametrization Model

We propose to represent the optical flow $(u(x, y, t), v(x, y, t))$ by the general over-parameterized space-time model

$$u(x, y, t) = \sum_{i=1}^{n} A_i(x, y, t)\phi_i(x, y, t),$$

$$v(x, y, t) = \sum_{i=1}^{n} A_i(x, y, t)\eta_i(x, y, t), \tag{1}$$

where, $\phi_i(x, y, t)$ and $\eta_i(x, y, t)$, $i = 1, \ldots, n$ are n basis functions of the flow model, while the A_i are space and time varying coefficients of the model. This is an obviously heavily over-parameterized model since for more than two basis functions, there are typically many ways to express the same flow at any specific location. This redundancy however will be adequately resolved by a regularization assumption applied to the coefficients of the model. The coefficients and basis functions may be general functions of space-time, however, they play different roles in the functional minimization process: The basis functions are fixed and selected a priori. The coefficients are the unknown functions we solve for in the optical flow estimation process. In our model, appropriate basis functions are such that the true flow could be described by approximately piecewise constant coefficients, so that most of the local spatio-temporal changes of the flow are induced by changes in the basis functions and not by variations of the coefficients. This way, regularization applied to the coefficients (as will be described later on) becomes meaningful since major parts in the optic flow variations can be described without changes of the coefficients. For example, rigid body motion has a specific optical flow structure which can explain the flow using only six parameters at locations with approximately constant depth. Let us start from conventional optical flow functionals that include a data term $E_D(u, v)$, that measures the deviation from the brightness constancy assumption, and a regularization (or smoothness) term $E_S(u, v)$ that quantifies the smoothness of the flow field. The solution is the minimizer of the sum of the data and smoothness terms

$$E(u, v) = E_D(u, v) + \alpha E_S(u, v). \tag{2}$$

The main difference between the diverse variational methods is in the choice of data and smoothness terms, and in the numerical methods used for solving for the minimizing flow field $(u(x, y, t), v(x, y, t))$. For the data term we shall use the functional

$$E_D(u, v) = \int \Psi((I(\mathbf{x} + \mathbf{w}) - I(\mathbf{x}))^2)d\mathbf{x}, \tag{3}$$

where, $\mathbf{x} = (x, y, t)^T$ and $\mathbf{w} = (u, v, 1)^T$. This is the integral measure used for example in (Brox et al. 2004) (omitting the gradient constancy term). The function $\Psi(s^2) = \sqrt{s^2 + \varepsilon^2}$ that is by now widely used, induces an approximate L_1 metric of the data term for a small ε. The smoothness term used in (Brox et al. 2004) is given by

$$E_S(u, v) = \int \Psi(\|\widetilde{\nabla}u\|^2 + \|\widetilde{\nabla}v\|^2)d\mathbf{x}, \tag{4}$$

where $\widetilde{\nabla} f \equiv (f_x, f_y, \omega_t f_t)^T$ denotes the weighted spatio-temporal gradient. ω_t indicates the weight of the temporal axis relative to the spatial axes in the context of the smoothness term ($\omega_t = 1$ is used in (Brox et al. 2004)). Inserting the over-parameterized model into the data term, we have

$$E_D(A_i) = \int \Psi\left(\left(I\left(x + \sum_{i=1}^{n} A_i\phi_i, y + \sum_{i=1}^{n} A_i\eta_i, t+1\right)\right.\right.$$
$$\left.\left. - I(x, y, t)\right)^2\right)d\mathbf{x}. \tag{5}$$

Our proposed smoothness term replaces (4) with a penalty for spatio-temporal changes in the coefficient functions,

$$E_S(A_i) = \int \Psi\left(\sum_{i=1}^{n} \|\widetilde{\nabla}A_i\|^2\right)d\mathbf{x}. \tag{6}$$

Notice that in (6), constant parameters of the model can describe changes of the flow field according to the chosen model as described in (1) (e.g. Euclidean, affine, etc.) without smoothness penalty, whereas in (4), any change in the flow field is penalized by the smoothness term. For the sake of simplicity of the resulting Euler–Lagrange equations, we have omitted writing explicit relative weights to the different coefficients in the smoothness term, the weighting is alternatively achieved by scaling the basis functions by appropriate factors as will be shown in the description of the motion models. Scaling a basis function by a small factor would mean that in order to achieve the same overall influence on the optical flow, the corresponding coefficient would have to make larger changes (proportional to the inverse of the factor). These larger changes would be suppressed by the regularization term. On the other hand, scaling a basis function by a large factor would scale down the changes required

from the corresponding coefficient in order to achieve the same overall change and therefore would result in less regularization for this specific term.

2.1 Euler–Lagrange Equations

For an over-parametrization model with n coefficients, there are n Euler–Lagrange equations, one for each coefficient. The Euler–Lagrange equation for A_q ($q = 1, \ldots, n$) is given by

$$\Psi'(I_z^2)I_z(I_x^+\phi_q + I_y^+\eta_q)$$
$$- \alpha \cdot \text{div}\left(\Psi'\left(\sum_{i=1}^{n} \|\tilde{\nabla}A_i\|^2\right)\hat{\nabla}A_q\right) = 0. \tag{7}$$

where $\hat{\nabla} f \equiv (f_x, f_y, \omega_t^2 f_t)^T$, and

$$I_x^+ := I_x\left(x + \sum_{i=1}^{n} A_i\phi_i, y + \sum_{i=1}^{n} A_i\eta_i, t+1\right),$$

$$I_y^+ := I_y\left(x + \sum_{i=1}^{n} A_i\phi_i, y + \sum_{i=1}^{n} A_i\eta_i, t+1\right), \tag{8}$$

$$I_z := I\left(x + \sum_{i=1}^{n} A_i\phi_i, y + \sum_{i=1}^{n} A_i\eta_i, t+1\right) - I(x, y, t).$$

2.2 The Affine Over-Parametrization Model

The affine model is a good approximation of the flow in large regions in many real world scenarios. We therefore first start with the affine model for our method. Note that we do not force the affine model over image patches as in previously considered image registration techniques, and here each pixel has "its own" independent affine model parameters. The affine model has $n = 6$ parameters,

$$\begin{aligned}
&\phi_1 = 1; &&\phi_2 = \hat{x}; &&\phi_3 = \hat{y}; \\
&\phi_4 = 0; &&\phi_5 = 0; &&\phi_6 = 0; \\
&\eta_1 = 0; &&\eta_2 = 0; &&\eta_3 = 0; \\
&\eta_4 = 1; &&\eta_5 = \hat{x}; &&\eta_6 = \hat{y},
\end{aligned} \tag{9}$$

where,

$$\hat{x} = \frac{\rho(x - x_0)}{x_0},$$
$$\hat{y} = \frac{\rho(y - y_0)}{y_0}, \tag{10}$$

and x_0 and y_0 are half image width and height respectively. ρ is a constant that has no meaning in an unconstrained optimization such as the Lucas–Kanade method. In our variational formulation, ρ is a parameter which weighs

the penalty of the x and y coefficients relative to the coefficient of the constant term in the regularization. An equivalent alternative is to add a different weight for each coefficient in (6).

2.3 The Rigid Motion Model

The optic flow of an object moving in rigid motion or of a static scene with a moving camera is described by

$$u = -\theta_1 + \theta_3 \hat{x} + \Omega_1 \hat{x}\hat{y} - \Omega_2(1 + \hat{x}^2) + \Omega_3 \hat{y},$$
$$v = -\theta_2 + \theta_3 \hat{y} + \Omega_1(1 + \hat{y}^2) - \Omega_2 \hat{x}\hat{y} - \Omega_3 \hat{x}, \quad (11)$$

where $(\theta_1, \theta_2, \theta_3)^T$ is the translation vector divided by the depth and $(\Omega_1, \Omega_2, \Omega_3)^T$ is the rotation vector. Here too, the number of coefficients is $n = 6$. The coefficients A_i represent the translation and rotation variables, $A_1 = \theta_1$; $A_2 = \theta_2$; $A_3 = \theta_3$; $A_4 = \Omega_1$; $A_5 = \Omega_2$; $A_6 = \Omega_3$, and the basis functions are $\phi_1 = -1$; $\phi_2 = 0$; $\phi_3 = \hat{x}$; $\phi_4 = \hat{x}\hat{y}$; $\phi_5 = -(1 + \hat{x}^2)$; $\phi_6 = \hat{y}$ and $\eta_1 = 0$; $\eta_2 = -1$; $\eta_3 = \hat{y}$; $\eta_4 = 1 + \hat{y}^2$; $\eta_5 = -\hat{x}\hat{y}$; and $\eta_6 = -\hat{x}$. Similar constraints on optical flow of rigid motion were first introduced by Adiv in (Adiv 1985). However, there the optical flow is a pre-processing step to be followed by structure and motion estimation, while our formulation uses the rigid motion model in the optimality criterion of the optical flow estimation process. Using the rigid motion model directly in the optimality criterion was previously suggested by Sekkati and Mitiche (2003) where the functional is explicitly written in terms of the model parameters. However, since the data term in (Sekkati and Mitiche 2003) is quadratic and uses the linearized brightness constancy assumption, it is expected to be more sensitive to outliers and prone to errors for motion fields of large magnitudes. Since the optical flow induced by camera rotation is independent of the depth it has a more global nature and therefore one may wish to penalize more severely for changes in rotation when considering the smoothness term. This can be done by scaling the basis functions multiplying the rotation coefficients by a factor between 0 and 1. Such a factor would require larger changes in the coefficients in order to achieve the same overall influence of the rotation on the optical flow. Such larger changes in the corresponding coefficients would be suppressed by the regularization term and therefore achieve a more global effect of the rotation. Note, that assuming rigid motion one could also extract the depth profile (up to scaling) from the above coefficients.

2.4 Pure Translation Motion Model

A special case of the rigid motion scenario can be thought of when we limit the motion to simple translation. In this case we have,

$$u = -\theta_1 + \theta_3 \hat{x},$$
$$v = -\theta_2 + \theta_3 \hat{y}. \quad (12)$$

The Euler–Lagrange equations of the rigid motion still applies in this case when considering only the first $n = 3$ coefficients and corresponding basis functions.

2.5 Constant Motion Model

The constant motion model includes only $n = 2$ coefficients, with

$$\phi_1 = 1; \qquad \phi_2 = 0,$$
$$\eta_1 = 0; \qquad \eta_2 = 1, \quad (13)$$

as basis functions. For this model there are two coefficients to solve for A_1 and A_2 which are the optic flow components u and v, respectively. In this case we obtain the familiar variational formulation where we solve for the u and v components. This fact can also be seen by substitution of (13) into (7), that yields the Euler–Lagrange equations used, for example, in (Brox et al. 2004) (without the gradient constancy term).

3 Numerical Scheme

In our numerical scheme we use a multi-resolution solver, by down sampling the image data with a standard factor of 0.5 along the x and y axes between the different resolutions. The solution is interpolated from coarse to fine resolution. Similar techniques to overcome the intrinsic non-convex nature of the resulting optimization problem were used, for example, in (Brox et al. 2004).[1] At the lowest resolution, we start with the initial guess of $A_i = 0$, $i = 1, \ldots, n$. From this guess, the solution is iteratively refined and the coefficients are interpolated to become the initial guess at the next higher resolution and the process is repeated until the solution at the finest resolution is reached. At each resolution, three loops of iterations are applied. At the outer loop with iteration variable k, we freeze the brightness constancy linear approximation in the Euler–Lagrange equations

---

[1]We compare our model and results to (Brox et al. 2004) since, to the best of our knowledge, that paper reports the most accurate flow field results for the Yosemite without clouds sequence.

$$\Psi'((I_z^{k+1})^2) \cdot I_z^{k+1} \cdot ((I_x^+)^k \phi_q + (I_y^+)^k \eta_q)$$

$$- \alpha \cdot \mathrm{div} \left(\Psi' \left(\sum_{i=1}^{n} \|\tilde{\nabla} A_i^{k+1}\|^2 \right) \hat{\nabla} A_q^{k+1} \right) = 0. \qquad (14)$$

Inserting the first terms of the Taylor expansion

$$I_z^{k+1} \approx I_z^k + (I_x^+)^k \sum_{i=1}^{n} dA_i^k \phi_i + (I_y^+)^k \sum_{i=1}^{n} dA_i^k \eta_i, \qquad (15)$$

where $A_i^{k+1} = A_i^k + dA_i^k$. The second inner loop—fixed point iteration—deals with the nonlinearity of Ψ with iteration variable l, it uses the expressions of Ψ' from the previous iteration in both the data and smoothness terms, while the rest of the equation is written with respect to the $l+1$ iteration

$$(\Psi')_{\mathrm{Data}}^{k,l} dq \left(I_z^k + \sum_{i=1}^{n} dA_i^{k,l+1} d_i \right)$$

$$- \alpha \cdot \mathrm{div}((\Psi')_{\mathrm{Smooth}}^{k,l} \hat{\nabla}(A_q^k + dA_q^{k,l+1})) = 0, \qquad (16)$$

where,

$$d_m := (I_x^+)^k \phi_m + (I_y^+)^k \eta_m, \qquad (17)$$

$$(\Psi')_{\mathrm{Data}}^{k,l} := \Psi' \left(\left(I_z^k + \sum_{i=1}^{n} dA_i^{k,l} d_i \right)^2 \right), \qquad (18)$$

and

$$(\Psi')_{\mathrm{Smooth}}^{k,l} := \Psi' \left(\sum_{i=1}^{n} \|\tilde{\nabla}(A_i^k + dA_i^{k,l})\|^2 \right). \qquad (19)$$

At this point, we have for each pixel n linear equations with n unknowns, the increments of the coefficients of the model parameters. The linear system of equations is solved on the sequence volume using Gauss–Seidel iterations. Each Gauss–Seidel iteration involves the solution of n linear equations for each pixel as described by (16). The discretization uses two point central difference for the flow components and four point central difference for the image derivatives as suggested in (Barron et al. 1994).

4 Experimental Results

In this section we compare our optical flow computation results to the best published results. For test sequences with ground truth, we use the standard measures of Average Angular Error (AAE) and Standard Deviation (STD). Our results are measured over all the pixels (100% dense). The angle is defined as:

$$\theta = \arccos \left(\frac{uu_g + vv_g + 1}{\sqrt{(u^2 + v^2 + 1)(u_g^2 + v_g^2 + 1)}} \right) \qquad (20)$$

Table 1 Parameter settings

Method	α	ρ	ω_t^2	ε	σ
Constant motion (3D)	16.0	–	9.0	0.001	0.8
Affine (2D)	58.3	0.858	0	0.001	0.8
Pure translation (2D)	51.0	0.575	0	0.001	0.8
Affine (3D)	32.9	1.44	0.474	0.001	0.8
Rigid motion (3D)	54.6	1.42	0.429	0.001	0.8
Pure translation (3D)	23.6	1.22	0.688	0.001	0.8

where u and v are the estimated optical flow components, and u_g and v_g represent the ground truth optical flow. The AAE is the average and STD is the standard deviation of θ over the image domain.

4.1 Parameter Settings

The parameters were set experimentally by an optimization process which numerically minimizes the weighted sum of AAE and STD measured on the Yosemite sequence. Such a parameter optimization or training process is usual in many other papers, for example (Roth and Black 2005). Brox et al. (2004), Papenberg et al. (2006) also seem to have an optimal parameter setting since in their parameter sensitivity analysis, each parameter change in any direction results in an increase of the AAE measure for the Yosemite sequence. We have found slightly different parameter settings for the 2D and 3D smoothness cases as shown in Table 1, where 2D refers to smoothness term with only spatial derivatives, while 3D refers to spatio-temporal smoothness term that couples the solution of the optic flow field at different frames (also known as the "two frames" versus "multi-frame" formulation). Here σ denotes the standard deviation of the 2D Gaussian pre-filter used for pre-processing the image sequence. We used 60 iterations of the outer loop in the 3D method and 80 iterations in the 2D method, 5 inner loop iterations and 10 Gauss–Seidel iterations.

4.2 Yosemite Sequence

We applied our method to the Yosemite sequence without clouds (available at http://www.cs.brown.edu/people/black/images.html), with four resolution levels. Table 2 shows our results relative to the best published ones. As seen in the table, our method achieves a better reconstructed solution compared to all other reported results for this sequence, both for the 2D and 3D cases. In fact, our result for the 2D case is good even compared to 3D results from the literature. Figures 4 and 5 show the solution of the affine parameters from which one can observe the trends of the flow changes with respect to the x and y axes. The solution of the pure translation parameters is shown in Fig. 6. The depth discontinuities in the scene are sharp and evident in all the parameters.

Figure 1 shows the ground truth optical flow and the results with the pure translation model. Figure 2 shows the image of the angular errors. Figure 3 shows both the histogram and the cumulative probability of the angular errors. Both figures demonstrating the typically lower angular errors of our method. Table 3 summarizes the noise sensitivity results of our method. We also coupled the affine over-parameterized model with our previous work on joint optic flow computation and denoising presented in (Nir et al. 2005) by iterating between optical flow computation on the denoised sequence and denoising with the current optical flow. A related con-

cept was first introduced by Borzi et al. (2002). This coupling provides a model with robust behavior under noise, that obtains better AAE measure under all noise levels compared to the best published results.

Our improvement relative to (Brox et al. 2004) results mostly from the over-parameterization and not by the additional smoothness weight parameter ω_t. Table 4 shows the results obtained with equal spatio-temporal smoothness weights ($\omega_t = 1$) as used in Brox et al. (2004). The AAE measure changes by approximately 1 percent for the pure translation model and 2 percent for the affine model, where, the ω_t parameter was changed by approximately 21 and 45 percent for the pure translation and affine models respectively.

4.3 Synthetic Piecewise Constant Affine Flow Example

For illustration purposes we also considered the piecewise affine flow over an image of size 100×100 having the ground truth shown in Fig. 8, and given by

For $x < 40$,

$$u = -0.8 - 1.6(x - 50)/50 + 0.8(y - 50)/50,$$

$$v = 1.0 + 0.65(x - 50)/50 - 0.35(y - 50)/50.$$

For $x \geq 40$,

$$u = 0.48 - 0.36(x - 50)/50 - 0.6(y - 50)/50,$$

$$v = 0.3 - 0.75(x - 50)/50 - 0.75(y - 50)/50.$$

The two images used for this test were obtained by sampling a grid of size 100×100 from frame 8 of the Yosemite sequence (denoted I_{yosemite}). The second image $I_2(x, y) = I_{\text{yosemite}}(x + \Delta x, y + \Delta y)$ and the first image is sampled at warped locations $I_1(x, y) = I_{\text{yosemite}}(x + \Delta x + u, y +$

Table 2 Yosemite sequence without clouds

Method	AAE	STD
Papenberg et al. 2D smoothness (Papenberg et al. 2006)	1.64°	1.43°
Brox et al. 2D smoothness (Brox et al. 2004)	1.59°	1.39°
Mémin and Pérez (2002)	1.58°	1.21°
Roth and Black (2005)	1.47°	1.54°
Bruhn et al. (2005)	1.46°	1.50°
Amiaz et al. 2D smoothness (Over-fine x4) (Amiaz et al. 2007)	1.44°	1.55°
Farnebäck (2000)	1.40°	2.57°
Liu et al. (2003)	1.39°	2.83°
Our method affine 2D smoothness	1.18°	1.31°
Govidu (2006)	1.16°	1.17°
Farnebäck (2001)	1.14°	2.14°
Our method constant motion 3D smoothness	1.07°	1.21°
Papenberg et al. 3D smoothness (Papenberg et al. 2006)	0.99°	1.17°
Brox et al. 3D smoothness (Brox et al. 2004)	0.98°	1.17°
Our method rigid motion 3D smoothness	0.96°	1.25°
Our method affine 3D smoothness	0.91°	1.18°
Our method pure translation 3D smoothness	0.85°	1.18°

Fig. 1 Optic flow of the Yosemite sequence. *Left*—ground truth. *Right*—our method with the pure translation model
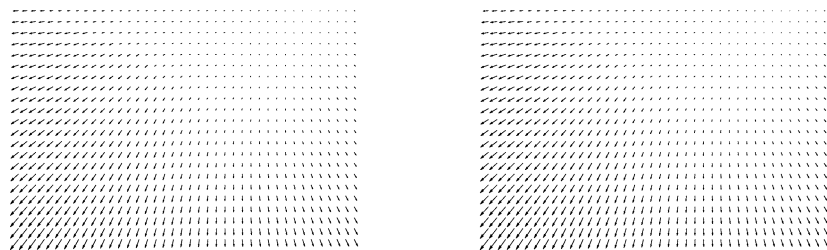
Fig. 2 Images of the angular error, white indicates zero error and black indicates an error of 3 degrees or above. *Left*—our method with the affine model. *Middle*—our method with the pure translation model. *Right*—the method of (Brox et al. 2004), optic flow field courtesy of the authors
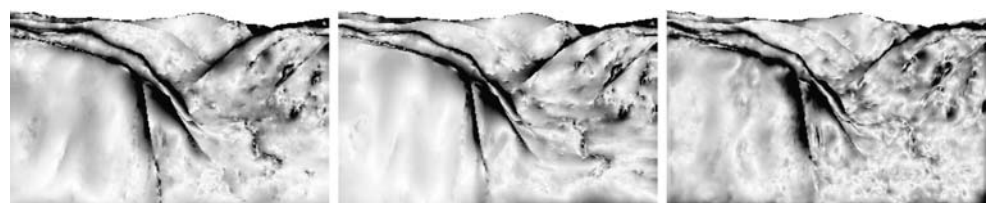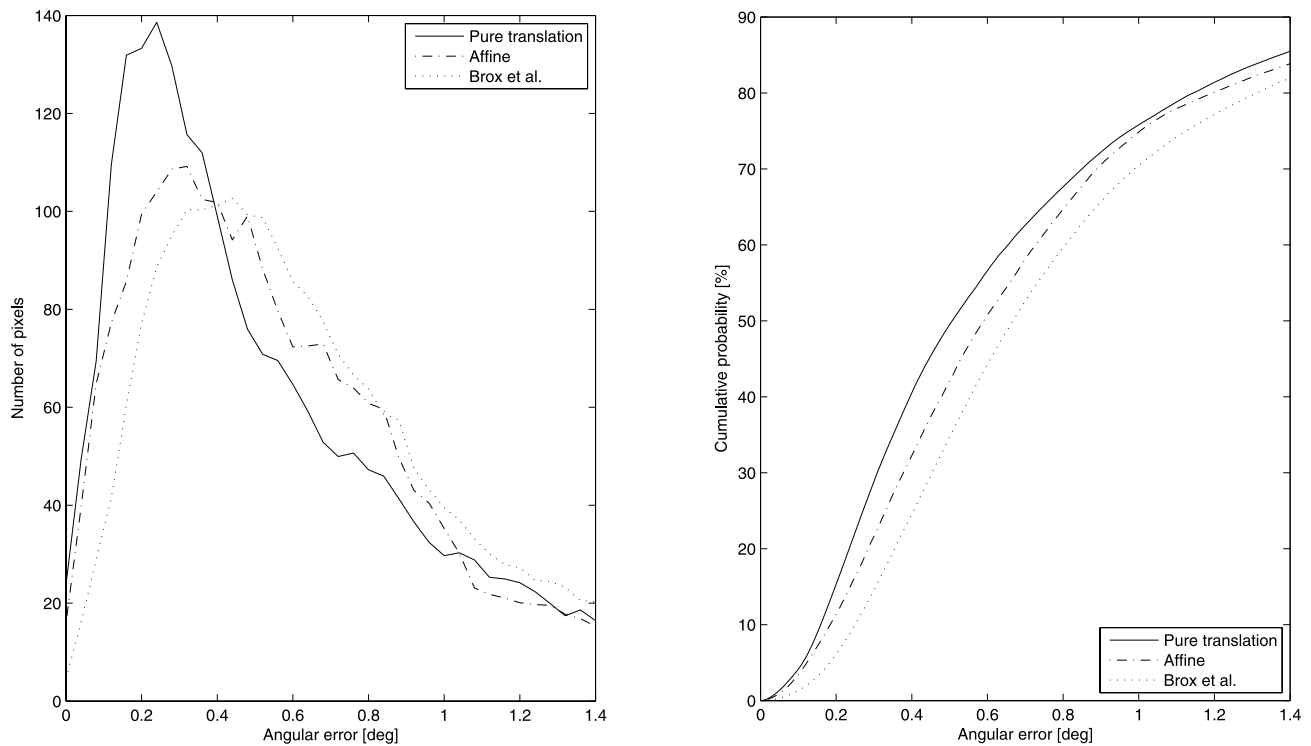
Fig. 3 Histogram (*left*) and Cumulative probability (*right*) of the angular errors. *Solid*—pure translation model. *Dash dot*—affine model. *Dotted*—optic flow obtained in (Brox et al. 2004) courtesy of the authors

Table 3 Yosemite without clouds—noise sensitivity results presented as $AAE \pm STD$

σ_n	Our method (affine)	Our method (affine) coupled with (Nir et al. 2005)	Results reported in (Brox et al. 2004)
0	$0.91 \pm 1.18°$	$0.93 \pm 1.20°$	$0.98 \pm 1.17°$
20	$1.59 \pm 1.67°$	$1.52 \pm 1.48°$	$1.63 \pm 1.39°$
40	$2.45 \pm 2.29°$	$2.02 \pm 1.76°$	$2.40 \pm 1.71°$

Table 4 Yosemite sequence without clouds—results obtained by using equal spatio-temporal smoothness weights ($\omega_t = 1$)

Method	AAE	STD
Our method affine 3D smoothness	0.93°	1.20°
Our method pure translation 3D smoothness	0.86°	1.18°

$\Delta y + v$) using bilinear interpolation. The constant shift values are: $\Delta x = 79$ and $\Delta y = 69$. The two images obtained are displayed in Fig. 7. The results exhibited in Table 5 show that our method with the affine over-parametrization outperforms the method of (Brox et al. 2004). This is to be expected since the true flow is not piecewise constant and the smoothness term in (Brox et al. 2004) penalizes for changes from the constant flow model, whereas, the affine over-parametrization model solves the optimization problem in the (correct) affine space in which it accurately finds the piecewise constant affine parameters solution of the problem, as shown in Figs. 9 and 10. One can notice that the discontinuity at pixel $x = 40$ is very well preserved due to the effective edge preserving L_1 based optimization. The re-

sulting optical flow shown in Fig. 11 accurately matches the ground truth.

4.4 Flower Garden Sequence

We also applied our method to the real image sequence "flower garden". The results obtained by the 2D affine model are shown in Fig. 12. We have checked our results by manually marking the coordinates of several corresponding points on the two images and comparing the motion with the optical flow results. We have found good match between manual tracking and the computed flow. The tree moves with a velocity of about -5.7 pixel/frame along the x direction.

The whole scene has low velocity in the y direction (the computed v is between -0.9 and 1.4 pixel/frame). The y component of the flow is about 1 pixel/frame in the upper section of the tree and almost zero in the lower section. In the area of the garden, the velocity is decreasing as the distance from the camera increases from about -3 pixel/frame in the
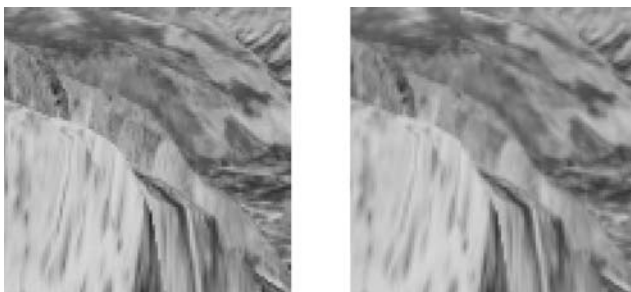
Fig. 4 Solution of the affine parameters for the Yosemite sequence—from left to right A_1, A_2, A_3
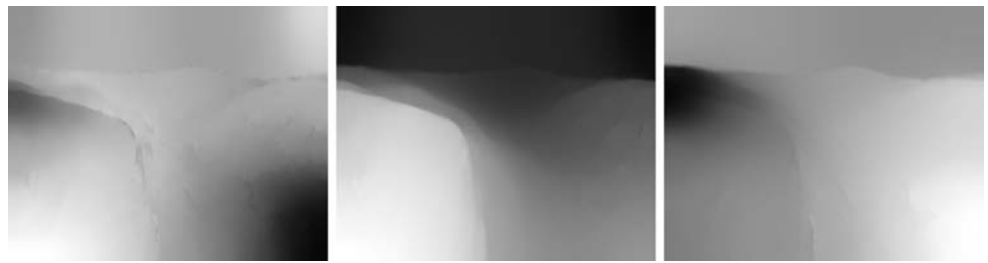
Fig. 5 Solution of the affine parameters for the Yosemite sequence—from left to right A_4, A_5, A_6

Fig. 6 Solution of the pure translation parameters for the Yosemite sequence—from left to right A_1, A_2, A_3
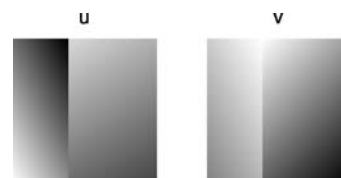

Fig. 7 The two images: *Left*—first frame. *Right*—second frame

Table 5 Synthetic piecewise affine flow test

Method	AAE	STD
Our implementation of the method of (Brox et al. 2004)	1.48°	2.28
Our method with the affine flow model	0.88°	1.67

Fig. 8 Synthetic piecewise affine flow—ground truth. *Left*—u component. *Right*—v component


**Fig. 9** Solution of the affine parameters—from left to right A_1, A_2, A_3

Fig. 10 Solution of the affine parameters—from left to right A_4, A_5, A_6

right lower section (-2 on the left) to about -1 pixel/frame in the upper part and in the area of the houses. The computed flow by the pure translation over-parameterized model shown in Fig. 13 produces similar flow field as the results with the affine model. The solution of the three pure translation model parameters is shown in Fig. 14.

Fig. 11 Ground truth for the synthetic test (*left*) and computed flow by the affine over-parametrization model (*right*)
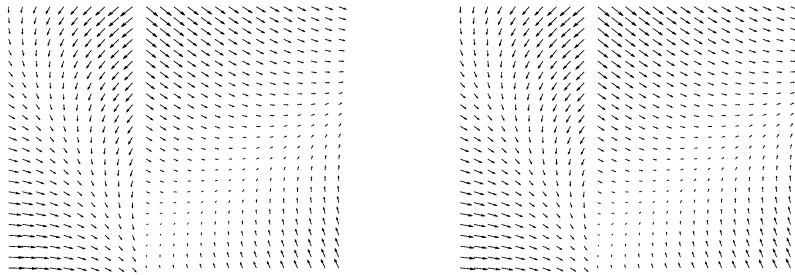
Fig. 12 Flower garden sequence: First image (*left*). Optical flow computed by the 2D affine over-parametrization model, *u* (*middle*) and *v* (*right*)
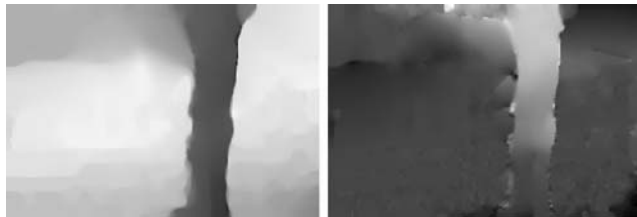
Fig. 13 Flower garden sequence: optical flow computed by the 2D pure translation over-parametrization model, *u* (*left*) and *v* (*right*)

4.5 Road Sequence

The two frames of this real sequence courtesy of the authors of (Vázquez et al. 2006) are shown in Fig. 15. The computed flow by our 2D over-parameterized affine model is shown in Fig. 16. The motion compensated image difference is shown in Fig. 17, our motion compensated reconstruction ratio is $PSNR = 38.17$ dB (more than 1.13 dB better than the results reported in (Vázquez et al. 2006) for the affine model and 0.58 dB better than their quadratic model). The measure of the motion compensated difference might be misleading for comparing the quality of different optical flow algorithms since one can generate a globally optimal algorithm for this measure which produces meaningless optical flow results. For example: suppose we find for each pixel in one image a corresponding pixel with the same gray value in the other image. In this case, the resulting optical flow is perfect in terms of $PSNR$, however, it can be arbitrary vectors connecting pixels with the same gray level belonging to different objects in the scene. For this example there is no ground truth motion and therefore we use this measure in order to compare with the results reported in (Vázquez et al. 2006).

5 Summary

We introduced a novel over-parameterized variational framework for accurately solving the optical flow problem. The flow field is represented by a general space-time model. The proposed approach is useful and highly flexible in the fact that each pixel has the freedom to choose its own set of model parameters. Subsequently, the decision on the discontinuity locations of the model parameters, is resolved within the variational framework for each sequence. Many variational approaches can be regarded as special cases of our method when one selects a constant motion model. Observe however that in most scenarios, the optical flow would be better represented by a piecewise constant affine model or a rigid motion model rather than a piecewise constant flow. Therefore, compared to existing variational techniques, the smoothness penalty term modeled by the proposed over-parametrization models yields better optic flow recovery performance as demonstrated by our experiments. In this paper, the derivation is for general space-time basis functions. We limited our experiments to spatial basis functions. Future work will focus on experiments with spatio-temporal basis functions. Incorporating learning of the basis functions (dictionaries) for specific scenes could be of great interest and useful for video compression. As a consequence of this work, motion segmentation based on optical flow should generally be replaced by segmentation in the higher dimensional parameter space as suggested by our initial results presented herein for the synthetic sequence. Although the models suggested in this paper were all over-parameterized, an under-parameterized model might also be used in this framework, for example in case one has prior knowledge regarding constraints between the *u* and *v* components (as in stereo matching or when we know the optic flow to be radial).
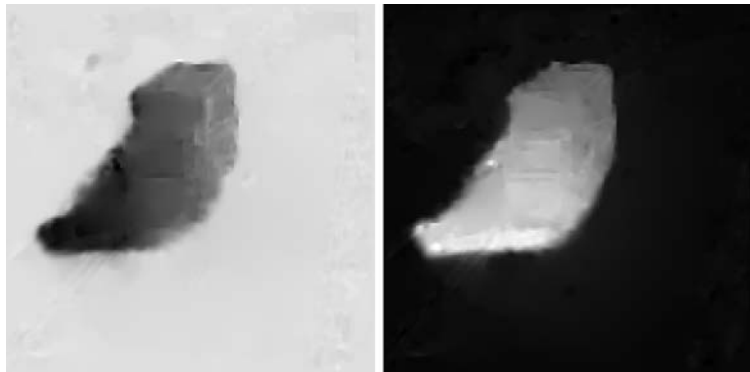
Fig. 14 Flower garden sequence: the solution of the pure translation parameter. From left to right *A*1, *A*2, *A*3


Fig. 15 Road sequence. Frame 1—*Left*. Frame 2—*Right*

Fig. 16 Optical flow of the road sequence. *u* is shown on the *left* and *v* on the *right*

Fig. 17 Difference of the frames. *Left*—difference of the original frames (*PSNR* = 23.73 dB). *Right*—motion compensated difference by the estimated optical flow (*PSNR* = 38.17 dB). Both images are scaled by the same linear transformation from errors to graylevels
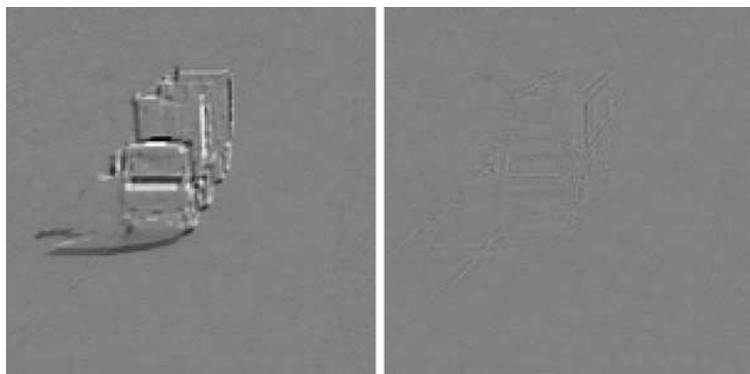
Table 6 Yosemite sequence without clouds—results obtained by our method when using bicubic image interpolation instead of bilinear

Method	AAe	STD
Our method affine 2D smoothness	1.15°	1.33°
Our method affine 3D smoothness	0.86°	1.17°
Our method pure translation 3D smoothness	0.82°	1.17°

Note added in proof

We further experimented with replacing the bilinear image interpolation with bicubic interpolation which consistently improved the results as can be seen in Table 6.

References

Adiv, G. (1985). Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *7*(4), 384–401.

Amiaz, T., & Kiryati, N. (2005). Dense discontinuous optical flow via contour-based segmentation. In *Proceedings of the international conference on image processing* (Vol. 3, pp. 1264–1267).

Amiaz, T., Lubetzky, E., & Kiryati, N. (2007). Coarse to over-fine optical flow estimation. *Pattern Recognition 40*(9), 2496–2503.

Ari, R. B., & Sochen, N. (2006). A general framework and new alignment criterion for dense optical flow. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (Vol. 1, pp. 529–536).

Barron, J. L., Fleet, D. J., & Beauchemin, S. S. (1994). Performance of optical flow techniques. *International Journal of Computer Vision*, *12*(1), 43–77.

Black, M., & Anandan, P. (1991). Robust dynamic motion estimation over time. In *Proceedings of the computer vision and pattern recognition* (pp. 292–302). Los Alamitos: IEEE Computer Society.

Black, M. J., & Anandan, P. (1996). The robust estimation of multiple motions: parametric and piecewise smooth flow fields. *Computer Vision and Image Understanding*, *63*(1), 75–104.

Borzi, A., Ito, K., & Kunisch, K. (2002). Optimal control formulation for determining optical flow. *SIAM Journal on Scientific Computing*, *24*(3), 818–847.

Brox, T., Bruhn, A., Papenberg, N., & Weickert, J. (2004). High accuracy optical flow estimation based on a theory for warping. In T. Pajdla & J. Matas (Eds.), *Lecture notes in computer science*: *Vol. 3024. Proceedings of the ECCV 2004* (pp. 25–36), Prague, Czech Republic. Berlin: Springer.

Brox, T., Bruhn, A., & Weickert, J. (2006). Variational motion segmentation with level sets. In: *Lecture notes in computer science: Vol. 3951. ECCV 2006, Part 1* (pp. 471–483).

Bruhn, A., Weickert, J., & Schnörr, C. (2005). Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods. *International Journal of Computer Vision*, *61*, 211–231.

Cremers, D., & Soatto, S. (2005). Motion competition: a variational approach to piecewise parametric motion segmentation. *International Journal of Computer Vision*, *62*(3), 249–265.

Deriche, R., Kornprobst, P., & Aubert, G. (1995). Optical flow estimation while preserving its discontinuities: a variational approach. In *Proceedings of the second Asian conference on computer vision* (Vol. 2, pp. 290–295).

Farnebäck, G. (2000). Fast and accurate motion estimation using orientation tensors and parametric motion models. In *Proceedings of the 15th international conference on pattern recognition* (Vol. 1, pp. 135–139).

Farnebäck, G. (2001). Very high accuracy velocity estimation using orientation tensors, parametric motion and simultaneous segmentation of the motion field. In *Proceedings of the 8th international conference on computer vision* (Vol. 1, pp. 171–177). Los Alamitos: IEEE Computer Society.

Govidu, V. M. (2006). Revisiting the brightness constraint: probabilistic formulation and algorithms. In *Lecture notes in computer science*: *Vol. 3953. Proceedings of the ECCV 2006, Part 3* (pp. 177–188). Berlin: Springer.

Horn, B. K. P., & Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, *17*, 185–203.

Ju, S., Black, M., & Jepson, A. (1996). Skin and bones: multi-layer, locally affine, optical flow and regularization with transparency. In *Proceedings of the computer vision and pattern recognition* (pp. 307–314).

Irani, M., Rousso, B., & Peleg, S. (1993). Robust recovery of egomotion. In *Computer analysis of images and patterns* (pp. 371–378).

Irani, M., Rousso, B., & Peleg, S. (1997). Recovery of ego-motion using region alignment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(3), 268–272.

Liu, H., Chellappa, R., & Rosenfeld, A. (2003). Accurate dense optical flow estimation using adaptive structure tensors and a parametric model. *IEEE Transactions on Image Processing*, *12*, 1170–1180.

Lucas, B., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th international joint conference on artificial intelligence* (pp. 674–679).

Mémin, E., & Pérez, P. (2002). Hierarchical estimation and segmentation of dense motion fields. *International Journal of Computer Vision*, *46*(2), 129–155.

Nagel, H. H. (1990). Extending the 'oriented smoothness constraint' into the temporal domain and the estimation of derivatives of optical flow. In *Lecture notes in computer science*: *Vol. 427. Proceedings of the ECCV* (pp. 139–148). Berlin: Springer.

Nagel, H. H., & Enkelmann, W. (1986). An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *8*, 565–593.

Nir, T., Kimmel, R., & Bruckstein, A. M. (2005). *Variational approach for joint optic-flow computation and video restoration*. CIS-2005-03 report, Technion.

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

Roth, S., & Black, M. J. (2005). On the spatial statistics of optical flow. In *Proceedings of the tenth international conference on computer vision (ICCV'05)* (*Vol*. 1, pp. 42–49), Washington, DC, USA. Los Alamitos: IEEE Computer Society.

Sekkati, H., & Mitiche, A. (2003). Dense 3d interpretation of image sequences: a variational approach using anisotropic diffusion. In *Proceedings of the ICIAP-03*.

Sekkati, H., & Mitiche, A. (2006). Joint optical flow estimation, segmentation, and 3d interpretation with level sets. *Computer Vision and Image Understanding*, *103*(2), 89–100.

Szeliski, R., & Coughlan, J. (1997). Hierarchical spline-based image registration. *International Journal of Computer Vision*, *22*(3), 199–218.

Vázquez, C., Mitiche, A., & Laganière, R. (2006). Joint multiregion segmentation and parametric estimation of image motion by basis function representation and level set evolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(5), 782–793.

Weickert, J., & Schnörr, C. (2001). Variational optic flow computation with a spatio-temporal smoothness constraint. *Journal of Mathematical Imaging and Vision*, *14*(3), 245–255.